CHAPTER 3

ROUGH SET, FUZZY SET, CBRS & SOFT SET

3.1 Introduction

In computer science, a ***rough set***, first described by a Polish Computer Scientist Zdzisław Pawlak, is a formal approximation of a crisp set (i.e., conventional set) in terms of a pair of sets which give the lower and the upper approximation of the original set. In the standard version of rough set theory (Pawlak 1982, 1991, 1996)[103,104 and 105], the lower and upper approximation sets are crisp sets, but in other variations, the approximating sets may be fuzzy sets. The main goal of the rough set analysis is induction of (learning) approximations of concepts. A rough set constitutes a sound basis for KDD. It offers mathematical tools to discover patterns hidden in data. It can be used for feature selection, feature extraction, data reduction, decision rule generation, and pattern extraction (templates, association rules) etc. Identifies partial or total dependencies in data, eliminates redundant data, and gives approach to null values, missing data, dynamic data and others.

In 1999, (Molodtsov 1999)[90] proposed ***soft set*** theory as a new mathematical tool for dealing with vagueness and uncertainties. At present, work on the soft set theory is progressing rapidly and many important theoretical models have been presented, such as soft groups (Aktas H et.al 2007)[4], softrings (Acar, U 2010)[3], soft semirings (Feng, F et.al 2008)[29], soft ordered semigroup (Jun, Y.B et.al 2010)[62] and exclusive disjunctive soft sets (Xiao, Z et.al 2010)[157]. The research on fuzzy soft set has also received much attention since its introduction by (P.K Maji et al. 2001)[98]. Several extension models including intuitionistic fuzzy soft sets (Maji, P.K et.al 2002)[99], interval-valued fuzzy soft sets (Yang, X.B et.al 2009)[160] and interval-valued intuitionistic fuzzy soft set (Jiang, Y et. Al 2010)[60] are proposed in succession. At the same time, researchers have also successfully applied soft sets to deal with some practical problems, such as decision making (Feng, F et.al 2010)[30], economy forecasting (Xiao, Z et.al 2009)[156], maximal association rules mining (Herawan, T., Mat Deris, M 2011)[50], etc.

The soft sets mentioned above, either in theoretical study or practical applications are based on complete information. However, incomplete information widely exists in practical problems. For example, an applicant perhaps misses age when he/she fills out an application form. Missing or unclear data often appear in questionnaire due to the fact that attendees give up some questions or cannot understand the meaning of questions well. In addition, other reasons like mistakes in the process of measuring and collecting data, restriction of data collecting also can cause unknown or missing data. Hence, soft sets under incomplete information become incomplete soft sets. In order to handle incomplete soft sets, new data processing methods are required.

In mathematics, ***fuzzy sets*** are sets whose elements have degrees of membership. Fuzzy sets were introduced by (Lotfi A. Zadeh and Dieter Klaua in 1965)[72] as an extension of the classical notion of set. At the same time, (Salii V.N 1965)[131] defined a more general kind of structure called an *L*-relation, which he studied in an abstract algebraic context. Fuzzy relations, which are used now in different areas, such as linguistics (De Cock, Bodenhofer &

Kerre 2000)[18] decision-making (Kuzmin 1982)[71] and clustering (Bezdek 1978)[11], are special cases of *L*-relations when *L* is the unit interval [0, 1].

In classical set theory, the membership of elements in a set is assessed in binary terms according to a bivalent condition an element either belongs or does not belong to the set. By contrast, fuzzy set theory permits the gradual assessment of the membership of elements in a set; this is described with the aid of a membership function valued in the real unit interval [0, 1]. Fuzzy sets generalize classical sets, since the indicator functions of classical sets are special cases of the membership functions of fuzzy sets, if the latter only take values 0 or 1. In fuzzy set theory, classical bivalent sets are usually called *crisp* sets. The fuzzy set theory can be used in a wide range of domains in which information is incomplete or imprecise, such as bioinformatics.

3.2 Rough Set Concepts and Notations

The concept of rough set is another approach to deal with imperfect knowledge. It was introduced by Z. Pawlak in 1982 (Pawlak 1982)[103]. From a philosophical point of view, rough set theory is a new approach to deal vagueness and uncertainty, and from a practical point of view, it is a new method of data analysis.

This method has the following important advantages:

- ✓ It is easy to understand.
- ✓ It finds reduced set of data (data reduction).
- ✓ It evaluates significance of data.
- ✓ It generates minimal set of decision rules from data.
- ✓ It provides efficient algorithms for finding hidden patterns in data.
- ✓ It offers straightforward interpretation of results.
- ✓ It can be used in both qualitative and quantitative data analysis and
- ✓ It identifies relationship that would not be found by using statistical methods.

Let $R \subseteq UxU$ denote an equivalence relation on U, that is, R is a reflexive, symmetric and transitive relation. The equivalence class of an element $x \in U$ with respect to R is the set of elements $y \in U$ such that xRy. If two elements x, y in U belong to the same equivalence class then the x and y are indistinguishable with respect to relation R.

Given an arbitrary set $A \subseteq U$ it may not be possible to describe 'A' precisely in the approximation space apr (R) = (U, R) instead one may only characterize 'A' by a pair of lower and upper approximations. This leads to the concept of rough sets. Define,

$$\underline{R}A = \bigcup \{ Y \in U/R : Y \subseteq A \};$$

and

$$\overline{R}A = \bigcup \{ Y \in U/R : Y \cap A \neq \varphi \}.$$

$\underline{R}A$ and $\overline{R}A$ are respectively called the R-lower and R-upper approximation of A with respect to R.

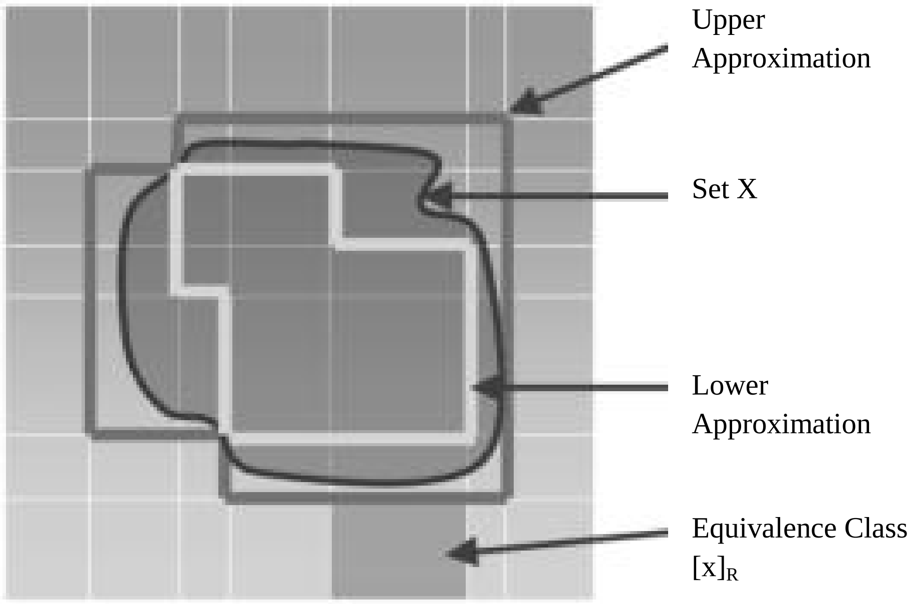

Fig: 3.1. Rough Set Approach

It can be noted that

$$\underline{R}A=\left\{x \in U:\left[x_R\right] \subseteq A\right\} \text{ and}$$

$$\overline{R}A=\left\{x \in U:\left[x\right]_R \cap X \neq \varphi\right\}.$$

The set $BN_R\left(A\right)=\overline{R}A - \underline{R}A$ is called the *R-boundary* of A. The set $\underline{R}A$ consists of all those elements of U which can with certainty be classified as elements of A, employing the knowledge R. The set $\overline{R}A$ consists of all those elements of U which can possibly be classified as elements of A, employing the knowledge R. Set $BN_R(A)$ is the set of elements which cannot be classified as either belonging to A or belonging to –A having the knowledge R, say that a set A is *R-definable* if and only if $\underline{R}A=\overline{R}A$. Otherwise A is said to be *R-rough*.

The borderline region is the undecidable area of the universe. Say that *X* is *rough* with respect to *R* if and only if $\underline{R}X \neq \overline{R}X$, equivalently $BN_R(X) \neq$ ¤ . *X* is said to be R-*definable* if and only if $\underline{R}X = \overline{R}X$, or $BN_R(X) =$ ¤ .

Let U ($\neq$ ¤) be a finite set of objects, called the universe and R be an equivalent relation over U . By U/R we denote the family of equivalence classes of R (or classification of U) referred to as categories or concepts of R and $\left[X\right]_R$ denotes a category is R containing an element $X \in U$ By a Knowledge base, we understand a relation system $K=(U,\Re)$, where U is as above and $\Re$ is a family of equivalence relations over U .

