Granular Soft Computing:

A Paradigm in Information Processing

Saroj K. Meher

Systems Science and Informatics Unit

Indian Statistical Institute, Bangalore

Granular computing (GrC): Outline

- Granular Spaces
- Fuzzy granulation
- Rough granulation
- Wavelet granulation
- Neural Granulation
- Basic issues
- Hybrid granular spaces
- Application to pattern classification
- Concluding remarks

Granular spaces

- Time/space
- Frequency
- Time-frequency
- Transformation
- Neural space

Historical notes

Main driving force of GrC

- Fuzzy set, and
- Rough set theories

However, the connections to other fields and the generality, flexibility, and potential of GrC have not been fully explored, such as

- Wavelet transform and others
- Neural networks

Concept of Flexibility Analysis

Fuzzy Sets and Information Granules
: Basic Concepts

Fuzzy Sets

FUZZY LOGIC—A BRIEF SUMMARY

Misconceptions about fuzzy logic.

Fuzzy logic is not fuzzy, In essence,

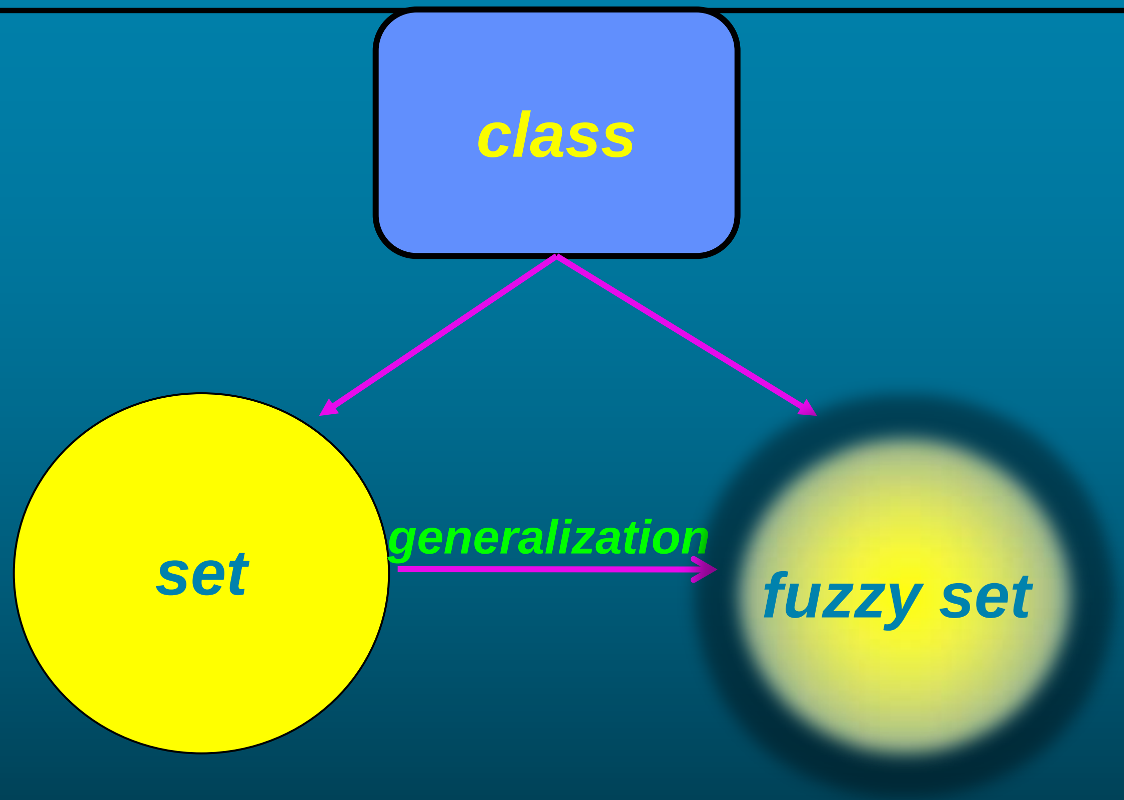
fuzzy logic is a precise logic of imprecision.

The point of departure in fuzzy logic—the nucleus of fuzzy logic, is the concept of a fuzzy set.

EXAMPLE: Age of a person

Young, Old

The concept of Fuzzy set (ZADEH 1965)



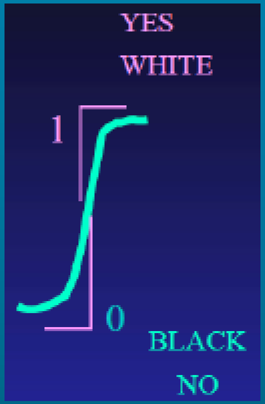
Informally, a fuzzy set, A, in a universe of discourse, U, is a class with a fuzzy boundary .

Continued….

- A set, A, in U is a class with a crisp boundary.
- A set is represented through association with a characteristic function c_A: U {0,1}
- A fuzzy set is represented through graduation, that is, through association with a membership function μ_A: U [0,1], with $\mu_A(u)$, uεU, representing the grade of membership of u in A.
- Membership in U is a matter of degree.
- In fuzzy logic everything is or is allowed to be a matter of degree.
- Fuzzy logic has nothing to do with randomness (probability),
- In essence, it deals with possibility called possibility theory.
- Mathematical objects that behave like fuzzy sets exist in probability theory. It does not mean that fuzziness is reducible to randomness.

