Project Report

On

FUZZY SIMILARITY RELATION AND IT'S APPLICATION IN FEATURE SELECTION

Sent in part fulfilment of the requirements for the B.Sc. in Mathematics degree.

Submitted by:-

SANSKRITI SAINI
RANJAN KUMAR YADAV
NAMAN BALIYAN

B.Sc. Mathematics (*6th* Semester)

Under the Supervision of

Dr. Shivam Shreevastava

and Co-supervision of

Mrs. Neelam Kumari

Division of Mathematics
School of Basic and Applied Science
GALGOTIAS UNIVERSITY, Uttar Pradesh

GALGOTIAS UNIVERSITY

(Established under Galgotias University Uttar Pradesh Act No. 14 of 2011)

School Of Basic And Applied Science

CERTIFICATE

This is to Certify that this project report work entitled “**Fuzzy similarity relation and its application in feature selection**” under our supervision. This work is fit for submission for the award of Bachelor Degree in Mathematics.

Shivam Shreevastava

Supervisor
Dr. Shivam Shreevastava

Co-Supervisor
Mrs. Neelam Kumari

Dean
School of Basic & Applied Sciences
(Galgotias University)

ACKNOWLEDGEMENT

We are really fortunate to have received this support during the creation of our report because the achievement of ultimate result of this project necessitated a great deal of direction and a lot of people's help. We appreciate their watching over all we did, so please accept our gratitude.

We are grateful to Dr. Sushant Shekhar for giving us the chance to work on a project at Galgotias university and for providing us with the assistance and direction necessary for us to successfully finish the project. Despite his hectic schedule overseeing corporate issues, he provided us with a lot of help and guidance, for which we are quite grateful.

We owe a debt of appreciation to our mentors Dr. Shivam Shreevastava and Mrs. Neelam Kumari, who showed a genuine interest in our project work and guided us through it, giving us all the information we needed to create a good system.

We won't forget to thank Dr. Varsha Gautam, the program chair for the Mathematics division, for her encouragement as well as for her prompt assistance and direction during the course of our project work.

We are lucky to have the ongoing support, encouragement, and direction of the mathematics department's teaching team, which enabled us to successfully complete our project work.

Preface:

A number of factors drive the employment of the dimensionality reduction (DR) step of a knowledge discovery in database (KDD) tool in various problem-solving systems. Numerous application challenges combine real-valued vectors to process data (for example, protein classification, bookmark etc.). Processing typically becomes impossible when these vectors have a significant degree of dimension. Thus, it is occasionally advantageous and frequently required to reduce the data dimensionally to a size that is more manageable while preserving most or all the original information. High dimensional complex processes can occasionally be controlled by a much smaller number of basic factors. Modelling these phenomena with the use of dimensionality reduction improves their clarity. The various information systems frequently contain a sizable amount of redundant, useless, or false information. Before any additional processing can be done, this must be removed. For instance, a pre-processing phase called data reduction is typically beneficial for the challenge of producing classification rules from huge volume information. This shortens the time needed to complete induction, improves the clarity of the resulting rules, and may increase classification accuracy. Semantics-destroying dimensionality reduction methods permanently change the meaning of the data, whereas semantics-preserving DR methods (like feature selection) try to preserve the original feature set's meaning. The basic goal of feature selection is to isolate a problem domain's smallest feature subset while maintaining an appropriate level of accuracy in re- presentation the original content. Sometimes several different features are used to choose from a huge number of feature combinations. As more qualities were involved, it might be assumed that the likelihood of having adequate information to distinguish between classes would increase. Unfortunately, if the size of the training dataset does not swiftly grow with the addition of each new feature, this is not the case. The "curse of dimensionality" is used to describe this. Data-mining algorithms are more likely to discover fake patterns that aren't often true when the dataset is high-dimensional. A well-organized and efficient reduction strategy is needed because the majority of systems use some level of reduction to deal with vast amounts of data.