For any subset P($\neq$ ¤) $\subseteq R$, the intersection of all equivalence relations in P is denoted by $IND(P)$ and is called the *indiscernibility relation over* P . The equivalence classes of

$IND(P)$ are called $P-$*basic knowledge* about U in K. For any $Q \in R$, Q is called a $Q-$elementary knowledge about U in K and equivalence classes of Q are called $Q-$*elementary concept of knowledge* $\Re$. The family of $P-$basic categories for all ¤ $\neq P \subseteq R$ will be called the *family of basic categories* in knowledge base K. By $IND(K)$, we denote the family of all equivalence relations defined in K. Symbolically, $IND(K) = \{IND(P) : ¤ \neq P \subseteq \Re\}$.

For any $X \subseteq U$ and an equivalence relation $R \in IND(K)$, we associate two subsets. $\underline{R}X = U\{Y \in U / R : Y \subseteq X\}$ and $\overline{R}X = U\{Y \in U / R : Y \bigcap X \neq ¤\}$, are called the $R-$*lower* and $R-$*upper approximations* of X respectively.

The $R-$*boundary* of X is denoted by $BN_R(X)$ and is given by $BN_R(X) = \overline{R}X - \underline{R}X$. The elements of $\underline{R}X$ are those elements of U which can be certainly classified as elements of X with the knowledge of R and $\overline{R}X$ is the set elements of X which can be possibly classified as elements of X employing knowledge of R. The borderline region is the un decidable area of the universe.

We say X is *rough* with respect to R if and only if $\underline{R}X \neq \overline{R}X$, equivalently $BN_R(X) \neq ¤$. X is said to be $R-$*definable* if and only if $\underline{R}X = \overline{R}X$, or $BN_R(X) = ¤$. So, a set is rough with respect to R if and only if it is not $R-$definable.

The set notion for both crisp and rough can now be defined as:

- A set X is called crisp (exact) with respect to R if and only if the boundary region of X is empty.
- A set X is called rough (inexact) with respect to R if and only if the boundary region of X is nonempty.

The definitions of set approximations presented above can be expressed in terms of granules of knowledge in the following way. The lower approximation of a set is union of all granules which are entirely included in the set; the upper approximation – is union of all granules which have non-empty intersection with the set; the boundary region of a set is the difference between the upper and the lower approximation of the set.

It is interesting to compare definitions of classical sets, fuzzy sets and rough sets. Classical set is a primitive notion and is defined intuitively or axiomatically. Fuzzy sets (Zadeh LA, 1965)[72] are defined by employing the fuzzy membership function, which involves advanced mathematical structures, numbers and functions. Rough sets are defined by approximations. Thus this definition also requires advanced mathematical concepts.

One can easily prove the following properties of approximations:

1. $\underline{R}(X) \subseteq X \subseteq \overline{R}(X)$
2. $\underline{R}(¤) = \overline{R}(¤) = (¤)$; $\underline{R}(U) = \overline{R}(U) = U$
3. $\overline{R}(X \cup Y) = \overline{R}(X) \cup \overline{R}(Y)$

4. $\underline{R}(X \cap Y) = \underline{R}(X) \cap \underline{R}(Y)$
5. $\underline{R}(X \cup Y) \supseteq \underline{R}X \cup \underline{R}Y$
6. $\overline{R}(X \cap Y) = \overline{R}X \cap \overline{R}Y$
7. $X \subseteq Y \rightarrow \underline{R}(X) \subseteq \underline{R}(Y)$ & $\overline{R}(X) \subseteq \overline{R}(Y)$
8. $\underline{R}(-X) = -\overline{R}(X)$
9. $\overline{R}(-X) = -\underline{R}(X)$
10. $\underline{R}\underline{R}X = \overline{R}\underline{R}X = \underline{R}X$
11. $\overline{R}\overline{R}X = \underline{R}\overline{R}X = \overline{R}X$

It is easily seen that the lower and the upper approximations of a set are, respectively the interior and closure of this set in the topology generated by the indiscernibility relation.

One can define the following four basic classes of rough sets, i.e., four categories of vagueness:

SN	Conditions	Classes of X
1	If $\underline{R}X \neq$ ¤ and $\overline{R}X \neq U$	Roughly R-definable
2	If $\underline{R}X =$ ¤ and $\overline{R}X \neq U$	Internally R-undefinable.
3	$\underline{R}X \neq$ ¤ and $\overline{R}X = U$	Externally R-undefinable
4	$\underline{R}X \neq$ ¤ and $\overline{R}X = U$	Totally R-undefinable

Table 3.1 Basic Classes of Rough Sets

The intuitive meaning of this classification is the following:

- A set X is roughly R-definable means that with respect to R able to decide for some elements of U that they belong to X and for some elements of U that they belong to –X.
- A set X is internally R-undefinable means with respect to R able to decide for some elements of U that they belong to –X, but unable to decide for any element of U whether it belongs to X.
- A set X is externally R-undefinable means that with respect to R able to decide for some elements of U that they belong to X, but we are unable to decide for any element of U whether it belongs to –X.
- A set X is totally R-undefinable means that with respect to R unable to decide for any element of U whether it belongs to X or –X.

A rough set X can be also characterized numerically by the following coefficient

$$\alpha_R(X) = \frac{|R_*(X)|}{|R^*(X)|}$$

called the accuracy of approximation, where |X| denotes the cardinality of $X \neq \emptyset$

Obviously $0 \leq \alpha_R(X) \leq 1$. If $\alpha_R(X) = 1$ then X is crisp with respect to R (X is precise with respect to R), and otherwise, if $\alpha_R(X) < 1$, X is rough with respect to R (X is vague with respect to R).

3.3 Covering Based Rough Set

There is a strong similarity between rough sets and covering based sets. However the origin of uncertainty is radically different. In the case of rough sets, attribute values are known but there are not enough attributes (or attribute domains are too coarse) to single out objects by their descriptions using these attributes. In the case of covering based sets there may be enough attribute values, but the lack of knowledge about objects forbids a precise numeration of the contents of sets defined by means of properties. It is clear that these sources of uncertainty being unrelated, they can be simultaneously present. This section describes this hybrid situation. It is shown here that imprecise observations of the attribute can be described, in a natural way, by coverings. To study direct covering-based extensions of rough sets proposed by (Zhu.W 2007)[163 and 164].

A cover is a generalization of the notion of partition. Using covers instead of partitions, covering based rough sets have been introduced by (Zakowski 1983)[162]. The covering based rough sets are models with promising potential for applications to data mining. There is only one lower approximation for such rough sets. However, there are as many as four different versions for the definition of upper approximation of such sets. Also, this has led to comparison of the different types of rough sets thus generated. Several properties of the different types of covering based rough sets have been derived by different researchers.

3.3.1 Definition and Notations

(William Zhu 2007)[75] gave the definitions of three other types of covering based rough sets which are given as below. First he said that the descriptions of lower and upper approximations of such sets could be done two ways. One way of description includes only essential characteristics related to class of elements and he called it as minimal description (*md*). Another way of description includes all characteristics related to class of elements and he called it as maximal description (*MD*). The union of minimal descriptions of element $x \in U$ is called close friends of element *x*. If a covering *C* of *U* contains only one minimal description $\forall x \in U$, then it will be called as unary covering.If a covering *C* of *U* contains covers of *x* which are subset of close friends of *x*, then it is called as point wise-covered covering.

The definitions of three types of covering based rough sets are given by keeping the lower approximation the same for all the three types and varying only the upper approximation for each type.

Definition 1: Let U be a universe of discourse and C be a family of subsets of U. C is called a cover of U if no subset in C is empty and UC = U. (U, C) the covering approximation space and the covering C is called the family of approximation sets.

Definition 2: Let (*U*, *C*) be an approximation space and x be any element of U. Then the Family Md *(x) = {K∈C: x∈K} S∈C* is called the minimal description of the object x.

Definition 3: For any set *X*ċ *U*, the family of sets *C*(X) = {K ∈ C : K* ċ *X}* is called the family of sets bottom approximating the set X.

Definition 4: The set *X** = *C* (X)* is called the lower approximation of the set X.

In second type of covering based rough set, if covering is replaced with partition then it will become traditional or basic rough set introduced by Pawlak. In covering based rough sets duality property never holds true whereas it holds true only in traditional rough set defined by Pawlak. Actually duality is the property that specifies the relationship between lower and upper approximation of the rough set. This relationship facilitates that if one of them is known then another one will be obtained from the known one. Three more covering based rough sets were also defined further. In this research work covering based rough sets in the context of multi granulation set up given by Liu are considered and used.

Definition 5: Let U be a universe of discourse and C be a family of subsets of U. C is called a *cover* of U if no subset in C is empty and $\cup$ C = U. We call (U, C) the *covering approximation space* and the covering C is called the *family of approximation sets*.

Definition 6: Let (U, C) be an approximation space and x be any element of U. Then the family Md (x) = $\{K \in C: x \in K \wedge \forall S \in C(x \in S \wedge S \subseteq K \Rightarrow K = S)\}$

Definition 7: For any set X $\subseteq$ U, the family of sets *of sets bottom approximating the set* X.

