Dominance-Based Rough Set Approach

Gaurav Raj

July 7, 2022

1 Rough Sets

1.1 A General Overview of Rough Sets

The idea of Rough Sets, as introduced by Pawlak (1982), has been crucial in the field of decision making when a vague description of the actions (objects in the decision making) is given. Vagueness, or ambiguity, in this data arises from information granulation. The rough sets philosophy is based on the fact that all objects in the universe can be associated with some attribute, which is used for the description of the object. These attributes, too, then take up a range of values, and various objects can be **indiscernible** with respect to one another in terms of attributes considered. In the Rough Set theory, the Universe is made up of blocks of indiscernible relations, which can then be used to classify any given real-world or abstract problem. The use of indiscernibility relations leads to information granulation.

Any subset X of the universe can be represented in terms of these blocks of indiscernibility relations, either as an exact union, or an approximation. In the second case, we use two metrics to define set X - a lower approximation and an upper approximation. For crisp sets, these are the same sets. The difference between the upper and lower approximations, intuitively, then determines the boundary region for set X (This boundary being an empty set for a crisp set). Moreover, the cardinality of the boundary region X can be used as a measure of the vagueness of set .

An analysis of multi-criteria decision problems using rough sets often yields results in the form of 'if.....then.....' statements, which are logically sound and easily understandable. We will discuss more on the information table and various other aspects of the Rough Set theory in the coming sections, with the help of an illustration.

1.2 Information Tables, Indiscernibility Relation, and Membership Functions

1.2.1 Information Tables and Indiscernibility Relations

The information present with a DM is often represented in the form of a table, with the rows being objects, and the columns being attributes. Thus, a cell represents a piece of information about an object, that is, the quality of the object wrt one criterion/attribute. Formally, an information table is the 4-tuple $S = \langle U, Q, V, f \rangle$, where U is a finite set of objects (universe). $Q = \{q_1, q_2, \dots q_m\}$ is a finite set of attributes, V_q is the domain of the attribute q, and $f : U \times Q \to V$ is a total function such that f(x,q)Vq for each qQ, xU is called information function.

To every (non-empty) subset of attributes P is associated an indiscernibility relation on U, denoted by I_p:

$$I_p = \{(x, y) \in U \times U : f(x, q) = f(y, q) \;\forall\, q \in P\}$$

In other words, if two or more objects have the same values for one more (p) attributes, then they are p-indiscernible. As stated earlier, these indiscernibility relations divide the Universe U into building blocks that can be used to characterize any subset of U. It is evident that the indiscernibility relation is an equivalence relation, i.e., it is symmetric, reflexive, and transitive. The family of all the equivalence classes of the relation I_p is denoted by $U|I_p$ and the equivalence class containing an element x by $I_p(x)$. The equivalence classes of the relation I_p are called P-elementary sets.

1.2.2 Approximations

As discussed in the Rough Set Theory, X is not a crisp set, rather, is defined by Upper and Lower Approximations. Let S be an information table, and U denote the Universe, and $X \subset U$. Then the P-lower and P-upper approximations of X are respectively defined as follows :

$$\underline{P}(X) = x \in U : I_p(x) \subsetneqq X$$

$$\bar{P}(X) = \bigcup_{x \in X} I_p(x)$$

In other words, $\underline{P}$, the lower approximation of set X is the set of all elements $x \in U$ such that they definitely belong to the set X and have are discernible with respect to every object outside of X. On the other hand, $\bar{P}(X)$, the upper approximation of the set X is the set of all elements which belong to X as well as those which are indiscernible to an element that truly belongs to X. In other words, $\underline{P}(X)$ is the largest union of the P-elementary sets included in X, while $\bar{P}(X)$ is the smallest union of the P-elementary sets containing X.

The **P-boundary** of X in S, denoted by Bn_p, is given by

$$Bn_p(X) = \bar{P}(X) - \underline{P}(X)$$

It is easy to see that the following relation holds : $\underline{P}(X) \subsetneqq X \subsetneqq \bar{P}(X)$. This means that if an object belongs to the lower approximation, $\underline{P}$, it definitely belongs to the set X, while if it belongs to $\bar{P}(X)$, it may or may not belong to X. For the elements of set $Bn_p(X)$, then, nothing can be said with respect to their belonging to X. we will soon see an illustration to understand the lower and upper approximations as well the boundary of X. Meanwhile, the following set relationship holds for the upper and lower approximations.

$$\underline{P}(X) = U - \bar{P}(U - X)$$

This property satisfied by the rough sets is often called the *Complementarity Property.*

A rough set, then is one for which $Bn_p \neq \phi$, and a set for which it is equal to ϕ is called a crisp set.