It is obviously desired to use a method that can reduce dimensionality utilizing data from the data set while maintaining the semantics one of the characteristics (namely, semantics-preserving). In order to uncover decrease data dependencies and the volume of features in a only Rough set theory (RST) may be used as such a tool because it only requires the data and no outside knowledge. .RST has in fact grown in popularity over the past 10 years and has been used in a variety of fields. The Pawlak-initiated classical rough set [1] model has been successfully used to pick a feature, attribute reduction, and rule learning technique. A crisp equivalence relation and crisp equivalence classes are used in the rough set model to specify the dependence function between the choice and conditional attributes that are accessible in the information system. The significance between the decision and conditional features is successfully established using the dependency function, and the classification potential of the features is evaluated. However, because the traditional rough set model only required nominal data, it was unable to be applied directly to real-valued datasets. To prevent the information loss, numerous generalizations of the rough set concept have been presented.

One of the most beneficial rough set extensions that may be used with real-valued datasets without altering the information system is fuzzy rough set. Fuzzy rough sets can often handle the fuzziness and ambiguity seen in datasets with continuous properties. The concept of fuzzy rough sets is made possible by combining rough (as proposed by Pawlak [1]) and fuzzy (as proposed by Zadeh [2]) sets, as reported by Dubois and Prade [3–4]. This offers a potent approach to dealing with the discretization issue and can be successfully used to reduce continuous attributes. In the proposed paradigm of fuzzy rough sets, two items are compared using a fuzzy similarity relation[5], which is specified by real-valued conditional characteristics. When there is a mismatch between conditional and decision attributes, or when a small number of samples have the same or nearly identical conditional attribute values but different labels, fuzzy rough sets are most frequently employed to address the issue. The evaluation of this contradiction is made possible by the assignment of each sample to a membership in the form of a decision class under lower approximations in fuzzy rough sets. The fundamental objective of feature selection methods based on fuzzy rough sets is to get a reduct while maintaining each sample's membership. For real-valued data, the combination of fuzzy and rough sets offers a crucial path for reasoning with uncertainty. The concepts of vagueness (for fuzzy sets) and indiscernibility (for rough sets) are combined to form fuzzy rough sets. Both of these concepts are a result of knowledge uncertainty. The absence of distinct divisions or limits within the data itself leads to vagueness. This is a common technique for people to communicate and reason. Numerous techniques have been developed to work around the drawbacks of the conventional rough set technique, including the fuzzy rough set idea.

We have read a number of articles about fuzzy rough aided feature selection for this project. The degree of dependence method and the approach based on a discernibility matrix are the two major types of techniques found in the literature. In this work, we mainly focused upon degree of dependency approach.

Abstract:

Owing to technology advancements and the rising expansion of electronically stored information, automated solutions are required to assist users in processing and maintaining this large volume of information. The primary sources of knowledge are subject matter experts and computer program that evaluate enormous amounts of data using machine learning. Knowledge extraction is a crucial process stage in the construction of clever and skilled systems. However, because of the noise and the volume of data, the knowledge extraction stage is extremely sluggish or perhaps impossible. The effectiveness of classifiers and the readability of data in machine learning algorithms both benefit from the decision of pertinent and characteristics without repetition. This process the term "feature selection" or attribute reduction. Numerous domains, such as the use of image processing, artificial intelligence, bioinformatics, data mining, natural language processing, etc., use feature selection in ways that are very relevant to expert and intelligent systems. The discretization process may result in some information being lost, rendering rough set theory unsuitable for attribute reduction of real-valued data sets, despite the fact that it has been employed effectively for attribute reduction. Real-valued data can be handled easily thanks to the numerous attribute selection algorithms that have been given, In addition, the integration of collection of blurry and rough theories. In the current study, we have compiled all prior research on the selection of fuzzy rough sets-based features.

Introduction:

The amount of digital data that can be used in the globe is always growing because we are living in a data-driven age the advancement of computers and database technology. In the current situation, all business organization constantly acquire data based on millions of observations across a range of themes, brands, predictor factors, and storage sites on a periodic basis as a result of the growth of internet-based technologies. Everyday quintillion of byte data are stored in several formats. nodal points like information relating to various banking and business transactions, bio-genetic information in health services, enormous amount of statistical data regarding mass population and satellite data information of global and regional climate changes. New tools are always a requirement to analyse and process this large volume data so as to enable the extraction of useful information from the entire information system. This extracted information is the source of information. Knowledge finding in databases (KDD)is an exploratory and automatic analysis and modelling of large volume data repositories. KDD is the suitable and organized process of identifying novel, useful, understandable, and valid patterns from large and complex information systems. The abundance of data available today and their accessibility makes knowledge discovery a matter of considerable, necessity, and importance. The KDD process can be divided into the following stages:

Data generation: An objective dataset is created or acquired from a certain domain. To achieve that, it involves combining several existing datasets to obtain the ideal set of samples.