Definition 8: Let $< U, C >$ be a covering approximation space. For a set $X \subseteq U$ *the set* $X_4 = X_* \cup \{K / K \in C \text{ and } K \cap (X - X_*) \neq$ is called the *fourth type of covering upper approximation of* X.

3.3.2 Application of Covering based Rough sets

Classifications of universes play a central role in the study of basic rough set theory. Extending the concept of approximation of sets, (Grzymala-Busse, J. 1988)[39] introduced the approximation of classifications. As pointed out by (Pawlak 1982)[103] these results of Busse establish that the two concepts, approximation of sets and approximation of families of sets (or classifications) are two different issues and that the equivalence classes of approximate classifications cannot be arbitrary sets. He has further stated that if positive examples of each category in the approximate classification then must also have negative examples of each category below a classification on a universe formally.

(Grzymala-Busse, J. 1988)[39] has established some properties of approximation of classifications in the form of four theorems. These results are irreversible by nature. However, these results have been further extended to obtain theorems establishing necessary and sufficient type of properties of approximation of classifications. From these general types of results, many interesting results have been obtained as corollaries, besides the results of Busse.

3.4 Concept of Soft Set

(Molodtsov 1991)[90] introduced the concept of soft set theory, which can be used as a generic mathematical tool for dealing with uncertainty.

Soft set is a parameterized general mathematical tool which deals with a collection of approximate descriptions of objects. Each approximate description has two parts, a predicate and an approximate value set. In classical mathematics, a mathematical model of an object is constructed and defines the notion of exact solution of this model. Usually the mathematical model is too complicated and the exact solution is not easily obtained. So, the notion of approximate solution is introduced and the solution is calculated. In the soft set theory, the opposite approach to this problem. The initial description of the object has an approximate nature, and do not need to introduce the notion of exact solution. The absence of any restrictions on the approximate description in soft set theory makes this theory very convenient and easily applicable in practice. Any parameterization can be used with the help of words and sentences, real numbers, functions, mappings and so on.

Soft sets (P.K. Maji, et.al 2002)[99] was firstly proposed by the idea of reduct and decision making using soft set theory the application of soft set theory to a decision making problem with the help of Pawlak's rough mathematics was presented. The reduction approach presented using Pawlak's rough reduction and a decision can be selected based on the maximal weighted value among objects related to the parameters.

(Zou 2008)[165] proposed a new technique for decision making of soft set theory under incomplete information systems. The idea is based on the calculation of weighted-average of all possible choice values of object and the weight of each possible choice value is decided by the distribution of other objects.

3.4.1 Definition and Notations

Definition 1: A pair (F,A) is called a soft set over U where F is a mapping given by

$$F : A \rightarrow P(U)$$

In other words, a soft set over U is a parameterized family of subsets of the universe U. For $X \in A$, F(X) may be considered as the set of X-elements of the soft set (F,A), or as the set of X-approximate elements of the soft set.

Definition 2: Let S = (U, A, V, f) be an information system and let B be any subset of A. Two elements $x, y \in U$ are said to be B-indiscernible (indiscernible by the set of attribute $B \in A$ in S) if and only if f (x, a) = f (y, a), for every $a \in B$.

Definition 3: Let S = (U, A, V, f) be an information system and let B be any subset of A. A rough approximation of a subset $X \in U$ with respect to B is defined as a pair of lower and upper approximations of X, i.e.

$$\{\underline{B}(X), \overline{B}(X)\}$$

Definition 4: Let S = (U, A,V, f) be an information system and let B be any subsets of A in information system S. Attribute $b \in B$ is called dispensable if

$$U/(B-\{b\}) = U/B$$

3.4.2 Applications of Soft Set

Soft set theory has potential applications in many different fields which include the smoothness of functions, game theory, operations research, Riemann integration, Peron integration, and probability theories, and measurement theory, see (H. Yang 2011)[48] for details. In (H. Yang 2011)[48] applications of soft set to stability and regularization, game theory and operational research are discussed in details. Extension of soft set theory to real analysis and its applications were also presented. Also, application of soft set theory to the problems of medical diagnosis in medical expert system was discussed.

Applications of soft set theory in other disciplines and real life problems are now catching momentum. (P.K. Maji, et.al 2002)[99] in year 2002, gave first practical application of soft sets in decision making problems. It is based on the notion of knowledge of reduction in rough set theory.

(N. Cagman and S. Enginoglu 2011)[92] defined soft matrices and their operations to construct a soft max-min decision making method which can be successfully applied to the problems that contain uncertainties.

(T. Herawanet al. 2010)[147] gave an alternative approach for attribute reduction in multi-valued information system under soft set theory. The work emphasized that based on the notion of multi-soft sets and AND operation, attribute reduction can be defined. It is shown that the reducts obtained are equivalent with Pawlak's rough reduction.

3.5 Fuzzy Sets

Fuzzy sets, introduced by (Zadeh LA, 1965)[72] have better modeling power than crisp sets. The graded membership assigned to elements provides additional modeling power to fuzzy sets.

3.5.1 Definition and Notations

Definition 1. A fuzzy set A defined over an universal set X is defined through its membership function μ_A as $\mu_A: X \rightarrow [0, 1]$, such that for any $x \in X$, $\mu_A(x)$ defines its membership value, which is a real number lying between 0 and 1.

Fuzzyness of concepts can be of two types; intrinsic fuzzyness and informational fuzzyness. Concepts like tall, young and beautiful have intrinsic fuzzyness, whereas concepts like trustworthy customer and efficient employee have informational fuzzyness. Both types of fuzzyness can occur in computer science.

Let X be a space of objects and x be a generic element of X. A classical set A, $A \subseteq X$, is defined as a collection of elements or objects $x \in X$, such that x can either belong or not belong to the set A. A fuzzy set A in X is defined as a set of ordered pairs

$$A = \{(x, \mu A(x)) | x \in X\}$$

Where µA (x) is called the membership function (MF) for the fuzzy set A. The MF maps each element of X to a membership grade between zero and one. Obviously (1) is a simple extension of the difinition of a classical set in which the characteristic function is permitted to have any values between zero and one.

3.5.2 Fuzzy Similarity Relation

(Slowinski and Vanderpooten 1997)[142] have considered binary relations, which are only reflexive and defined upper approximation and lower approximation of sets with respect to such relations. Gradually they improved the other properties of symmetry and transitivity and compared their results with those of existing ones.

In this chapter a fuzzy binary relation on U, which is a fuzzy similarity relation.

Definition 2. Let U be any universal set. Let R be a fuzzy binary relation defined on U. Then R is a fuzzy similarity relation defined as follows.

(i) R is said to be fuzzy reflexive if and only if
$R(x,x) = 1$ for all $x \in U$, $R(x, y) \in [0,1]$, for $y \in U$;

(ii) R is said to be fuzzy symmetric if and only if
$R(x,y) = R(y,x)$ for all $x,y \in U$; and

(iii) R is said to be fuzzy max-min transitive if and only if
$R(x,z) \geq \max_{y \in U} \min \{R(x,y), R(y,z)\}$, for all $x, y, z \in U$.

Note 1. In this case a max-min transitivity is a transitive relation, in general case, if $R(x,y) = R(y,z) =1$ then $R(x,z) \geq \max \min [R(x,y), R(y,z)] = 1$, i.e. $R(x,z) =1$.

3.5.3 Similarity Using Fuzzy Reflexive Relation

Let R be a fuzzy binary reflexive relation on U. For any $x,y \in U$, shall say x is fuzzy similar to y written as xRy if and only if $R(x,y) \geq 0$. Then R is a fuzzy binary reflexive relation on U.

Define the following two sets associated with any $x \in U$:

$R(x) = \{y \in U: yRx \text{ or } R(y,x) \geq 0\}$, and

$R^{-1}(x) = \{y \in U: xRy \text{ or } R(x,y) \geq 0\}$.

Here R(x) is called the *fuzzy similarity class of x* with respect to R, which consists of the set of objects which are fuzzy similar to x consider here that an object from one fuzzy similarity class may be fuzzy similar to an object of another fuzzy similarity class.

The fuzzy similarity classes are R(x) and $R^{-1}(x)$, where R(x) is the set of objects which are similar to x. Since some cases, the fuzzy similarity relation does not exist, the inverse relation of R, denoted by R^{-1} is considered so that a fuzzy binary relation will be a fuzzy symmetric relation.

The fuzzy similarity relations will be fuzzy symmetric if $R(x,y) = R(y,x)$, where $R(x,y)$, $R(y,x) \in [0, 1]$. This type of fuzzy similarity does not exist in real life application.

3.5.4 Fuzzy Ambiguity

Consider a non-empty set $X \subseteq U$ and a fuzzy binary relation R defined on a universe U. Again consider the followings cases.

Case I. If $x \in X$, but $\exists\, y \notin X$ such that x is fuzzy similar to y, i.e. $R(x, y) \not> 0$, then x will be discarded from X, by the information provided by R.