1.2.3 Parameters related to rough sets

Various parameters are used to define the degrees of vagueness etc. for a rough set, and they majorly involve the cardinality of the upper and lower approximations of the set X. Some of these are listed as follows :

- **Accuracy of Approximation**
  The accuracy of approximation of X is given by the following ratio (for $X \neq \phi$)

  $$\alpha_P(X) = |\underline{P}(X)|/|\bar{P}(X)|$$

  where $|X|$ denotes the cardinality of a finite set X. The result is, obviously, $0 \leq \alpha_p(X) \leq 1$; if $\alpha_p(X) = l$, X is an ordinary (exact) set with respect to P; if $\alpha_p(X) < l$, X is a rough (vague) set with respect to P.

- **Quality of Approximation**
  The quality of approximation of X is given by the following ratio (for $X \neq \phi$)

  $$\gamma_P(X) = |\underline{P}(X)|/|X|$$

  The quality $\gamma_p(X)$ represents the relative frequency of the objects correctly classified using the attributes from P. Moreover, we have $0 \leq \gamma_p(X) \leq \alpha_p(X) \leq 1$; $\gamma_p(X) = O$ iff $_p(X) = O$ and $\gamma_p(X) = 1$ iff $_p(X) = 1$

- **Rough Membership Function**
  From the viewpoint of a particular object $X \in U$, it may be interesting, to use the available in-formation to assess the degree of its membership to a subset X of U. The subset X can be identified with concept of knowledge to be approximated. Using the rough set approach one can calculate the membership function $\nu_X^p(x)$ (rough membership function) as

$$\nu_X^p(x) = |X \cap I_p(x)|/|I_p(x)|$$

  . The value of $\nu_X^p(x)$ may be interpreted analogously to conditional probability and may be understood as the degree of certainty (credibility) to which x belongs to X. Observe that the value of the membership function is calculated from the available data, and not subjectively assumed, as it is the case of membership functions of fuzzy sets. The The following relationships hold for the rough membership function and the approximations of the set X :

$$\underline{P}(X) = \{x \in U : \nu_X^p(x) = 1\}$$
$$\bar{P}(X) = \{x \in U : \nu_X^p(x) > 0\}$$
$$Bn_p(X) = \{x \in U : 0 < \nu_X^p(x) < 1\}$$
$$\underline{P}(U - X) = \{x \in U : \nu_X^p(x) = 0\}$$

  The results are quite intuitive : For an element to be part of the lower approximation of the set, it must be a part of the set X as well as the the equivalence classes generated by the indiscernibility relation on P, where as it doesn't have to be so for the upper approximation.

1.2.4 Reduction of attributes

Dependence of Attributes is one of the most important concepts for practical applications of the Rough Set Approach. Intuitively, a set of attributes $T \subseteq Q$ totally depends on a set of attributes $P \subseteq Q$ (notation P→T) if all the values of the attributes from T are uniquely determined by the values of the attributes from P, that is, if a functional dependence exists between evaluations by the attributes from P and by the attributes from T. In other words, the partition generated by the attributes from T is "finer" than that generated by the attributes from T, so that it is sufficient to use the attributes from P to build the partition $U|I_T$. Formally, T totally depends on P iff $I_P \subseteq I_T$. Therefore, T is totally (partially) dependent on P if all (some) elements of the universe U may be univocally assigned to classes of the partition $U|I_T$, using only the attributes from P.
Another issue of great practical importance is that of "superfluous" data in an information table. Superfluous data can be eliminated, in fact, without deteriorating the information contained in the original table. Let $P \subseteq Q$ and $p \in P$. It is said that attribute p is superfluous in P if $I_P = I_{P-\{p\}}$; otherwise, p is indispensable in P.

1.3 An Example

Instead of going through the theory of decision rules and decision tables, we will understand it by the following example. The example will also help us understand the things we talked about earlier, including approximations and reducts. The following table describes 6 warehouses on the following attibutes :

- A_1, capacity of the sales staff
- A_2, perceived quality of goods,
- A_3, high traffic location,
- A_4 warehouse profit or loss.

As discussed before, the components of the information table S are: U=1,2,3,4,5,6, Q=A_I, A_2, A_3, A_4, V_1=high, medium, low, V_2=good, medium, V_3=no, yes, V_4=profit, loss, the information function f(x,q), taking values f(1,A_1)=high, f(1,A_2)=good, and so on.

Warehouse	A_1	A_2	A_3	A_4
1	High	Good	No	Profit
2	Medium	Medium	No	Loss
3	Medium	Medium	No	Profit
4	Low	Medium	No	Loss
5	High	Good	Yes	Loss
6	Medium	Medium	Yes	Profit

It can be observed that the warehouses have all different descriptions with respect to the attributes $A_1, A_2, A_3 and A_4$, so that they can be distinguished (discerned) by means of the information supplied by the attributes considered. Formally, the indiscernibility relation based on all four attributes is

$$I_Q = \{(1,1),(2,2),(3,3),(4,4),(5,5),(6,6)\}$$

and, therefore, there is no two distinct warehouses x and y such that $(x, y) \in I_Q$. However, warehouses 2 and 3 are indiscernible in terms of the attributes from $P = \{A_1, A_2, A_3\}$, since they have the same values on the three attributes. Formally, the indiscernibility relation based on P is, thus, $I_p = \{(1,1),(2,2),(2,3),(3,2),(3,3),(4,4),(5,5),(6,6)\}$. Similarly, warehouses 2, 3 and 4 are indiscernible with reference to the attributes of $P' = \{A_2, A_3\}$, and so on, considering all the possible subsets of attributes from Q.