Data cleaning entails a variety of operations to produce processed data, including attribute discretization, noise removal, and missing value imputation, among others. Enhancing the general quality of any information that may be gleaned from the information technology is the main concern.

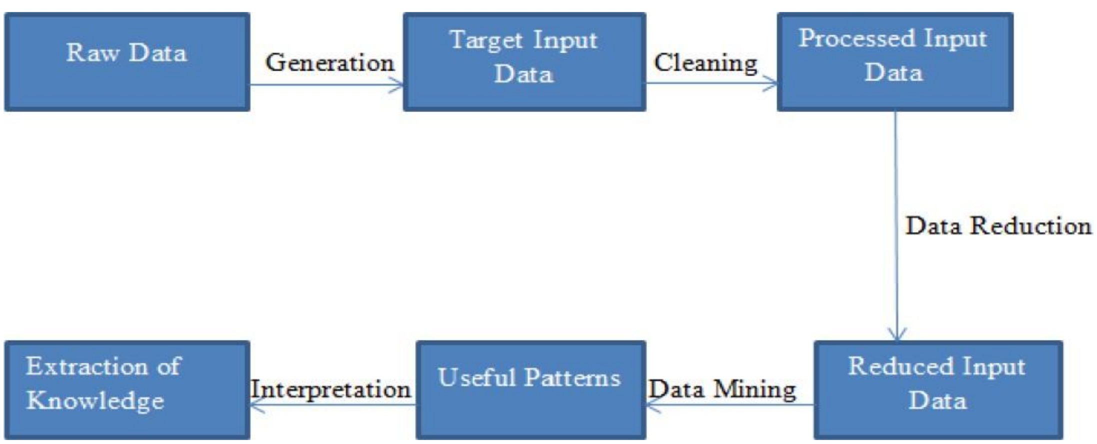
Data reduction: The amount of data points or instances and the number of features are both growing exponentially. Redundancy is a common feature of datasets that hinders knowledge discovery and may even cause errors in the process as a whole. Finding valuable features to describe the complete data set and removing irrelevant characteristics are the key goals of this phase as well as redundant features and similar data points. This step also aids in preserving the time, storage and cost whenever data mining is being done and also improves the interpretability is a data.

Data mining: The activity of sorting through voluminous data to find intriguing patterns and extract knowledge is known as data mining. Data mining is a group of strategies employed in an automated way to thoroughly examine and provide the fore complicated relationships in a vast amount of datasets. Databases may serve as the data sources, the Web, data warehouses, other information repositories, or data that are fed into the system dynamically.

Interpretation/evaluation: After the discovery of knowledge, It is assessed in terms of novelty, validity, usability, and simplicity. Some of the earlier steps might need to be repeated for this. Figure 1.1 depicts the full procedure in its entirety, with the dimensionality reduction (DR) step serving as the system's preprocessing stage.

Curse Of Dimensionality:

Data production is reaching new heights. This increase can be observed in virtually all spheres of human endeavor, ranging from the data produced on a everyday activities, such as business and financial transactions, telephone calls, etc. to more advanced and intricate data, such as molecular datasets, medical records, astronomical data, genome data, etc. These databases could include a wealth of useful information that has yet to be found. This increase in data may take one of two forms: either an increase in the number of samples or instances, or an increase in the number of features that are tracked down and computed. This rise has made processing datasets with hundreds of thousands of features or attributes necessary for many real-world applications. Some of these datasets are also accessible to the public at UCI.