Fuzzy Sets and Flexibility

FUZZY SETS



- Classical set : $\mu \in \{0,1\}$ Hard
- Fuzzy set : $\mu \in [0,1]$ Soft

$\mu A(x)$: degree of belonging of x to A or degree of possessing some imprecise property represented by A

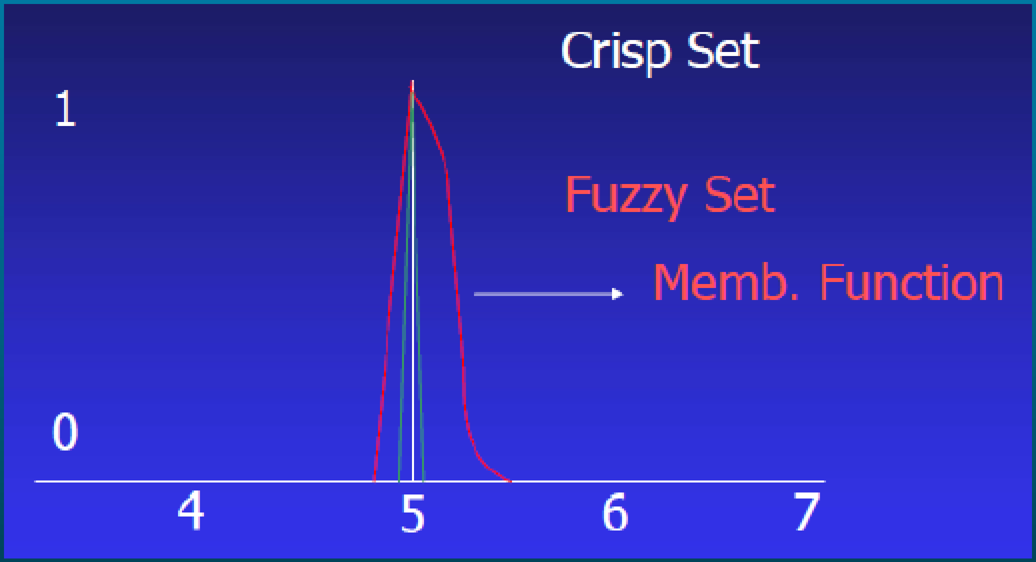
Example : tall man, long street, large number, sharp corner, very young, skin colour etc.

- Fuzzy set is a Generalization of classical set theory

⇒ Greater flexibility in capturing faithfully various aspects of incompleteness or imperfection in a situation.

Fuzzy Sets and Flexibility

Meeting at 5 PM



Membership Function: Context Dependent

Fuzzy Sets and Flexibility

FUZZY SETS

Flexibility of fuzzy set theory is associated with the Concept of **μ**

μ : A measure of compatibility of an object with the concept represented by fuzzy set.

- μ_{TALL} = 0.3 means

  Compatibility of some one with the set ``TALL´´ NOT the prob. that some one is TALL

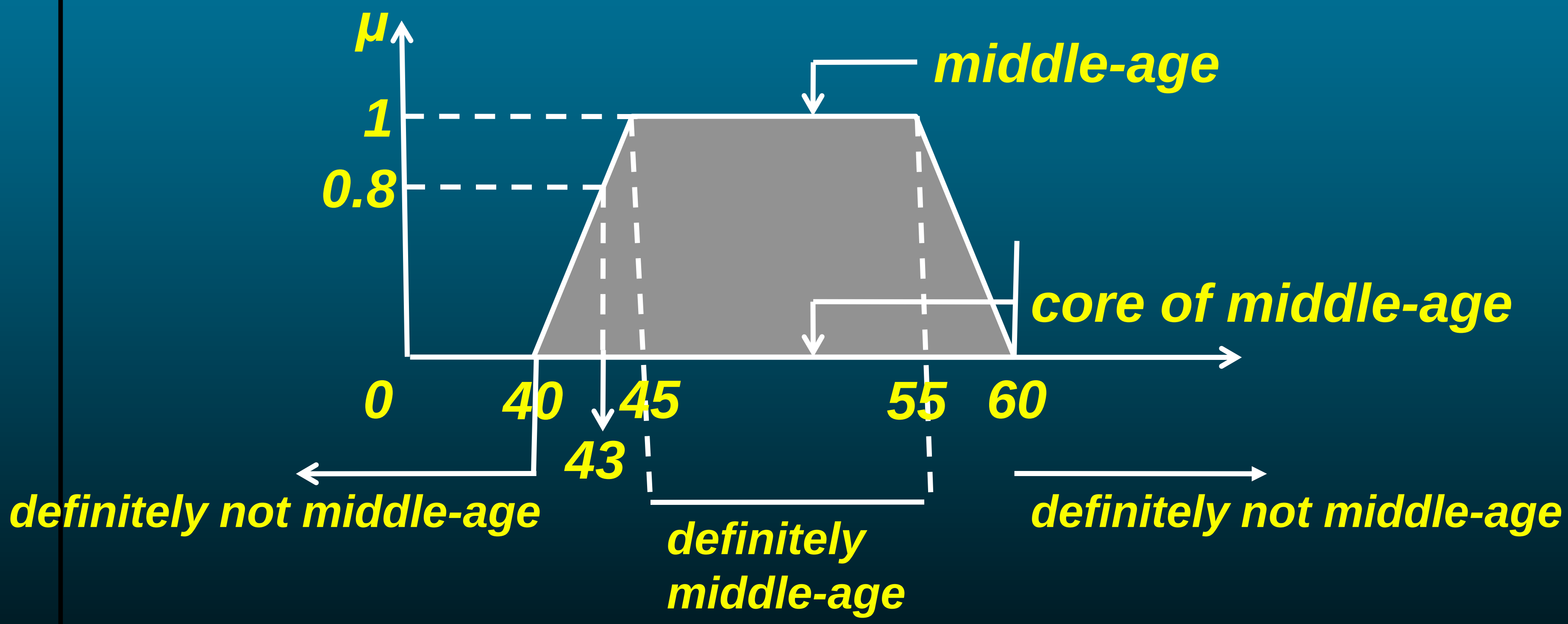
  i.e., 0.3 is the extent to which the concept ``TALL´´ must be stretched to fit him