Case II. If $x \notin X$, but $\exists\, y \in X$ such that x is fuzzy similar to y, i.e. $R(x,y) > 0$, then x will be included in X, by the information provided by R.

Definition 3.3.3. Consider a subset $X \subseteq U$ and a fuzzy binary relation R defined on U, define the extended types of ambiguities as follows:

(i) An element $x \in U$ is an *fuzzy ambiguous object of type-I* with respect to R if and only if $x \in X$ and $R^{-1}(x) \cap (U \backslash X) \neq \varphi$. (1)

(ii) An element $x \in U$ is an *fuzzy ambiguous object of type-II* with respect to R if and only if $x \in U \backslash X$ and $R^{-1}(x) \cap X \neq \varphi$. (2)

Definition 3. An element $x \in U$ is said to be in X without ambiguity with respect to a fuzzy binary reflexive relation R and if $R^{-1}(x) \subseteq X$ and $x \in X$.

Such objects are called *fuzzy positive objects* of X with respect to R. (3)

Definition 4. An element $x \in U$ is said to be not in X without ambiguity with respect to a fuzzy binary reflexive relation R and $R^{-1}(x) \subseteq U \backslash X$ and $x \in U \backslash X$. Such objects are called *fuzzy negative objects* of X with respect to R. (4)

For the given $X \subseteq U$ and a fuzzy binary relation R, any object $x \in U$ belongs to one and only one of the following categories:

i) fuzzy positive objects of X with respect to R
ii) fuzzy ambiguous objects of type-I of X with respect to R
iii) fuzzy ambiguous objects of type-II of X with respect to R
iv) fuzzy negative objects of X with respect to R.

So, for any $X \subseteq U$, a fuzzy binary relation R induces a partition of U into the above four categories.

Theorem 1. Consider a non-empty subset $X \subseteq U$, a fuzzy binary reflexive relation R defined on U, then $\bigcup_{x \in X} R(x)$ corresponds to the set of fuzzy positive or fuzzy ambiguous objects.

The above result shows that the set of fuzzy positive or ambiguous objects can be described as a union of fuzzy similarity classes.

Theorem 2. Consider a non-empty subset $X \subseteq U$, a fuzzy binary reflexive relation R defined on U, Then

$$\bigcup_{x \in X} R(x) = \{x \in U: R^{-1}(x) \cap X \neq \varphi \}. \quad (5)$$

The above result shows the characterization of fuzzy positive and fuzzy ambiguous objects in other way.

3.5.5 Generalized Definition of Rough Approximation on Fuzzy Similarity Relations

The following definitions are given for the lower and upper approximations of a set based on the fuzzy binary reflexive relation.

Definition 3.4 Suppose $X \subseteq U$ and R is a fuzzy binary reflexive relation defined on U. Then define the lower approximation of X as

$$R_*(x) = \{x \in U: R^{-1}(x) \subseteq X \} \quad (6)$$

and the upper approximation of X as

$$R^*(x) = \bigcup_{x \in X} R(x) . \quad (7)$$

Theorem 3.5 Suppose $X \subseteq U$ and R is a fuzzy binary reflexive relation defined on U. Then $R_*(x) \subseteq X \subseteq R^*(x)$. (8)

3.5.6 Fuzzy Partitions over Lower and Upper Approximations on X

The fuzzy partitions of U can be defined using the lower and upper approximations as follows.

For any subset $X \subseteq U$,

(i)	fuzzy positive objects	:	$R_*(x)$
(ii)	fuzzy ambiguous objects of type-I	:	$X \setminus R_*(x)$
(iii)	fuzzy ambiguous objects of type-II	:	$R^*(x) \setminus X$
(iv)	fuzzy negative objects	:	$U \setminus R^*(x)$

From above, it is clear that X does not admit any fuzzy ambiguous object with respect to R if and only if $R_*(x) = R^*(x)$.

3.5.7 Fuzzy Sets vs. Rough Sets

Fuzzy sets are sets that have elements with degrees of membership. In other words, in fuzzy logic an element of a set has a degree of belonging or membership to those particular set (Zadeh 1965)[72] introduced fuzzy sets as an expansion of the classical concept of a set.

In classical set theory, the membership of elements in a set is evaluated in binary terms in that i.e., either it is a member of that set or it is not a member of that particular set. Fuzzy set theory allows the steady evaluation of the membership of elements in a set and this is illustrated with the help of a membership function allowed to fall within the interval [0, 1]. A fuzzy set is therefore a generalized version of a classical set. Conversely, a classical set is a special case of the membership functions of fuzzy sets which only permit values 0 or 1. Thus far, fuzzy set theory has not generated any results that differ from the results from a probability or classical set theory.

(Zhang and Shao 2006)[167] developed a self-learning method to identify a fuzzy model and to extrapolate missing rules. This was done using the modified gradient descent method and confidence measure. The method can simultaneously identify a fuzzy model, revise its parameters and establish optimal output fuzzy sets. When tested on a classical truck control problem, the results obtained show the usefulness and accuracy of the advanced method.

(Coulibaly and Evora 2007)[16] investigated the multilayer perception network, the time-lagged feed- forward network, the generalized radial basis function network, the recurrent neural network, the time delay recurrent neural network and the counter-propagation fuzzy-neural network along with different optimization methods for estimating daily time series with missing daily total precipitation records and extreme temperatures. The results obtained revealed that the multi-layer perception, the time-lagged feed-forward network and the counter-propagation fuzzy-neural network offered the highest accuracy in estimating missing precipitation values. The multi-layer perception was found to be successful at imputing missing daily precipitation values. In addition, the multi-layer perception was most suitable for imputing missing daily maximum and minimum temperature values. The counter-propagation fuzzy-neural network was similar to the multi-layer perception at imputing missing daily maximum temperatures; sadly, it was less useful for estimating the minimum temperature. The recurrent neural network and time delay recurrent neural network were found to be the least appropriate method for imputing both daily precipitation and the extreme temperature records, while the radial basis function was good at approximating maximum and minimum temperature.

A rough set is defined as a proper estimation of a classical (crisp) set through a pair of sets that offer the lower and the upper approximation of the original set. The lower approximations are sets that are similar to sets that have already been observed in the information system while the higher approximation sets are sets that can only be inferred either strongly or weakly from the information system. These lower and upper approximation sets are crisp sets in the classical description of rough set theory (Pawlak 1991)[104], but in other variants, the approximating sets may also be fuzzy sets.

The distinguishing features between fuzzy sets and rough sets are that while fuzzy sets operate via membership functions, rough sets operate through upper and lower approximation sets (Chanas & Kuchta, 1992)[13]. The similarity between the two approaches is that they are both designed to deal with the vagueness and uncertainty of the data. For example, (Jensen and Shen 2004)[59] introduced dimensionality reduction that retained semantics using both the rough and fuzzy-rough based approaches. Some researchers, such as (Deng et al..2007)[19] introduced a method that combines both rough and fuzzy sets. The reason for this hybridization process is that even though these techniques are similar, each one offers its own unique advantages. Such hybrid fuzzy rough sets scheme has been applied to model completed problems such as breast cancer detection. The following section describes rough sets, which is a method that will be adopted in this chapter.

3.6 Rough Equality of Sets

Extending the idea of equality of sets in crisp set theory, where two sets are said to be equal if and only if they have the same elements, three types of rough or approximate equalities have been introduced by (Pawlak, Z., and Skowron, 2007)[106] and state these definitions.

3.6.1 Definition

Let $K = (U, \Re)$ be a knowledge base, $X, Y \subseteq U$ and $R \in IND(K)$. We say that

(1) Two sets X and Y are *bottom R-equal* $(X \eqsim_R Y)$ if $\underline{R}X = \underline{R}Y$

(2) Two sets X and Y are *top R-equal* $(X \simeq_R Y)$ if $\overline{R}X = \overline{R}Y$

(3) Two sets X and Y are *R-equal* $(X \approx_R Y)$ if $(X \eqsim_R Y)$ and $(X \simeq_R Y)$. Equivalently, $\underline{R}X = \underline{R}Y$ and $\overline{R}X = \overline{R}Y$.

Drop the suffix R in the notations to make them look simpler.

It can be easily verified that the relations bottom R-equal, top R-equal and R-equal are equivalence relations on *P(U)*, the power set of *U*.

The concept of approximate equality of sets refers to the topological structure of the compared sets but not the elements they consist of. Thus sets having significantly different elements may be rough equal. In fact, if $X \eqsim Y$ then $\underline{R}X = \underline{R}Y$ and as $X \subseteq \underline{R}X$, $Y \subseteq \underline{R}Y$ X and Y can differ in elements of $X - \underline{R}X$ and $Y - \underline{R}Y$. However, taking for example $U = \{x_1, x_2, .., x_8\}$ and $R = \{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5, x_6\}, \{x_7, x_8\}\}$, that the two sets $X = \{x_1, x_3, x_5\}$ and $Y = \{x_2, x_4, x_6\}$ are top R-equal, though $X \cap Y =$ ¤ .