Feature Selection:

A preprocessing phase in KDD called feature selection [22-91] selects the best subset of available features based on a predetermined criterion. The criterion takes feature subset measurement specifics into account. The decision to choose a certain criterion is influenced by the goal of feature selection or attribute reduction. When the values of the other assessment metrics are identical, an optimal subset can be a small subset that provides the best estimate of total prediction accuracy. In some situations, one must find a subset with the stated number that best meets the requirement if the number of features is given (like in data visualization and projection, where this number is typically 2 or 3). (such as degree of dependency, information gain, consistency, etc.). A dimensionality reduction method's choice or extraction of an Ideal feature subset is always a relative concept. to a particular feature evaluation criterion. Generally speaking, various criteria may result in various ideal feature groups. However, every requirement, seeks to quantify how well a feature or combination of characteristics can differentiate between several class labels. One of the major issues with actual data analysis is uncertainty. Incompleteness and ambiguity in class definitions are a few origins of this uncertainty. In this context, modelling and spreading uncertainty have become more popular because to the possibility idea proposed by rough set theory. It has been used for fuzzy rule extraction and modelling, feature selection, classification, clustering, and reasoning with uncertainty. Utilizing feature selection techniques, undesirable, redundant, and irrelevant attributes that do not contribute to the accuracy of any prediction model are found and removed from the dataset. The model should have the fewest possible attributes to reduce its complexity. A brute force method cannot be utilised for feature selection because there are nCm number of feature subsets with m features from a collection of n total features.

Benefits of Feature Selection:

Several advantages of feature selection may include:

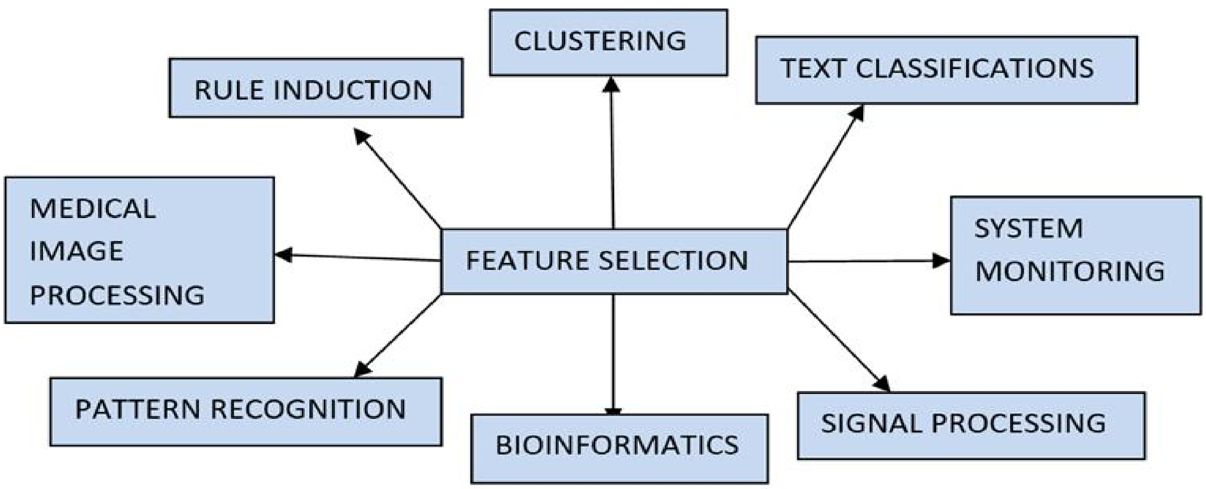
1. Data visualization: pattern and the data's trends can be visualized and recognized more quikly by reducing the dimensionality of data. It is found to be important in scenarios where there are little features a significant effect on data results. Algorithms for learning may not be able to discriminate key aspects distinguish the feature from the rest. or attribute set, hence it produces quite complicated models, the interpretation of which is very expensive.
2. The measurement and storage requirements: in various domains where cost and time-expense of taking the measurements for data attainment is significant, only fewer features or attributes are desirable due to the expense involved. Also, in domains where large datasets are encountered, and space becomes an issue, manipulated To accommodate storage needs, data size must be reduced.
3. Training and utilization times: feature selection for big data results in smaller data size which in turn significantly enhances the machine learning algorithms' training and classification step running times.
4. Prediction performance: feature selection in several cases enhance overall accuracy of classifiers, through the removal of misleading, redundant and noisy features. When machine learning algorithms are data for unknown samples may be less accurate because trained models may not be able to distinguish certain properties or qualities.