- As μ ↑, Amount of Stretching Concept ↓
- FUZZINESS IS ANALOGOUS TO ELASTICITY

EXAMPLE—MIDDLE-AGE, IMPRECISION OF MEANING

Imprecision of meaning = elasticity of meaning

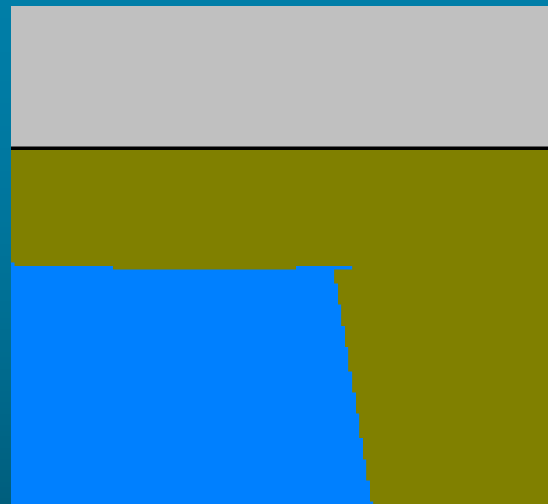
Elasticity of meaning = fuzziness of meaning



Fuzzy information granulation

Granules

- Crisp



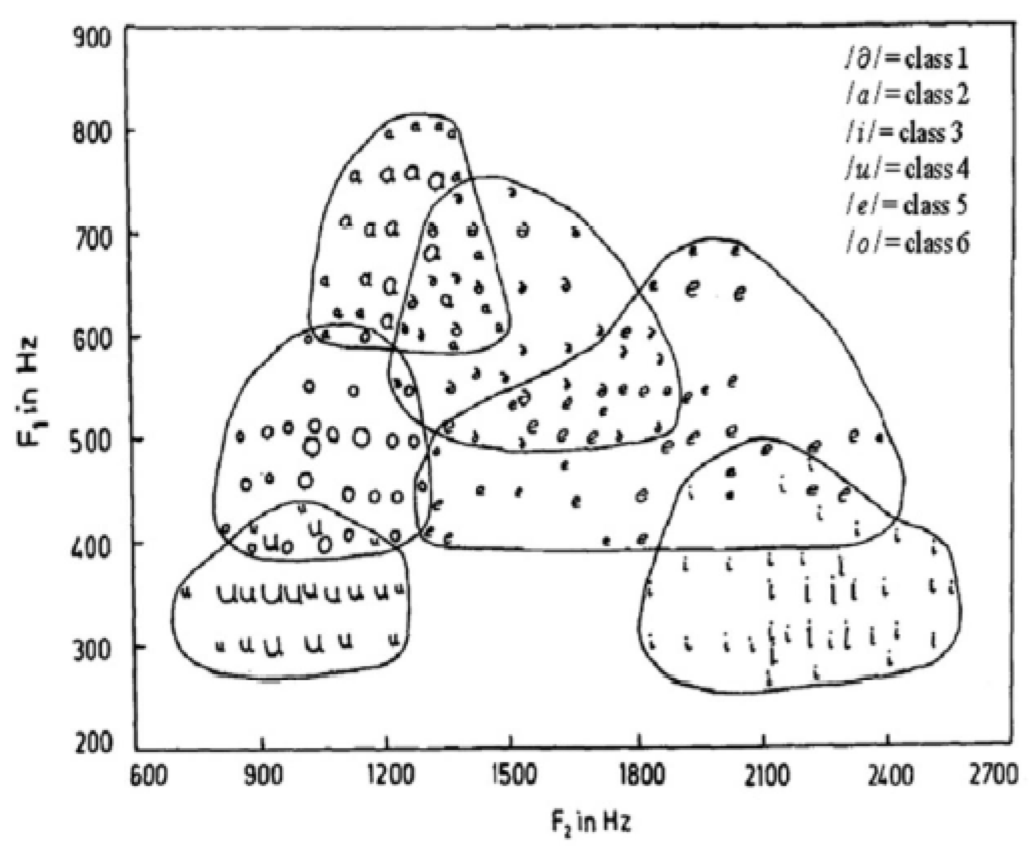
- Fuzzy

Fuzzy Information Granulation (FIG)

FIG deals with

- Imprecise representation of information,
- Problems having insufficient information.

