**FUZZY ROUGH BIREDUCTS:->**

Feature selection(general)

**Feature selection** is the process of reducing the number of input variables when developing a predictive model.

It is desirable to reduce the number of input variables to both reduce the computational cost of modelling and, in some cases, to improve the performance of the model.

There are often many features involved, and combinatorially large numbers of feature combinations, to select from. Note that the number of feature subset combinations with m features from a collection of N total features is N!/[m!(N − m)!]. It might be expected that the inclusion of an increasing number of features would increase the likelihood of including enough information to distinguish between classes. Unfortunately, this is not necessarily true if the size of the training dataset does not also increase rapidly with each additional feature included. A high-dimensional dataset increases the chances that a learning algorithm will find spurious patterns that are not valid in general. More features may introduce more measurement noise and, hence, reduce performance

There are many different approaches being utilised for this task but we will be discussing about the fuzzy rough feature selection based approach here:

As most datasets contain real-valued attributes, it is necessary to perform a discretization step beforehand. This is typically implemented by standard fuzzification techniques enabling linguistic labels to be associated with attribute values. It also aids the modelling of uncertainty in data by allowing the possibility of the membership of a value to more than one linguistic label.

Fuzzy-rough sets encapsulate the related but distinct concepts of vagueness (for fuzzy sets [408]) and indiscernibility (for rough sets), both of which occur as a result of uncertainty in knowledge [86]. A fuzzy-rough set is defined by two fuzzy sets, fuzzy lower and upper approximations, obtained by extending the corresponding crisp rough set notions. In the crisp case, elements that belong to the lower approximation (i.e., have a membership of 1) are said to belong to the approximated set with absolute certainty. In the fuzzy-rough case, elements may have a membership in the range [0, 1], allowing greater flexibility in handling uncertainty

Fuzzy-rough feature selection (FRFS) [168] provides a means by which discrete or real-valued noisy data (or a mixture of both) can be effectively reduced without the need for user-supplied information

If the fuzzy-rough reduction process is to be useful, it must be able to deal with multiple features, finding the dependency among various subsets of the original feature set.

///////////////////////**//////////////////////after lower and upper approximations:-**

Now the problem we face while using the fuzzy-rough approach is that int the case of conventional RSAR, a reduct is defined as a subset R of the features that have the same information content as the full feature set A . In terms of the degree of dependency function this means that the values γ (R) and γ (A) are identical and equal to 1 but in the case of fuzzy-rough approach this is not necessarily true because as we encounter some uncertainty when objects belong to multiple fuzzy equivalence which results in decrease in total dependency which inturn means that the degree of dependency value may be less than 1

Quickreduct algorithm:

Quick reduct algorithm helps us to calculate reduct without exhaustively generating all possible subsets which would obviously be very expensive . It starts off with an empty set and adds those attributes one at a time which give maximum increase in degree of dependence and this procedure goes on until the degree of dependence (ɣ) cannot be increased further and if that is the case the process stops and gives us the required reduct .

Flow of algorithm:-

Step1:We set the initial reduct to an empty set R and the value of ɣbest to 0.

Step2:Set the value of ɣprev equal to ɣbest;

Step3:We consider another set T=R

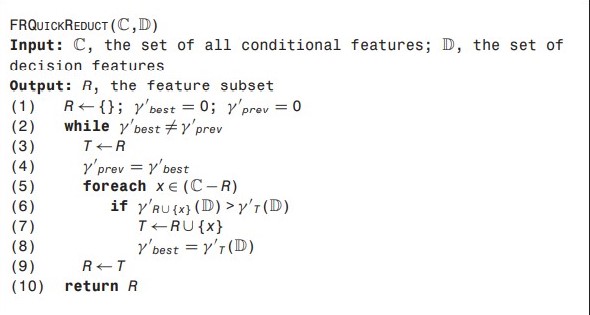
Step 3: We loop through all the values of x that are not already present in R and calculate the value of ɣ for RU{x} for each of these and check if ɣRU{x}> ɣT

if that is true we set the value of T=RU{x} and ɣbest= ɣT

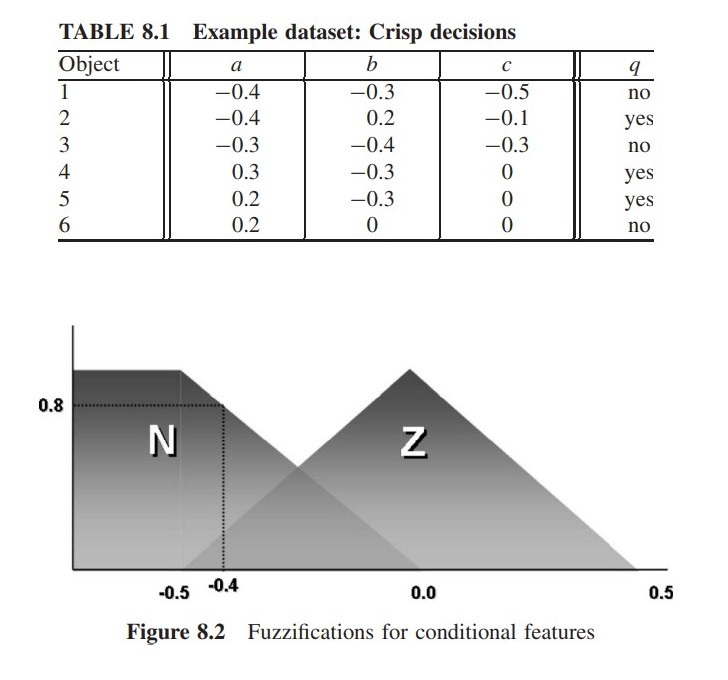
Step 4: R=T

Step5:if ɣbest!= ɣprev we go back to step 2

Step 6:R is the reduct.