Applications of Feature Selection:

Data mining, image recognition, signal processing, text classification, bioinformatics, system monitoring, data clustering, and rule induction are a few examples of applications for feature selection or attribute reduction. Peaking is a phenomenon that frequently occurs when classifiers are trained using a small number of sample sets. The classification rate of the classifier stops increasing or, in some situations, even starts to decline after a period if the number of qualities or features in the data is increased. For instance, dermatologists only diagnose malignant melanomas with a clinical accuracy of between 65 and 85 percent when it comes to melanoma. Automated techniques for identifying skin tumours can be as accurate as 95% when feature selection algorithms are used. The field of gene expression micro-arrays, a technique that can assist in analyzing the expression hundreds of genes at different levels in a same experiment, is another area in which feature selection or attribute reduction is put to use. It's utilized to tell healthy cells apart from malignant ones. Feature selectors are employed in these situations to the gene databases' size, which would otherwise be too large for further processing, must be decreased. The amount of structural and functional data obtained by studying the human genome has multiplied during the past few years. These datasets, which would otherwise be too large for further processing, are reduced in size using feature selectors. Other bioinformatics applications Included in this are QSAR (Quantitative Structure Activity Relationship), whose major objective is to establish theories connecting chemical characteristics of molecules to their molecular activity, and splice site prediction, which identifies junctions between coding and non-coding portions of DNA.

The usage of if-then production rules is the most popular method for creating expressive and human readable representations of knowledge. There are still no comprehensive and general expert criteria for classifying feature patterns into their underlying classes in real-world issue domains. Using an appropriate feature selection approach, the rule induction process can be sped up by lowering rule complexity. A number of inferential measurement systems are also developed utilizing data-based methodologies. The models are used to infer the value of the target class label. This shows that the internal models of inferential systems are significantly influenced by the quality of the data used to build such models. As a result, it is anticipated that complex application issues, including industrial plant diagnosis and precise monitoring, would produce a significant number of characteristics, many of which will be useless and unneeded for further processing. Additionally, it is found that measuring these qualities has a cost. Therefore, it is always advantageous to have an intelligent system that can choose the elements that are most important in order to build the most reliable and accurate model for future processing. The graph might offer a feature.

The most crucial component of the clear approaches to feature selection is the utilization of user-provided data. This in and of itself is a serious flaw because several feature selections need the user to input noise levels in advance. Some algorithms merely rank features, leaving it up to the user to choose a subset In some circumstances, users must provide a cutoff point that establishes when the algorithm should terminate as well as the minimum number of characteristics that must be selected. But each of them requires the user to develop a judgement based on their own preferences. Rough set theory is one method that has been effective in accomplishing data reduction. Over the past 20 years, RST has become more and more popular and has been used to a variety of industries. (For instance, clustering, expert systems, system monitoring, and classification.)



Information System:

Information systems are described by the quadruple IS = (U, AT, V, h), where U = (u1, u2,.., un) is a finite collection of non-empty things, commonly referred to as the universe of discourse, and AT = a1, a2,.., an is a non-empty finite set of attributes.

Decision system:

IS = (U, AT, V, h) is referred to be a decision system if AT is the union of C and D, where C is a set of conditional attributes and D is a set of decision attributes. Table 1 gives a decision-making system illustration.

Example of a Decision System in Table 1.

Attributes / Objects	*a*	*b*	*c*	*d*	Q
0	1	0	2	2	0
1	0	1	1	1	2
2	2	0	0	1	1
3	1	1	0	2	2
4	1	0	2	0	1
5	2	2	0	1	1
6	2	1	1	1	2
7	0	1	1	0	1

Rough Set:

While requiring less knowledge overall, rough sets can be utilize to condensely extract knowledge from a subject while still maintaining the information's content. The primary benefit of rough set analysis is that all that is needed to get started is the information that has already been provided [10]. A key RST principle that can be used in feature selection is discernibility.

Rough set theory (RST) is a powerful mathematical tool for dealing with uncertainty. It may be used to quickly extract knowledge from any area because it shrinks the system while maintaining the original information content.