Real Data set



Scatter plot of VOWEL data in F1-F2 plane

Fuzzy information granulation

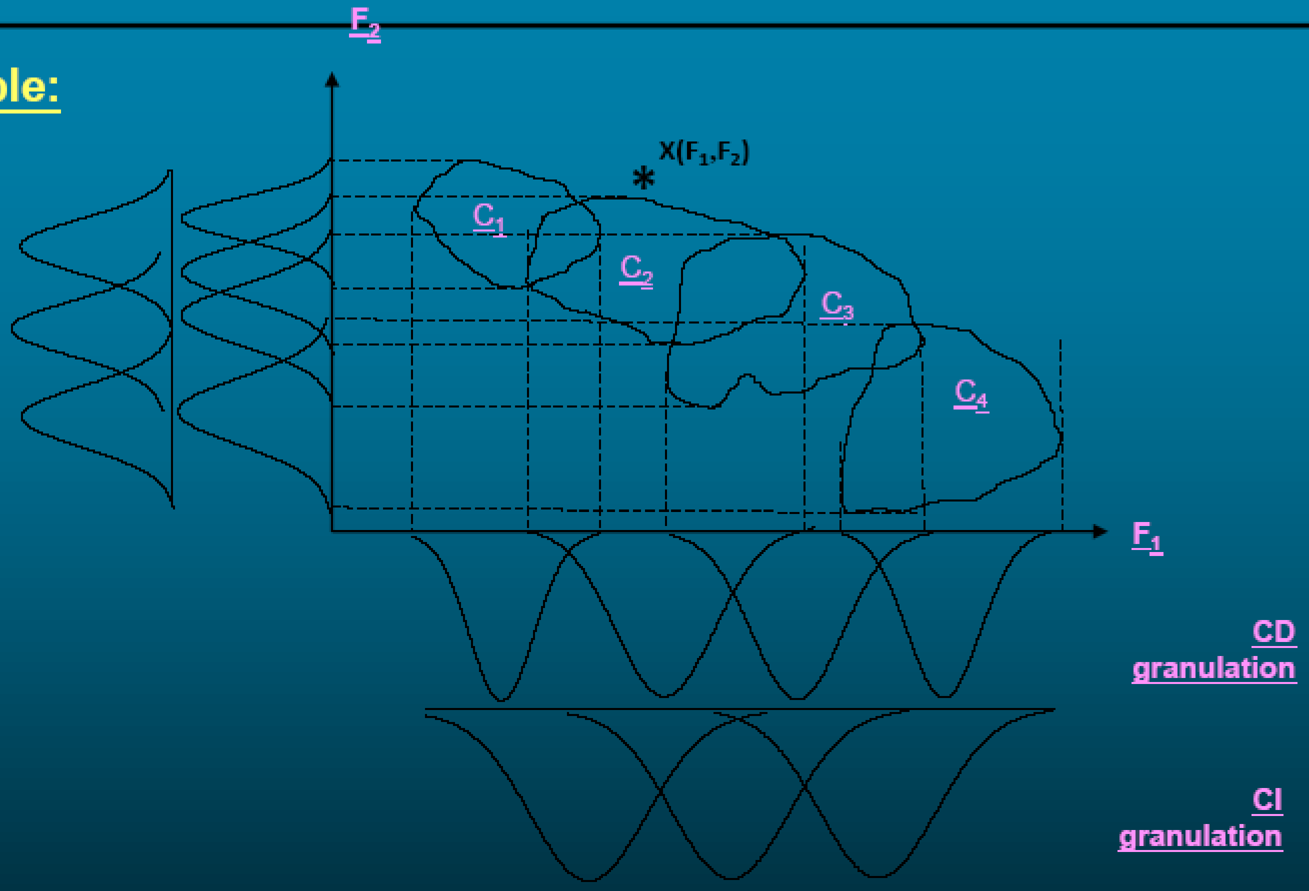
Fuzzy Granules

- Class-dependent (CD)
  - Each feature explores its class belonging information to different classes,
  - Features are described by the fuzzy sets (equal to the number of classes) that characterizes corresponding number of fuzzy granules along the axis.

- Class-independent (CI)
  - With CI granulation, each feature is described with some defined number fuzzy membership functions over the whole space,
  - The generated granules thus, does not take care of the class belonging information of features to different classes.