As noted by Pawlak, this concept of equality of sets is of relative character that is things are equal or not equal from our point of view depending on what we know about them. So, in a sense the definition of rough equality refers to our knowledge about the universe.

3.6.2 Properties

The following properties of rough equalities have been proved.

(1) $X \eqsim Y$ if and only if $X \cap Y \eqsim X$ and $X \cap Y \eqsim Y$.

(2) $X \simeq Y$ if and only if $X \cap Y \simeq X$ and $X \cap Y \simeq Y$.

(3) If $X \simeq X'$ and $Y \simeq Y'$ then $X \cup Y \simeq X' \cup Y'$.

(4) If $X \eqsim X'$ and $Y \eqsim Y'$ then $X \cap Y \eqsim X' \cap Y'$.

(5) If $X \simeq Y$ then $X \cup -Y \simeq U$.

(6) If $X \eqsim Y$ then $X \cap -Y \eqsim$ ¤ .

(7) If $X \subseteq Y$ and $Y \simeq$ ¤ then $X \simeq$ ¤ .

(8) If $X \subseteq Y$ and $X \simeq U$ then $Y \simeq U$.

(9) $X \simeq Y$ if and only if $-X \eqsim -Y$.

(10) If $X \eqsim$ ¤ or $Y \eqsim$ ¤ then $X \cap Y \eqsim$ ¤ .

(11) If $X \simeq U$ or $Y \simeq U$ then $X \cup Y \simeq U$.

In the following two properties of lower and upper approximations of rough sets, that inclusions hold and equality does not hold true in general:

(12) $\underline{R}X \cup \underline{R}Y \subseteq \underline{R}(X \cup Y)$ and

(13) $\overline{R}(X \cap Y) \subseteq \overline{R}X \cap \overline{R}Y$.

The following results provide necessary and sufficient conditions for equalities to hold in (3.4.1) and (3.4.2). In these results take $\{E_1, E_2, ..., E_n\}$ as a partition of an universe U with respect to an equivalence relation R and $X_1, X_2, ..., X_m$ are subsets of U.

Theorem 1. That has

(14) $$\bigcup_{i=1}^{m} \underline{R}(X_i) \subset \underline{R}\left(\bigcup_{i=1}^{m} X_i\right)$$

if and only if there exists at least one E_j such that $X_i \cap E_j \subset E_j$ for $i = 1, 2, ..., m$ and $\bigcup_{i=1}^{m} (X_i) \supseteq E_j$.

Corollary 1. Equality holds in (2.2.14) if and only if there exists no E_j such that $X_i \cap E_j \subset E_j$, $i = 1, 2, ..., m$ and $\bigcup_{i=1}^{m} X_i \supseteq E_j$.

Theorem 2 that has

(15) $$\overline{R}\left(\bigcap_{i=1}^{m} X_i\right) \subset \bigcap_{i=1}^{m} \overline{R}(X_i)$$

is that there exist at least one E_j such that $X_i \cap E_j \neq$ ¤ , $i = 1, 2, ..., m$ and $\left(\bigcap_{i=1}^{m} X_i\right) \cap E_j =$ ¤ .

Corollary 2. Equality holds if and only if there is no E_j such that $X_i \cap E_j \neq$ ¤ , $i = 1, 2, ..., m$ and $\left(\bigcap_{i=1}^{m} X_i\right) \cap E_j =$ ¤ .

It has been noted that the properties (3.6.1) to (3.6.2) fail to hold if X is replaced by E or vice versa. However, the following observations with regards to their interchange:

I. The properties (7) to (11) hold true under the interchange. That is,

(7)' $X \subseteq Y$ and $Y \eqsim ¤ \Rightarrow X \eqsim ¤$

(8)' $X \subseteq Y$ and $X \eqsim U \Rightarrow Y \eqsim U$

(9)' $X \eqsim Y$ if and only if $-X \simeq -Y$

(10)' If $X \simeq ¤$ or $Y \simeq ¤$ then $X \bigcap Y \simeq ¤$ and

(11)' If $X \eqsim U$ or $Y \eqsim U$ then $X \bigcup Y \eqsim U$.

II. The properties (5) and (6) holds true in the following form:

Let $BN_R(Y) = ¤$ (that is, Y is an R-definable set). Then

(5)' If $X \eqsim Y$ then $X \bigcup -Y \eqsim U$.

(6)' If $X \simeq Y$ then $X \bigcap -Y \simeq ¤$.

III. (i) The properties (3.6.1) and (3.6.2) hold if conditions of Corollary 2 hold with m = 2.

(iii) The properties (3.6.1) and (3.6.2) hold if conditions of Corollary 1 hold with m = 2.

So,

(1)' $X \simeq Y$ if and only if $X \bigcap Y \simeq X$ and $X \bigcap Y \simeq Y$.

(2)' $X \eqsim Y$ if and only if $X \bigcup Y \eqsim X$ and $X \bigcup Y \eqsim Y$.

(3)' $X \eqsim X'$ and $Y \eqsim Y' \Rightarrow X \bigcup Y \eqsim X' \bigcup Y'$.

(4)' $X \simeq X'$ and $Y \simeq Y' \Rightarrow X \bigcap Y \eqsim X' \bigcap Y'$.

3.7 Rough Equivalence of Sets

Here we introduce the concept of rough equivalence. In fact, we shall introduce three types of rough equivalence.

3.7.1 Definitions

1. That two sets *X* and *Y* are *bottom R-equivalent* if and only if both $\underline{R}X$ and $\underline{R}Y$ are ¤ or not¤ together (we write, *X* is b eqv. to *Y*).

2. That two sets *X* and *Y* are *top R-equivalent* if and only if both $\bar{R}X$ and $\bar{R}Y$ are *U* and not-*U* together (we write, X is t_eqv. to *Y*).

3. That two sets *X* and *Y* are *R-equivalent* if and only if *X* and *Y* are both bottom R-equivalent and top R-equivalent (X is eqv. *Y*).

   In these notations excluded the reference to R for simplicity.

3.7.2 Elementary Properties

1. It is clear from the definitions above that R-equality (bottom, top, total) implies R-equivalence (bottom, top, total) respectively.

2. Obviously the converses are not true.

3. Bottom R-equivalence, top R-equivalence and R-equivalence are equivalence relations on P (U).

4. As mentioned earlier, if two sets are rough equal to each other than by using our knowledge may be able to say that they are approximately equal. That is they have some close features which are enough to assume that they are approximately equal. Equivalently, with our present knowledge say that both have the same set of positive elements and same set of negative elements.

If two sets are roughly equivalent then by using our present knowledge, may not be able to say whether two sets are approximately equal as described above. But, it can say that they are approximately equivalent, that is both the sets have or have not positive elements with respect to R and both the sets have or have not negative elements with respect to R.

3.8 RST Application in Data Analysis

Rough set data analysis is one of the main techniques arising from RST; it provides a manner for gaining insight into the underlying data properties. The rough set model has several appealing advantages for data analysis. It is based on the original data only and does not rely on any external information, i.e. no assumptions about data are made. It is suitable for analyzing both quantitative and qualitative features leading to highly interpretable results.

In RST a training set can be represented by a table where each row represents an object and each column represents an attribute. This table is called an information system, more formally, it is a pair S= (U, A), where U is a not empty finite set of objects called the universe and A is a not-empty finite set attributes. The basic concepts of RST are the lower and upper approximations of a subset X $\subseteq$ U. These were originally introduced with reference to an indiscernibility relation *IND(B),* where objects x and y belong to *IND(B)* if and only if x and y are indiscernible from each other by features in B.

3.8.1 Information Systems

A data set is represented as a table, where each row represents a case, an event, a patient, or simply an object. Every column represents an attribute (a variable, an observation, a property, etc.) that can be measured for each object; the attribute may be also supplied by a human expert or the user. Such table is called an information system. Formally, an information system is pair S = (U, A) where U is a non-empty finite set of objects called the universe and A is a non-empty finite set of attributes such that $a:U \rightarrow V_a$ for every $a \in A$. The set V_a is called the value set of a.