1. Indiscernible Relation:

An associated equivalence relation RB for every subset of the attributes $B \subseteq AT$ is defined as

$$R_B = \{(x, y) \in U \times U \mid \forall b \in B, b(x) = b(y)\}$$

If (x, y) belongs to R_B, then x and y are considered as identical for B. In this instance, the equivalence relation is B-indiscernible and has equivalence classes of type $[x]_B$.

2. Upper and lower approximations:

Consider an information system to be (U, AT, V, f). For X belongs to U ; the following descriptions of the lower approximation $R_B \downarrow X$ and the upper approximation $R_B \uparrow X$ can be used to estimate as

$$R_B \downarrow X = x \in U | \, [x]_B \subseteq X$$

$$R_B \uparrow X = x \in U | \, [x]_B \cap X \neq \emptyset$$

The duo ($R_B \downarrow X$, $R_B \uparrow X$) is referred to as a rough set.

Rough Set based Feature Selection:

A decision attribute in the decision system (U,C,D,V,f) should be Q belongs to D. The equivalence classes of [x] are decision classes Q. The objects from U that fall within the P-positive region are those for which P values clearly predict the decision class given P belongs to C. (POS P).

$$.\, POS_P(Q) = \bigcup_{x \in X} R_P \downarrow [x]_{R_Q}$$

Now, the dependency degree of Q on P is:

$$\gamma P\,(Q) = \frac{|POS_P(Q)|}{|U|}$$

Decision system is said to be consistent if $\gamma c\,(Q) = 1.$

Reduct:

If a subset P of C satisfies the requirements for the degree of dependence based reduct,

$$\gamma_P = \gamma_C \text{ with } \gamma_{Pa} \neq \gamma_P, \forall a \in P.$$

Initial state of the reduct set is an empty set. The degree to which the decision attribute depends on the reduct set is then determined by gradually introducing conditional qualities into the reduct set. The suggested approach only selects conditional qualities that maximise the decision attribute's degree of dependency.

Fuzzy Sets:

Each component of a set is given a membership value, or the degree to which it belongs to the set, via fuzzy sets. A fuzzy set in X is a mapping with the form X to [0, 1].

Let E indicate the age of hospital patients in years, where E = 1, 2,..., 90. The fuzzy set A = Youth E can then be found by using a membership grade as:

$$\mu_A = \begin{cases} \dfrac{1}{1+\left(\dfrac{1}{5}x-5\right)^2}, & 25 < x \\ 1, & otherwise \end{cases}$$

Properties For two fuzzy sets A_1 and A_2 in X,

1. $A_1 \subseteq A_2 \; iff \; (\forall x \in X)\left(\mu_{A_1}(x) \leq \mu_{A_2}(x)\right)$.
2. $\mu_{A_1 \cup A_2}(x) = sup\{\mu_{A_1}(x), \mu_{A_2}(x)\}$
3. $\mu_{A_1 \cap A_2}(x) = inf\{\mu_{A_1}(x), \mu_{A_2}(x)\}$
4. $\mu_{A_1{}^c}(x) = 1 - \mu_{A_1}(x)$
5. If X is finite, the cardinality of the fuzzy set A_1 is calculated as

$$|A_1| = \sum_{x \in X} \mu_{A_1}(x)$$

Fuzzy Rough Set Theory:

Indistinguishability (for imperfect sets) and ambiguity (for fuzzy sets) are two separate concepts that can be found in datasets and have been successfully used in a variety of industries.

1. Fuzzy tolerance relation:

A fuzzy relation in X is a fuzzy set defined as Ry(x) = R (x, y) for every x belongs to X, the fuzzy set Ry is the R-foreset of y for each y in X. If R is a fuzzy relation that is symmetric and reflexive, then

R (x, x) = 1 and R (x, y) = R (y, x)

R is known as a fuzzy tolerance relation and Ry is known as the fuzzy tolerance class of y,

2. Fuzzy Triangular norm:

A triangular norm or t-norm T is an increasing, associative and commutative mapping from $[0,1] \times [0,1] \rightarrow [0,1]$ satisfying $T(1,x) = x, \forall\, x \in [0,1]$. A few widely used t-norms are: $T_M(x,y) = \min\{x,y\}$ and $T_L(x,y) = \max\{0, x+y-1\}$ (Lukasiewicz t-norm), for $x,\ y$ in $[0,1]$.