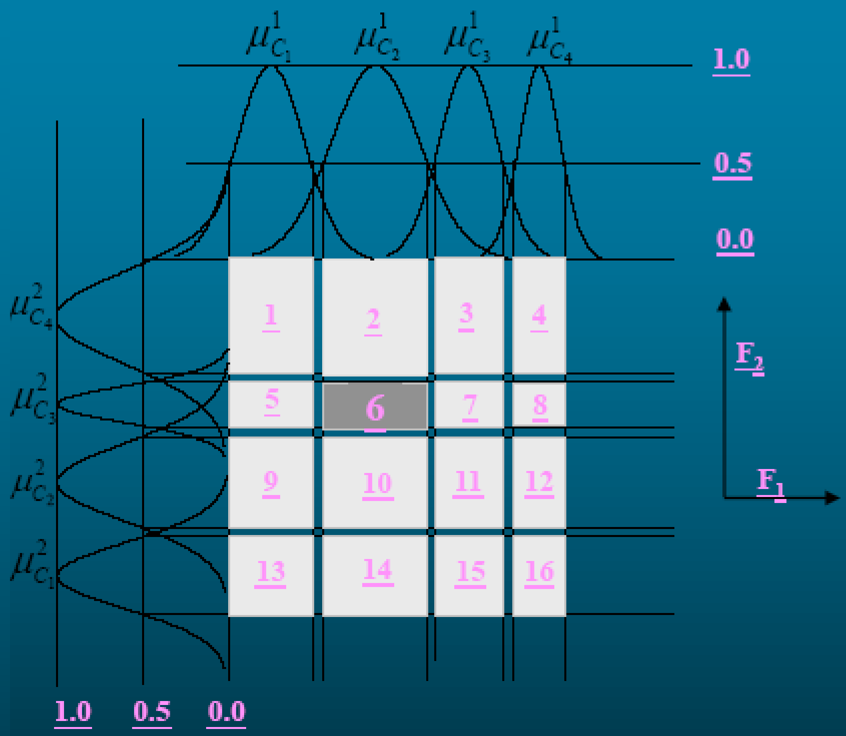
Fuzzy information granulation

Example:



Fuzzy granulation of features F_1 and F_2 that characterizes granules for four overlapping classes.

Fuzzy information granulation



Example of Granules Generation:

Each feature *F* is represented by *C* (= number of classes) [0,1]-valued membership functions (π-type in present case) representing *C* fuzzy sets or characterizing *C* fuzzy granules along each axis; there by constituting C^n fuzzy granules in an *n*-dimensional feature space.

Figure: Generation of granules from class-wise (class dependent) fuzzy representation of the features F_1 and F_2. The figure represents the granules for four overlapping classes. The shaded regions (16 nos.) indicate the granules. For example the region (granule no 6) indicates a crisp granule obtained by α-cuts (α = 0.5 in present case) on the $\mu^1_{C_2}$ and $\mu^2_{C_3}$. The granules shape / size are variable in nature and depend on the overlapping nature of classes and class-wise feature distribution.

Fuzzy information granulation

F with *n-numeric features* ($F = [F_1, F_2, ..., F_n]$)

CI: Granulate the feature values that characterizes them in terms of some combination of membership values in the linguistic property sets *low, medium* and *high*.

$$\mathbf{F} = [\mu_{low(F_1)}(F), ..., \mu_{high(F_n)}(F)]$$

CD: Each feature is described in terms of its fuzzy membership values corresponding to L (total number of classes) linguistic fuzzy sets.Thus, an n-dimensional pattern vector is expressed as(n x L)-dimensional vector and is given by

$$\begin{aligned}\mathbf{F} = [&\mu_1^1(F_1), \mu_2^1(F_1), \ldots, \mu_c^1(F_1), \ldots, \mu_{\mathbf{L}}^1(F_1);\\ &\mu_1^2(F_2), \mu_2^2(F_2), \ldots, \mu_c^2(F_2), \ldots, \mu_{\mathbf{L}}^2(F_2);\\ &\mu_1^n(F_n), \mu_2^n(F_n), \ldots, \mu_c^n(F_n), \ldots, \mu_{\mathbf{L}}^n(F_n)] \quad (c = 1, 2, \ldots, \mathbf{L}),\end{aligned}$$

Concept of Flexibility & Uncertainty Analysis

Rough Sets and Information Granules
: Basic Concepts

Rough Sets

Historical notes

Rough sets perspectives:

1982, Pawlak introduced the notion of rough sets.

1998, the GrC view of rough sets was discussed by many researchers (Lin, Pawlak, Skowron, Y.Y. Yao, and many more).