Example 1: Let us investigate the Middle East solution problem using our approach. Let us consider the ten parties (objects) in this problem.

	a	b	c	d	e	f	g	h	i	j	k	l
Egypt	1	0	0	1	0	0	0	1	1	1	1	0
Israel	0	1	1	1	1	1	1	0	1	1	0	1
Jordan	1	1	0	1	0	0	0	1	1	0	1	1
Lebanon	1	1	0	1	1	0	0	1	1	0	1	0
Palestine	1	1	0	1	1	0	0	1	1	0	1	0
Syria	1	1	0	1	1	0	0	1	1	1	1	1
KSA	1	1	0	1	1	0	0	1	1	1	1	1
Iraq	1	0	0	1	0	0	0	1	1	1	1	0
Kuwait	0	1	1	1	1	1	1	0	1	1	0	1
Qater	1	0	1	1	0	0	0	1	1	0	1	0

Table-3.2 Information System before Iraq war

	a	b	c	d	e	f	g	h	i	j	k	l
Egypt	1	0	0	1	0	0	0	1	1	1	1	0
Israel	0	1	1	1	1	1	1	0	1	1	1	0
Jordan	1	1	1	1	1	0	1	1	1	0	0	1
Lebanon	1	1	0	1	0	0	0	1	1	1	1	1
Palestine	1	1	0	1	1	0	0	1	1	0	1	0
Syria	1	1	0	1	1	0	0	1	1	1	1	0
KSA	1	1	0	1	1	0	0	1	1	1	1	1
Iraq	0	1	1	1	1	1	1	0	1	1	1	0
Kuwait	0	1	1	1	1	1	1	0	1	1	1	0
Qater	0	1	1	1	1	1	1	0	1	1	1	0

Table-3.3 Information System after Iraq war

(1) Egypt (2) Israel (3) Jordan (4) Lebanon (5) Palestine (6) Syria (7) Iraq (8) Kuwait (9) Qatar (10) Saudi Arabia

The relations between those parties are determined by the following twelve issues (attributes).

(a) Return of Golan Heights of Syria.

(b) Israeli military outposts on the Golan Heights.

(c) Israeli occupation zone in south Lebanon.

(d) Free access to all religious centers.

(e) Arab countries grant citizenship to Palestinians who choose to remain within their borders.

(f) Israeli retains East Al-Quads.

(g) Isolation and division of Al-Quads.

(h) Autonomous Palestinian state on the West Bank and Gaza.

(i) Return of the West Bank & Gaza to Arab rule.

(j) Israeli military outpost along the Jordan River.

(k) Road map.

(l) The segregation wall.

The two information systems (Table-3.2, Table-3.3) summarize all the participants' opinion on the previous twelve issues before and after Iraq war. If the participant is against the issue put 0 and if the participant is neutral or favorable toward the issue assign that by 1.

In Table-3.3 the equivalence classes are:

{1, 8} {2, 9} {6, 7} {4, 5} {3} {10}

In Table-3 the equivalence classes are:

{1} {2, 8, 9, 10} {3, 4, 5}

From these equivalence classes, rule sets can be induced and are categorized into certain and possible categories due to the feature of lower and upper approximations of RST (Slowinski and Stefanowski, 1989)[141]. Certain rules and possible rules may be propagated separately. These rule sets will be useful to make necessary classifications for the approximations with the associated degree of membership.

The degree of membership of any party (Israel) with respect to a set A = {Egypt, Kuwait, Iraq, Qatar, Palestine} can be measured from the above information systems (Tables) as specified below.

$$\frac{|\{1, 5, 8, 9, 10\} \cap \{2, 9\}|}{|\{2, 9\}|} = 1/2$$

To confine the discussions and for more can be referred (Fiskel, J, 1980)[27] and (Grzymala-Busse, J, 1988)[35]. However, in real time, the classifications need to be derived from different sets of attribute-values (other than 0, 1). Below some of the approximation methods using RST and covering based RST classifications.

3.9 Approximations of Classifications

As pointed out by (Pawlak, 2007)[95], the approximation of classifications of (Grzymala Busse, 2004)[45] establish two concepts, approximation of sets and approximation of families of sets (or classifications). Refer (Tripathy et.al 2009)[148] below a classification on a universe formally.

Definition 3.7.1 Let F = $\{X_1, X_2, \ldots X_n\}$ be a family of non empty sets defined over U. That F is a *classification of* U if and only if $X_i \cap X_j = \emptyset$ for $i \neq j$ and $\bigcup_{i=1}^{n} X_i = U$.

Definition 3.7.2 Let F be as above and R be an equivalence relation over U. Then $\underline{R}$ F and $\overline{R}$ F denote respectively the R-*lower* and R-*upper approximations* of the family F and are defined as

$$\underline{R}F=\{\underline{R}X_1, \underline{R}X_2,\ldots,\underline{R}X_n\},$$

$$\overline{R}F=\{\overline{R}X_1, \overline{R}X_2,\ldots,\overline{R}X_n\}.$$

3.10 Covering Based Approximations of Classifications

Covering based approximations of classifications is an extension of the basic approximations of classifications (Degang, C et.al 2007)[20], which are to deal with the same lower approximations for all types of covering based approximations and different types of upper approximations for different types (Tripathy, B. K et.al.2009)[149].

3.10.1 Definition

Definition 1 Let U be a domain, C a family of subsets of U. If none subsets in C is empty, and $\cup C$ = U, then C is called a covering of U, and the ordered pair < U,C > is called a covering approximation space.

Definition 2 Let < U,C > be a covering approximation space, x $\in$ U, then the set family $\{K \in C | x \in K \Lambda (\forall S \in C \Lambda x \in S \wedge S \subseteq K \Rightarrow K = S)\}$ is called the minimal description of x and denoted by md(x).

Definition 3 Let U be a finite nonempty domain, C is a covering of U, for any X$\subseteq$ U ,the covering upper and lower approximation set families of X are respectively defined as

$$\overline{C}(X) = \{x \in U | \cap md(x) \cap X \neq \Phi\},$$
$$\underline{C}(X) = \{x \in U | md(x) \subseteq X\}.$$

3.10.2. Covering reducts

Definition 4 Let U be a finite nonempty domain, C is a covering of U, K $\in$ C, if K is the union of sets in C – {K}, then K is called a reducible element of C, otherwise K is called an irreducible element of C.

Definition 5 Let U be a finite nonempty domain, C is a covering of U, the covering after removing all the reducible element of C is called the reduct of C, and denoted by reduct(C).

This new algorithm to extract rules from incomplete information system based on the covering of rough sets models. The method mainly focuses on the way to estimate unknown attribute values.

Let F = {X1, X2… Xn} be a classification of U. Then for any other cover C of U define:

$F_* = \{X_{1*}, X_{2*}, ..., X_{n*}\}$ and $F_i^* = \{X_{1i}^*, X_{2i}^*, ..., X_{ni}^*\}$ called the *C-lower approximation* and *C-upper approximation of type-i* (i=1, 2, 3) of the classification F respectively; where X_{i*} *and* X_{ij}^* *are* respectively the C-lower approximation and C-upper approximation of type-j (j= 1, 2, 3) of X_i , i = 1, 2,…n respectively.

Example 1: Let us consider a very simple information system shown in the Table-3.4. The set of objects U consists of seven objects: x_1, x_2, x_3, x_4, x_5, x_6, x_7, and the set of attributes(2) includes: Age and Lower Extremity Motor Score (LEMS).

	Age	**LEMS**
x_1	16-30	50
x_2	16-30	0
x_3	31-45	1-25
x_4	31-45	1-25
x_5	46-60	26-49
x_6	16-30	26-49
x_7	46-60	26-49

Table-3.4 Information System 1

The process is known as supervised learning. Information systems of this kind are called decision systems. A decision system (a decision table) is any information system of the form S = (U, A ∪ {d}), where d ∉ A is the decision attribute. The elements are called conditional attributes or simply conditions. The decision attribute may take several values though binary outcomes are rather frequent.

Example 2: Let us consider a decision system presented in Table-3.4. The table includes the same seven objects as in the previous example and one decision attribute (Walk) with two Boolean values: Yes, No.

	Age	**LEMS**	**Walk**
x_1	16-30	50	Yes
x_2	16-30	0	No
x_3	31-45	1-25	No
x_4	31-45	1-25	Yes

Table-3.5 Information System 2

3.10.3 Indiscernibility Relation

A decision system expresses all the knowledge about the model. This table may be unnecessarily large in part because it is redundant in at least two ways. The same or indiscernible objects may be represented several times, or some of the attributes may be superfluous.

Some relevant issues are discussed below:

Let S = (U, A) be an information system, and $B \subseteq A$. A binary relation $IND_S(B)$ defined in the following way

$$IND_S(B) = \{ (x,x') \in U^2 \mid \forall\, a \in B\ a(x) = a(x') \}$$

is called the B-indiscernibility relation. It is easy to see that $IND_S(B)$ is equivalence relation. If $(x,x') \in IND_S(B)$, then objects x and x' are indiscernible from each other by attributes from B. The equivalence classes of the B-indiscernibility relation are denoted $[x]_B$. The subscript S in the indiscernibility relation is usually omitted if it is clear which information system is meant.

Some expressions of standard rough sets do not require from a relation to be transitive. Such a relation is called tolerance relation or similarity.

Example 3: In order to illustrate how a decision system (Table-3.4) defines an indiscernibility relation, the following three non-empty subsets of the conditional attributes: {Age}, {LEMS} and {Age, LEMS}.

If consideration the set {LEMS} then objects x_3 and x_4 belong to the same equivalence class; they are indiscernible. From the same reason, x_5, x_6 and x_7 belong to another indiscernibility class. The relation IND defines three partitions of the universe.

$$IND(\{Age\}) = \{\{x_1,x_2,x_6\}, \{x_3,x_4\}, \{x_5,x_7\}\},$$

$$IND(\{LEMS\}) = \{\{x_1\}, \{x_2\}, \{x_3,x_4\}, \{x_5,x_6,x_7\}\},$$

$$IND(\{Age, LEMS\}) = \{\{x_1\}, \{x_2\}, \{x_3,x_4\}, \{x_5,x_7\}, \{x_6\}\}.$$

3.10.4 Set Approximation

This defines formally the approximations (Zhu, W 2007)[146] of a set using the discernability relation.