Let us take the set of attributes P = $\{A_1, A_2, A_3\}$. Using this set, we need to approximate the set X of warehouses that made a profit. Clearly, $U|I_P$ = 1, 2,3, 4, 5, 6, and X = 1,3,6. The lower and upper approximations for the set X with respect to P are then given by :

$$\underline{P}(X) = 1, 6, \bar{P}(X) = 1, 2, 3, 6, Bn_P = 2, 3$$

The accuracy of the approximation is thus given by 0.5, while the quality of the approximation is given by 0.67. It is then quite intuitive to conclude the following :

1. Warehouses 1 and 6, from the P-lower approximation of X, certainly belong to the set of warehouses that make a profit,

2. Warehouses 1,2,3 and 6, from the P-upper approximation of X, could belong to the set of warehouses that make a profit,

3. Warehouses 2 and 3, from the P-boundary of X, represent cases of uncertain membership to the set of warehouses that make a profit.

We can do the same exercise for the set Y of warehouses that made a loss, and we get

$$\underline{P}(Y) = 4, 5, \bar{P}(Y) = 2, 3, 4, 5, Bn_P = 2, 3$$

with the accuracy of the approximation and the quality of the approximation same as before.

Let us consider now the following subsets of $Q : P = \{A_1, A_2, A_3\}, R = \{A_1, A_2\}, T = \{A_1, A_3\}, W = \{A_2, A_3\}$. It is then easy to observe that $I_p = I_R, I_T = I_P and I_T \neq I_P$. This means that R and T are reducts of P, while W is not. In other words, R and T are minimal subsets of P that induce the same partition of the elements of U as the set of attributes P. It can also be observed that in the core of P, defined by $R \cap T$, there is attribute A_1, which is then indispensable for the approximation of the class of warehouses that make a profit (and also for the class of warehouses that make a loss), while other attributes from R and T may be mutually exchanged.

We will now introduce decision tables and decision rules, that happen to be the very end results of the rough set theory and the familiar $if \ldots statement \ldots$ that we saw earlier. These statements are derived for all occurring combinations of the attribute assignments and what they imply for the final decision attribute of the object.

If in the set of attributes Q, condition attributes $C = \{A_1, A_2, A_3\}$ and decision attribute $D = \{A_4\}$ were distinguished, the information table could be seen as a decision table.If in an information table the attributes of set Q are divided into condition attributes $(setC \neq \phi) and decision attributes (setD \neq \phi), C \cap D = Q and C \cup D = 0$, such a table is called a decision table.

In order to explain the evaluations of the decision attribute by means of the evaluations of the condition attributes, one can represent the information table as a set of decision rules. Such a representation of the aforementioned information table gives the following rules:

- $if\ f(x, A_1) =$ high $and\ f(x, A_2) =$ good $and\ f(X, A_3) =$ no, $then\ f(x, A_4) =$ profit. In linguistic terms, "if the capacity of the sales staff is high and the perceived quality of goods is good and the location is not in high traffic conditions, then the warehouse makes a profit". The other conclusions are listen (without reference to their linguistic meanings) next.
- $if\ f(x, A_1) =$ medium $and\ f(x, A_2) =$ medium $and\ f(X, A_3) =$ no, $then\ f(x, A_4) =$ loss.
- $if\ f(x, A_1) =$ medium $and\ f(x, A_2) =$ medium $and\ f(X, A_3) =$ no, $then\ f(x, A_4) =$ profit.
- $if\ f(x, A_1) =$ low $and\ f(x, A_2) =$ medium $and\ f(X, A_3) =$ no, $then\ f(x, A_4) =$ loss.
- $if\ f(x, A_1) =$ medium $and\ f(x, A_2) =$ good $and\ f(X, A_3) =$ yes, $then\ f(x, A_4) =$ loss.
- $if\ f(x, A_1) =$ high $and\ f(x, A_2) =$ medium $and\ f(X, A_3) =$ yes, $then\ f(x, A_4) =$ profit.

Now, with the use of induction, these decision rules can be further simplified and written in a more condensed form as follows :

- $if\ f(x, A_1) =$ high, $then\ f(x, A_4) =$ profit
- $if\ f(x, A_1) =$ low, $then\ f(x, A_4) =$ loss
- $if\ f(x, A_1) =$ medium $and\ f(x, A_2) =$ good, $then\ f(x, A_4) =$ loss