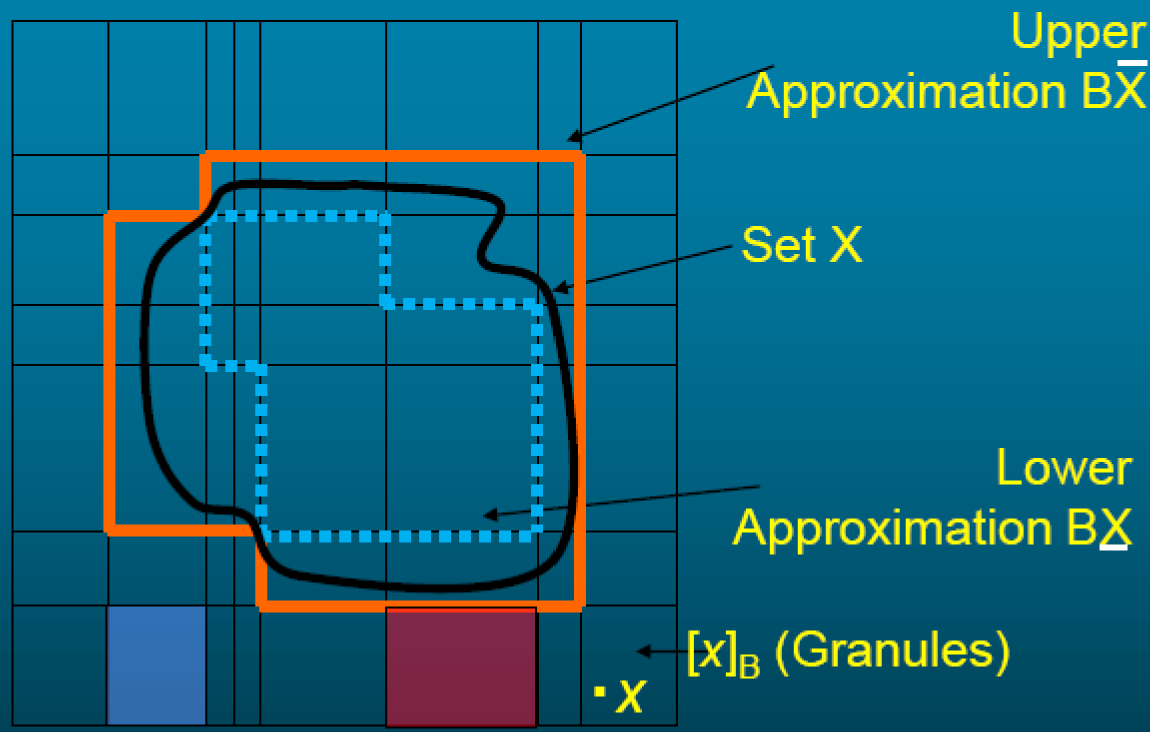
Rough set theory can be viewed as a concrete example of GrC.

Rough Sets

- Offer mathematical tools to *discover* hidden patterns in data.
- Fundamental principle of a rough set-based learning system is to *discover redundancies* and *dependencies* between the given features of a data to be classified.
- Approximate a given concept both *from below* and *from above*, using *lower* and *upper approximations*.
- Rough set learning algorithms can be used to obtain rules in IF-THEN form from a *decision table*.
- Extract Knowledge from data base (decision table w.r.t. objects and attributes → remove undesirable attributes (knowledge discovery) → analyze data dependency → minimum subset of attributes (reducts))

Rough Sets

$\Omega_B \subseteq U$



$[x]_B$ = set of all points belonging to the same granule as of the point x in feature space Ω_B.

⇨ $[x]_B$ is the set of all points which are *indiscernible* with point x in terms of feature subset B.

Rough Sets

Approximations of the set $X \subseteq U$ w.r.t feature subset B

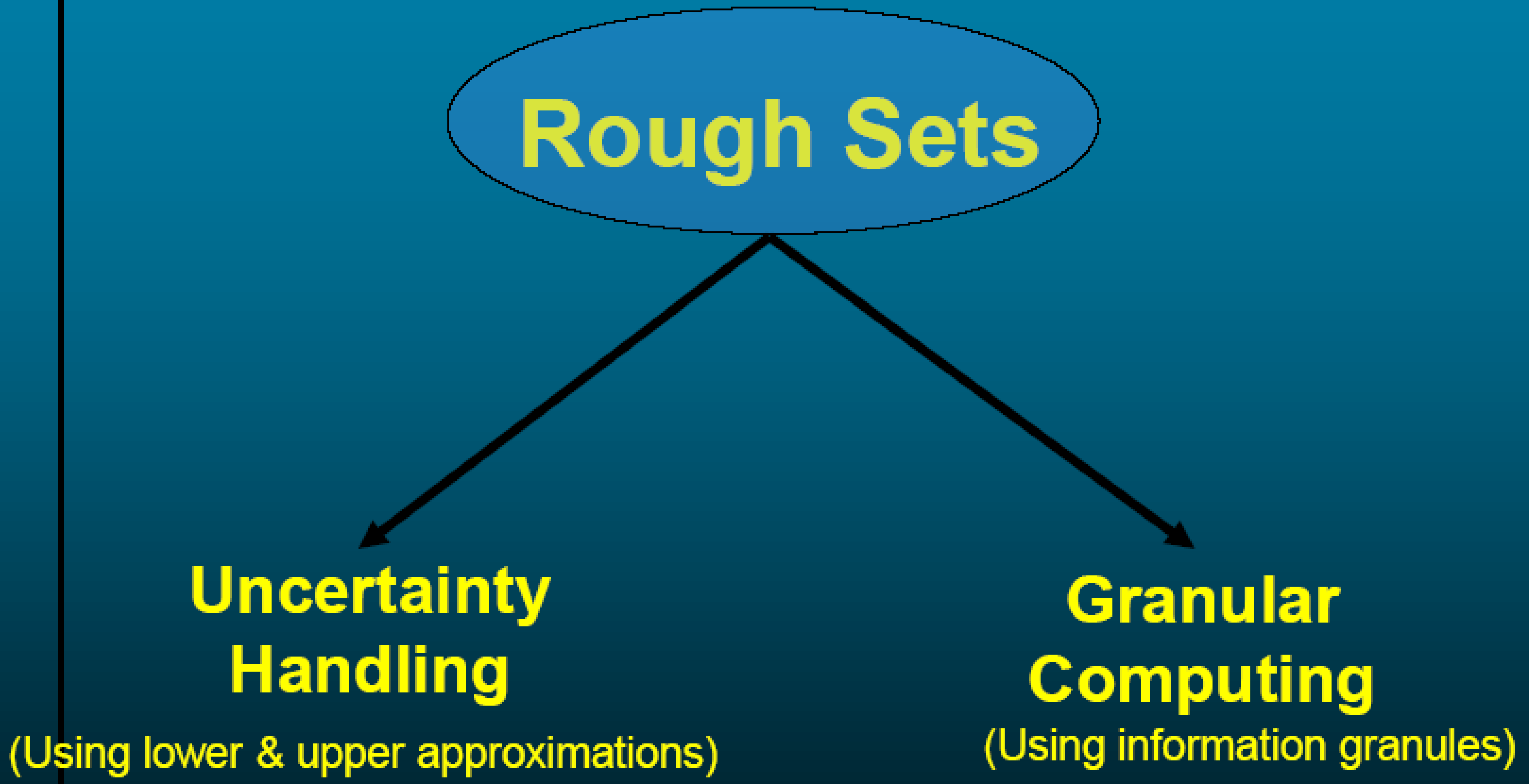
B-lower: $\underline{B}X$ = $\{x \in U : [x]_B \subseteq X\}$ Granules definitely belonging to *X*