Let S=(U, A) be an information system and let $B \subseteq A$ and $X \subseteq U$. Approximate a set X using only the information contained in the set of attributes B by constructing the B-lower and B-upper approximations of X, denoted $\underline{B}X$ and $\overline{B}X$ respectively, where $\underline{B}X = \{x \mid [x]_B \subseteq X\}$ and $\overline{B}X = \{x \mid [x]_B \cap X \neq \phi\}$.

Analogously as in a general case, the objects in $\underline{B}X$ can be with certainly classified as members of X on the basis of knowledge in B, while the objects in $\overline{B}X$ can be only classified as possible members of X on the basis of knowledge in B. The set $BN_B(X) = \overline{B}X - \underline{B}X$ is called the B-boundary region of X and thus consists of those objects that cannot decisively classify into X on the basis of knowledge in B. The set U- $\overline{B}X$ called is called the B-outside region of X and consists of those objects which can be with certainly classified as do not belonging to X (on the basis of knowledge in B). A set is said to be rough (respectively crisp) if the boundary region is non-empty (respectively empty).

Example 4: Let X = {x: Walk(x) =Yes}, as given by Table 3.4. In fact, the set X consists of three objects: x_1, x_4, x_6. Now, to describe this set in terms of the set of conditional attributes A = {Age, LEMS}. Using the above definitions, the following approximations: the A-lower approximation $\underline{A}X=\{x_1, x_6\}$, the A-upper approximation $\overline{A}X = \{x_1, x_3, x_4, x_6\}$, the A-boundary region $BN_S(X) = \{x_3, x_4\}$ and the A-outside region U- $\overline{A}X=\{x_2, x_5, x_7\}$. It is easy to see that the set X is rough since the boundary region is not empty. The graphical illustration of approximations of the set X together with the equivalence classes contained in the corresponding approximations is shown in Figure-3.1

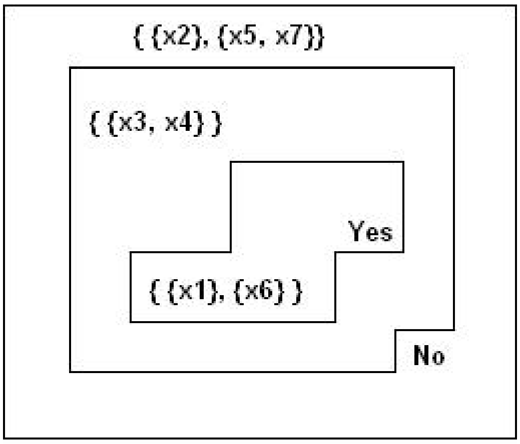

Figure-3.2 Approximations

It is easily seen that the lower and the upper approximations of a set, are respectively, the interior and the closure of this set in the topology generated by the indiscernibility relation.

The lower approximation of X in A is the greatest definable set in A, contained in X. The upper approximation of X in a is the least definable set in A containing X. Time complexity of algorithms for computing lower and upper approximations of any set X is $O(n^2)$, where n is the cardinality of set U of examples. Specific problems addressed by the theory of rough sets are:

3.10.5 List of properties of approximations:

1. $\underline{B}(X) \subseteq X \subseteq \overline{B}(X)$
2. $\underline{B}(\phi) = \overline{B}(\phi) = \phi$, $\underline{B}(U) = U$
3. $\overline{B}(X \cup Y) = \overline{B}(X) \cup \overline{B}(Y)$
4. $\underline{B}(X \cap Y) = \underline{B}(X) \cap \underline{B}(Y)$
5. $X \subseteq Y$ implies $B(X) \subseteq B(Y)$ and $B(X) \subseteq \overline{B}(Y)$
6. $\underline{B}(X \cup Y) \supseteq \underline{B}(X) \cup \underline{B}(Y)$
7. $\overline{B}(X \cap Y) \subseteq \overline{B}(X) \cap \overline{B}(Y)$
8. $\underline{B}(-X) = -\overline{B}(X)$
9. $\overline{B}(-X) = -B(X)$

3.11 Some More Application in RST:

This section provides a brief overview of some of the many applications of rough set theory. There are several properties of rough sets that make the theory an obvious choice for use in dealing with real problems; for example, it handles uncertainty present in real data through approximations and also does not require threshold information in order to operate (as is the case with many current techniques).

3.11.1 Prediction of Business failure

Attempts to develop business failure prediction models began seriously sometime in the late 1960s and continue through today. Although there has been much research in this area, there is still no unified well specified theory of how and why corporations fail. Financial organizations need these predictions for evaluating firms of interest. Many methods have been used for the purpose of bankruptcy prediction, such as logic analysis, discriminate analysis and probit analysis (F. E. H. Tay, L. Shen 2002)[26]. A comprehensive review of the various approaches to modeling and predicting this is presented in (A. I. Dimitras et.al 1996)[1]. Although some of these methods led to satisfactory models, they suffered from limitations, often due to unrealistic statistical assumptions. Because of this, the rough set model, with its aim of keeping model assumptions to a minimum, appeared to be a highly useful approach for the analysis of financial information tables. Rough set-based failure prediction was investigated in (R. Slowinski et.al 1999) [121] and (A. I. Dimitras et.al 1999)[1]. In these investigations, the rough set approach was evaluated against several other methods, including C4.5 (J. R. Quinlan 1993)[112], discriminate analysis and logic analysis. For the rough approach, decision rules were generated from the reducts produced by analysis of the financial information. All methods were then evaluated on data from the previous three years.

The rough set model was found to be more accurate than discriminate analysis by an average of 6.1% per case, using a minimal set of reduced rules. It also outperformed C4.5, but performed similarly to logic analysis. A comparative study of the rough sets model versus multi-variable discriminate analysis (MDA) can be found in (A. Szladow, D. Mills 1993)[2]. It was demonstrated that through the use of rough set theory, the prediction of corporate bankruptcy was 97.0% accurate - an improvement over MDA which achieved an accuracy of 96.0%.

3.11.2 Financial Investment

Trading systems have been built using rough set approaches. In (R. Golan, D. Edwards 1993)[114], (R. Golan 1995)[115] and (W. Ziarko et.al 1993)[151], the rough set model was applied to discover strong trading rules that reflect highly repetitive patterns in data. Historical data from the Toronto stock exchange in 1980 was used for the extraction of trading rules for five companies. Experts confirmed that the extracted rules described the stock behavior and market sensitivity of these companies. Depending on a roughness parameter, the rules generated were either \general" or \exact". The general rules were all recognized relationships in the investment industry, whereas the exact rules made less sense.

In the work reported in (J. Bazan 1994)[52], the problem of how to deduce rules that map the financial indicators at the end of a month to the stock price changes a month later was addressed. This was based on 15 market indicators. From this study, only a satisfactory performance was achieved with many issues still to be tackled, such as data filtration and how to handle missing data. In (J. K. Baltzersen 1996)[56], research was carried out into rough set reduct analysis and rule construction for forecasting the total index of the Oslo stock exchange. This also achieved satisfactory results, with a highest accuracy of 45%.

Research has been carried out on building trading systems for the S&P index (M. Ruggiero 1994)[86]. Here, a hybrid system was developed that incorporated both neural networks and rough sets. Rules generated by rough sets were used to supervise neural networks to correct for possible errors in predictions. This system reduced drawdown by 25-50% and increased the average winner/loser ratio by 50-100%.

Rough sets have also been applied to financial decision analysis and explanation for an industrial development bank, ETEVA (R. Slowinski et.al 1994, 1995)[119,120]. The bank was interested in investing its capital in firms, whilst reducing the risk involved in such an investment. To achieve this, a rough set-based firm assessment system was constructed that decided, based on a number of financial ratios, whether a company was acceptable, unacceptable or uncertain. An information table was constructed with the help of the financial manager of ETEVA. From this, the rough set-generated rules revealed the financial policy applied in the selection of firms. The rules can also be used to evaluate new firms that seek financing from the bank.

3.11.3 Bioinformatics and Medicine

A common and diagnostically challenging problem facing emergency department personnel in hospitals is that of acute abdominal pain in children. There are many potential causes for these pains most are usually non-serious. However, the pain may be an indicator that a patient has a serious illness, requiring immediate treatment and possibly surgery. Experienced doctors will use a variety of relevant historical information and physical observations to assess children. Such attributes occur frequently in recognizable patterns, allowing a quick and efficient diagnosis. Inexperienced physicians, on the other hand, may lack the knowledge and information to be able to recognize these patterns. The techniques developed in (K. Farion et.al 2004)[63] provide a rough set-based clinical decision model to assist such inexperienced physicians. In this research, rough sets are used to support diagnosis by distinguishing between three disposition categories: discharge, observation/further investigation, and consult. Preliminary results show that the system gives accuracy comparable to physicians, though it is dependent on a suitably high data quality.