3. Fuzzy Implicator:

An implicator is any mapping $I : [0,1] \times [0,1] \rightarrow [0,1]$, satisfying $I(0,0) = 1$ and $I(1,x) = x, \forall\, x \in [0,1]$. Moreover, I needs to be decreasing in its first, and increasing in its second component. A few widely used implicators are- $I_M(x,y) = \max\{1-x, y\}$ (Kleene- Dienes implicator) and $I_L(x,y) = \min\{1, 1-x+y\}$ (Lukasiewicz implicator) $\forall\, x, y \in [0,1]$.

4. Fuzzy Lower and Upper Approximations:

The Fuzzy Rough Set Theory (FRST) is an alternate method for feature selection. It proposes to calculate the similarity between the objects using a fuzzy relation R in U, i.e., $R : U \times U \to [0, 1]$ is a mapping that assigns to each distinct pair of objects their corresponding degree of similarity. Given a set $X \subseteq U$ and a fuzzy similarity relation R, the lower and upper approximation of X by R can be calculated in several ways. A general definition is the following:

$$(R \downarrow X)(x) = \inf_{y \in U} I(R(x, y), X(y))$$

$$(R \uparrow X)(x) = \sup_{y \in U} T(R(x, y), X(y))$$

Now, positive region can be defined by:

$$\mu_{POS_P(Q)} = \sup_{x \in X} \mu_{R_P} \downarrow X$$

Here, $P \subseteq C, X \subseteq U$ and Q is the set of decision attributes.

Now, degree of dependency can be given by:

$$\gamma_P\ (Q) = \frac{|\mu_{POS_P(Q)}(x)|}{|U|} = \frac{\sum_{x \in X} \mu_{POS_P(Q)}(x)}{|U|}$$

Selection of fuzzy rough set-based features:

The search for specific reducts, or subsets of conditional qualities, is one of the core components of rough set theory for knowledge acquisition. Since attribute reduction is a crucial preprocessing step in data mining, which aims to extract a limited number of rules from a data set with a predetermined purpose, it is a significant topic in rough set theory. However, it is not well suited for real-valued information systems since it necessitates discretization, which causes some information to be lost. Dubois and Prade merged rough set and fuzzy set to create fuzzy rough set in order to address this flaw in rough set theory. Following that, a number of researchers offered several methods for fuzzy rough aided attribute reduction and used them to solve classification issues. These methods are covered in this section.

Equivalence classes in the rough set based feature selection can be extended to the fuzzy rough based feature selection by introducing a fuzzy similarity relation, R in U, which determines degree of similarity of two objects on R.

CONCLUSION AND FUTURE SCOPE:

In this work, we studied several research articles based upon the feature selection techniques using fuzzy similarity relations. First, we need to define a suitable and well defined fuzzy similarity relation between two entities of a given decision system. Then FRS using defined relation in the form of upper as well as lower approximations. Positive region can be found by taking supremum value among all membership grades of lower approximations. Here, we have used degree of dependency based approach for finding the reduct of given dataset.

REFERENCES:

1 Pawlak, Z. (2012). *Rough sets: Theoretical aspects of reasoning about data* (Vol. 9). Springer Science& Business Media.

2 Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, *8*(3), 338-353.

3 Dubois, D., & Prade, H. (1990). Rough fuzzy sets and fuzzy rough sets. *International Journal of GeneralSystem*, *17*(2-3), 191-209.

4 Dubois, D., & Prade, H. (1992). Putting rough sets and fuzzy sets together. In *Intelligent DecisionSupport* (pp. 203-232). Springer, Dordrecht.

5 R. Jensen and Q. Shen, "Tolerance-based and Fuzzy-Rough Feature Selection," 2007 IEEE International Fuzzy Systems Conference, 2007 , pp.16,doi:10.1109/FUZZY.2007.4295481.

6 Tanmoy Som, Shivam Shreevastava, Anoop Kumar Tiwari, Shivani Singh. "Fuzzy Rough Set Theory- Based Feature Selection", Wiley, 2020

7 Tsang, E.C.C., Degang Chen, D.S. Yeung, Xi-Zhao Wang, and J. Lee. "Attributes Reduction Using Fuzzy Rough Sets", IEEE Transactions on Fuzzy Systems, 2008.