B-upper: $\overline{B}X$ = $\{x \in U : [x]_B \cap X \neq \phi\}$ Granules definitely and possibly belonging to *X*

If $B\overline{X} = B\underline{X}$, *X* is *B-exact* or *B-definable*

Otherwise it is *Roughly definable*

Rough Sets


Rough Sets
Uncertainty Handling
(Using lower & upper approximations)
Granular Computing
(Using information granules)

Granular Computing: Computation is performed using *information granules* and not the data points (objects)

- Information compression
- Computational gain

Rough granulation

Pawlak's rough set (PaRS)

Neighborhood rough set (NRS)

Characteristics

NRS covers the space, and PaRS partitions,

Number of granules

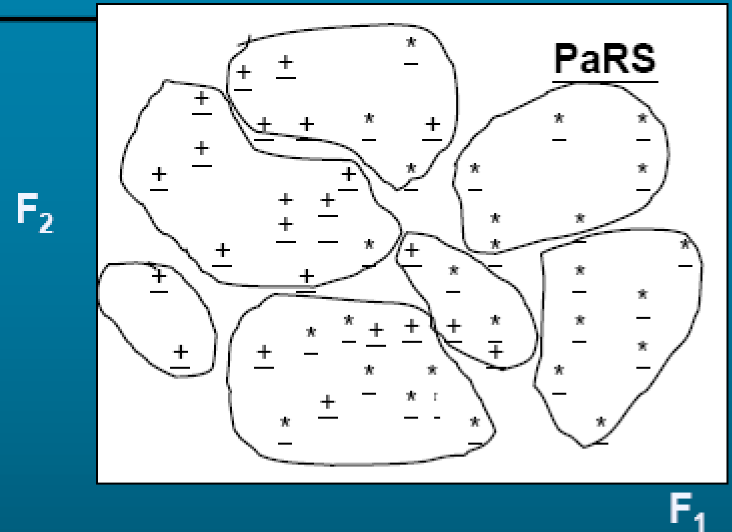
== number of patterns (NRS)

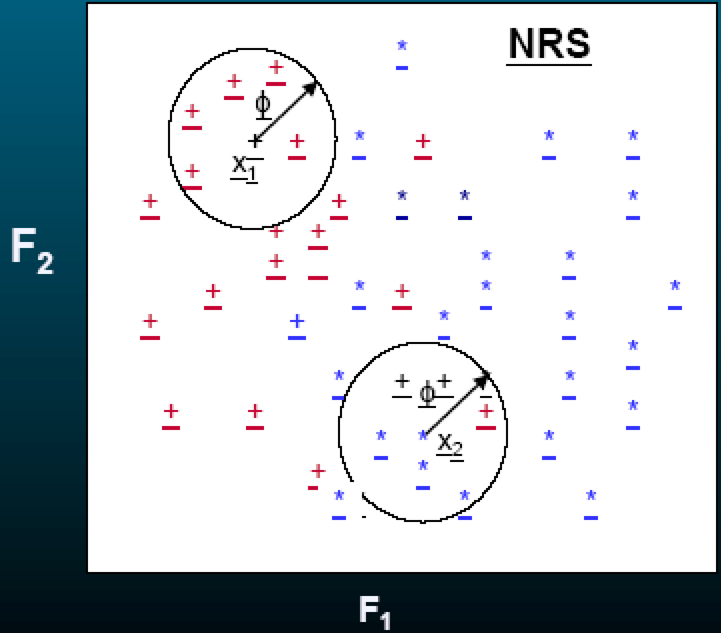
<= number of patterns (PaRS),

Granules shape and size
(granularity of data analysis)

controlled (NRS) with two parameters, such as
shape (Δ) and size (ϕ)

not under control (PaRS)