Rough set data analysis is also applied to the problem of extracting protein-protein interaction sentences in biomedical literature (F. Ginter et.al 2004)[27]. Due to the abundance of published information relevant to this area, manual information extraction is a formidable task. This approach develops decision rules of protein names, interaction words, and their mutual positions in sentences. To increase the set of potential interaction words, a morphological model is developed, generating spelling and infection variants. The performance of the method is evaluated using a hand-tagged dataset containing 1894 sentences, producing a precision-recall break-even performance of 79.8% with leave-one-out cross-validation.

Automated classification of calculated electroencephalogram (EEG) parameters has been shown to be a promising method for detection of intra-operative awareness. In (M. Ningler et.al 2004)[84], rough set-based methods were employed to generate classification rules resulting in satisfactory accuracy rates of approximately 90%.

3.11.4 Fault Diagnosis

A rough set approach for the diagnosis of valve faults in a multi-cylinder diesel engine is investigated in (L. Shen et.al 2000)[75] .The use of rough sets enabled the diagnosis of several fault categories in a generic manner. A decision table was constructed from attributes extracted from the vibration signals, with four operational states studied among the signal characteristics: normal, intake valve clearance too small, intake valve clearance too large, and exhaust valve clearance too large. Three sampling points were selected for the collection of vibration signals. The results demonstrated that the system is quite effective for such fault diagnosis, and the extracted rules correspond well with prior knowledge of the system. In (K. Mannar and D. Ceglarek 2004)[64], a rough set-based method for continuous failure diagnosis in assembly systems is presented. Sensor measurements were used to construct a diagnosis table from which rough set rules were extracted.

3.11.5 Spatial and Meteorological Pattern

Classification Sunspot observation, analysis and classification form an important part in furthering knowledge about the sun, the solar weather, and its effect on earth. Certain categories of sunspot groups are associated with solar fares. Observatories around the world track all visible sunspots in an effort to early detect fares. Sunspot recognition and classification are currently manual and labor intensive processes which could be automated if successfully learned by a machine. The approach presented in (S. H. Nguyen et.al 2005)[130] employs a hierarchical rough set-based learning method for sunspot classification. It attempts to learn the modified Zurich classification scheme through rough set-based decision tree induction.

The resulting system is evaluated on sunspots extracted from satellite images, with promising results. In (J. F. Peters et.al 2003)[53], a new application of rough set theory for classifying meteorological radar data is introduced. Volumetric radar data is used to detect storm events responsible for severe weather. Classifying storm cells is a difficult problem as they exhibit a complex evolution throughout their lifespan. Also, the high dimensionality and imprecision of the data can be prohibitive. Here, a rough set approach is employed to classify a number of meteorological storm events.

3.11.6 Music and Acoustics

A dominance-based rough set approach, an extension of rough sets to preference-ordered information systems, was used in (J. Jelonek et.al 2004)[55] to generate preference models for violin quality grading. A set of violins were submitted to a violinmaker's competition and evaluated by a jury according to several assessment criteria. The sound of the instruments was recorded digitally and then processed to obtain sound attributes. These features, along with jury assessments were analyzed by the rough set method, generating preference models. It was shown that the jury*0*s rankings were well approximated by the automated approach.

A decision system employing rough sets and neural networks is presented in (B. Kostek et.al 2004)[8]. The aim of the study was to automatically classify musical instrument sounds on the basis of a limited number of parameters, and to test the quality of musical sound parameters that are included in the MPEG-7 standard. The use of wavelet-based parameters led to better audio retrieval efficiency. The classification of musical works is considered in (M. P. Hippe 2002)[85] based on the inspection of standard music notations. A decision table is constructed, with features representing various aspects of musical compositions (objects), such as rhythm disorder, beat characteristics and harmony. From this, classification rules are induced (*via* rough set rule induction) and used to classify unseen compositions.

3.11.7 Feature Selection

Feature selection methods can be used to identify and remove unneeded, irrelevant and redundant attributes from data that do not contribute to the accuracy of a predictive model or may in fact decrease the accuracy of the model.

The discussion above has focused on actual practical applications of rough set theory. This section is concerned with the theoretical advancement of feature selection within the rough set community.

As indicated previously, the work on rough set theory offers a formal methodology that can be employed to reduce the dimensionality of datasets, often as a preprocessing step to assist other tasks like learning from data (Q. Shen 2003)[107]. The QuickReduct algorithm provided earlier is a typical example of rough set-assisted feature selection tools. Such a method helps select the most information rich features in a dataset, without transforming the data, all while attempting to minimize information loss during the selection process. Computationally, the approach is highly efficient as it involves simple set operations only. Thus, it represents one of the most successful applications of rough sets. However, it is reliant upon a discrete dataset; important information may be lost as a result of quantization of the underlying numerical features (that real-world problems typically have).

Feature selection is different from dimensionality reduction. Both methods seek to reduce the number of attributes in the dataset, but a dimensionality reduction method do so by creating new combinations of attributes, where as feature selection methods include and exclude attributes present in the data without changing them. Examples of dimensionality reduction methods include Principal Component Analysis, Singular Value Decomposition and Sammon's Mapping.

It is natural, then, to apply its extensions to this area. Such research has been carried out in (R. Jensen, Q. Shen 2004 , 2007)[116,117], where a reduction method was proposed based on fuzzy extensions to the positive region and dependency function based on fuzzy lower approximations. A greedy hill-climber is used to perform subset search, using the fuzzy dependency function both for subset evaluation and as a stopping criterion. The method was used successfully within a range of problem domains, including web content classification and complex system monitoring.

Optimizations are given in (R. B. Bhatt, M. Gopal 2005)[113] to improve the performance of the method. In (R. B. Bhatt, M. Gopal 2005)[113], a compact computational domain is proposed to reduce the computational effort required to calculate fuzzy lower approximations for large datasets, based on some of the properties of fuzzy connectives. Fuzzy entropy is used in (N. M. Parthalain et.al 2006)[93] to guide the search toward smaller reducts. In In (R. B. Bhatt, M. Gopal 2005)[113], an alternative search algorithm is presented that alleviates some of the problems encountered with a greedy hill-climber approach. This problem is also tackled in (R. Jensen, Q. Shen 2005)[117] *via* the use of a novel ant colony optimization-based framework for feature selection. A genetic algorithm is used in (L. Zhou et.al 2006)[76] for search based on the fuzzy dependency function within a face recognition system with promising results. The work in (G. C. Y. Tsang et.al 2005)[36], (X. Z. Wang et.al 2005)[155] improves upon these developments by formally defining relative reductions for fuzzy decision systems. A discernibility matrix is constructed for the computation of all such reductions. As the resulting discernibility matrix is crisp, some information may have been lost in this process. Additionally, there are complexity issues when computing discernibility matrices for large datasets.

However, in the crisp rough set literature there have been methods proposed that avoid this (Jue. Wang, Ju. Wang 2001)[61], and similar constructions may be applicable here. Feature selection algorithms, based on the generalization of fuzzy approximation spaces to fuzzy probability approximation spaces are introduced in (Q. Hu et.al 2006)[108]. This is achieved through the introduction of a probability distribution on the universe. Information measures for fuzzy indiscernibility relations are presented in (Q. Hu et.al 2006)[109] for the computation of feature importance. Reducts are computed through the use of a greedy selection algorithm similar to QuickReduct.

3.12 Conclusion

This study, it has discussed the Rough set, Fuzzy set, Covering based Rough set and Soft set theory, was proposed by (Z. Pawlak)[103], (Lotfi A. Zadeh 1965)[72] and (Molodtsov 1999)[90] as an approach to knowledge discovery from incomplete, vagueness and uncertain data. The rough set approach to processing of incomplete data is based on the lower and the upper approximation, and the theory is defined as a pair of two crisp sets corresponding to approximations.

The main advantage of rough set theory in data analysis is that it does not need any preliminary or additional information concerning data, such as basic probability assignment in Dempster-Shafer theory, grade of membership or the value of possibility in fuzzy set theory. The Rough Set approach to analysis has many important advantages such as (Pawlak, 1991)[104] Finding hidden patterns in data; Finds minimal sets of data (data reduction); Evaluates significance of data; Generates sets of decision rules from data; Facilitates the interpretation of obtained result Different problems can be addressed though Rough Set Theory, however during the last few years this formalism has been approached as a tool used with different areas of research.

There has been research concerning the relationship between Soft Set theory and rough sets and fuzzy sets. Soft set theory has also provided the necessary formalism and ideas for the development of some propositional machine learning systems. It has also been used for knowledge representation; data mining; dealing with imperfect data; reducing knowledge representation and for analyzing attribute dependencies. Soft set theory has found many applications such as power system security analysis, medical data, finance, voice recognition and image processing; and one of the research areas that have successfully used Soft Set is the knowledge discovery or Data Mining in database.