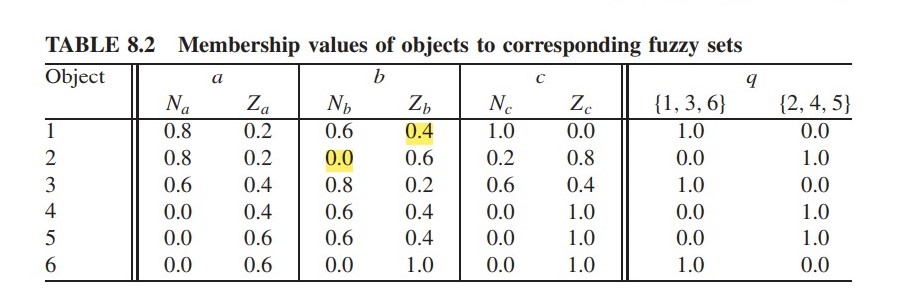


Example:



Fuzzification techniques:

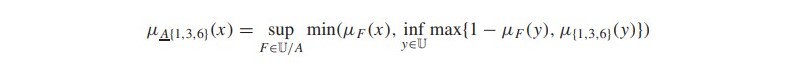
Fuzzification is a step to determine the degree to which an input data belongs to each of the appropriate fuzzy sets via the membership functions.

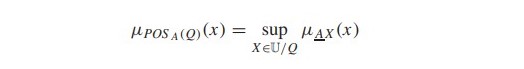


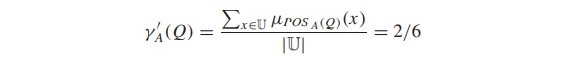
1-Initially the reduct R is empty that means R={} amd the value of ɣbest=0.

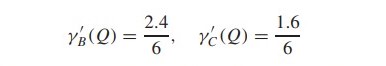
2-Value of ɣprev=ɣbest=0;

3-The first step is to calculate the lower approximations of the sets A, B, and C and calculate the value of degree of dependence of all these attributes









4-We can clearly see that the value of ɣA is greater than that of ɣB and ɣC hence ɣbest=2.4/6 and therefore we add b to R hence R becomes R={b}.

5-ɣprev=ɣbest=2.4/6

6-Now we add a to R and calculate the value of ɣ{a,b}and the add c to R and calculate values of ɣ{b,c}



7-As the value of ɣ{a,b} is greater ɣ{b,c} and also is greater than ɣbest we set the value of ɣbest =ɣ{a,b} and add a to R hence R becomes {a,b}

8-ɣprev=ɣbest

9-Now only c is left hence we add c to R and calculate ɣ{a,b,c} ;



10-It can be seen that the value of ɣ{a,b,c} is not greater that ɣbest hence we don’t add c to R and end the process here.

11- R={a,b}

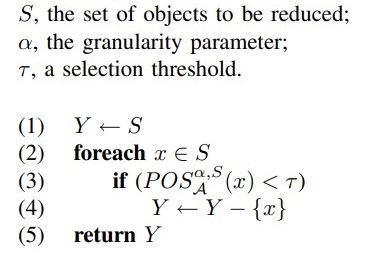
**Instance selection:**

Sometimes when we are dealing prohibitively high number of training instances, the use of instance selection becomes very important to make the volume of data manageable and to remove misleading training instances to get a better picture of the scenario we are dealing with.

**Fuzzy-rough instance selection:**

For this we use the information present in the positive region to determine how useful instances are and whether they can be removed by setting a threshold value.

**Algorithm used :**



The algorithm evaluates the degree of membership of each object x to the positive region and if this is less than the threshold, then the object can be removed. When an object membership is less than threshold, this means that there is some uncertainty as to which class this object truly belongs. If all such objects are removed, then there is no inconsistency exhibited by the remaining objects.

**Algo explanation:-**

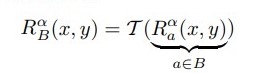
* Now to express the approximate equality between two instances with respect to each other , we take a fuzzy relation Ra for x,y in X and define it as follows:

****

Here l(a) id the range of attribute a and a(x) is the value of attribute a for instance x.

The parameter α (α ≥ 0) determines the granularity of Ra (tells us about the scale or detail in a set of data)

* For any subset B of A, the fuzzy B-indiscernibility relation is:

****

* The lower approximation Rα B↓A of a fuzzy set A in S by means of a fuzzy relation Rα B is defined by, for y ∈ S:



* We can then define the fuzzy B-positive region POSα,S B by, for y in S,