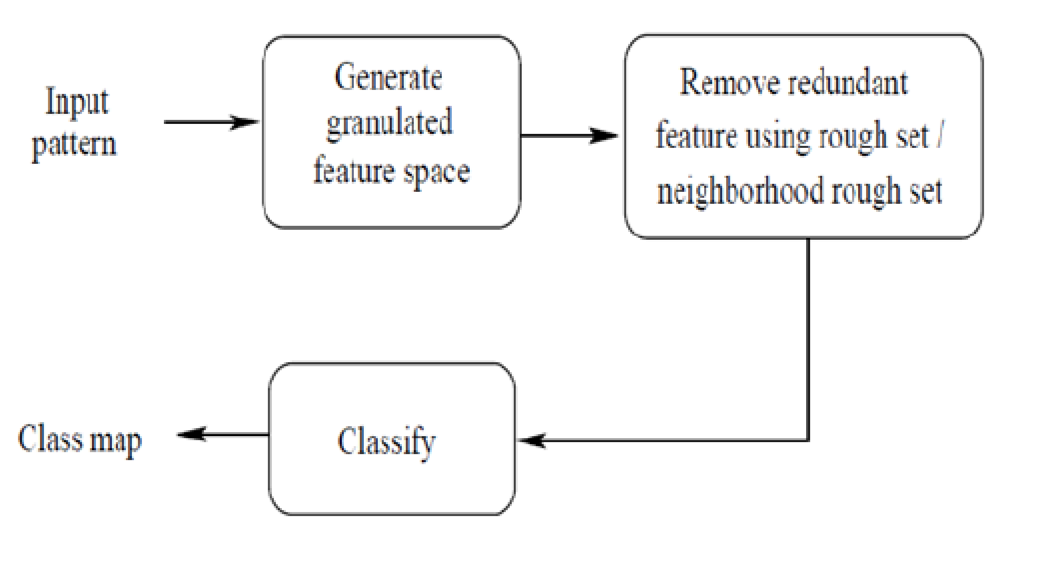
Two neighborhood granules centered at samples x_1 and x_2 in F1- F2 feature space. ϕ is the radius of the granules and $\Delta(xi, xj) \leq \phi$. Granules' shape & size are determined by p-norm distance function (Δ) and threshold ϕ.

Neighborhood rough granulation

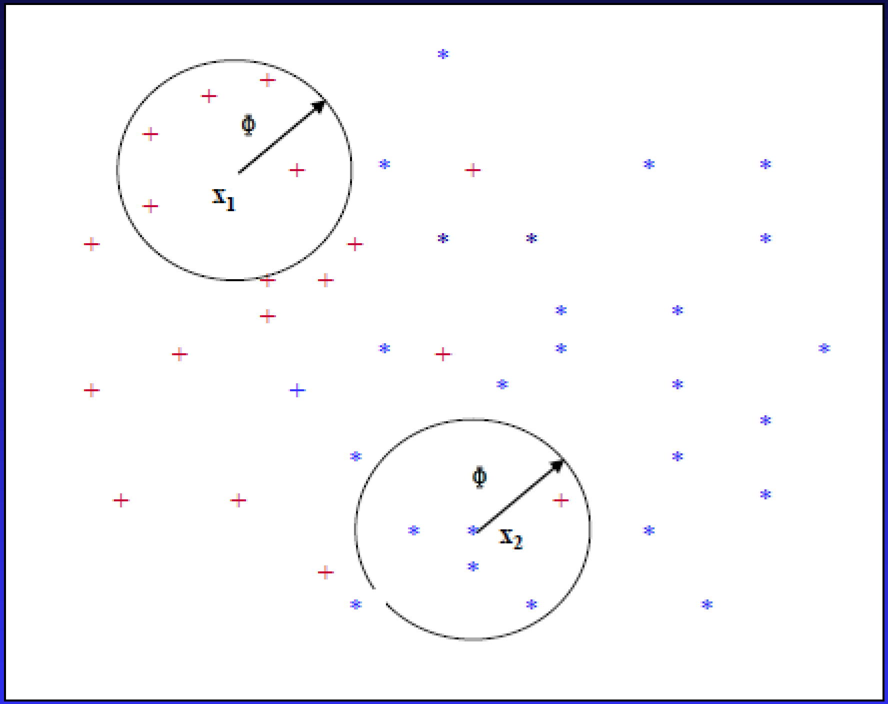
Advantages

- Significance of features vary with the granular size and granularity levels,
- Selects different feature subsets with the change of neighborhood shape and size,
- No need for feature value discretisation,
- Explores local/contextual information in an improved manner.

Classification model



Neighborhood Granule Generation for two overlapping classes

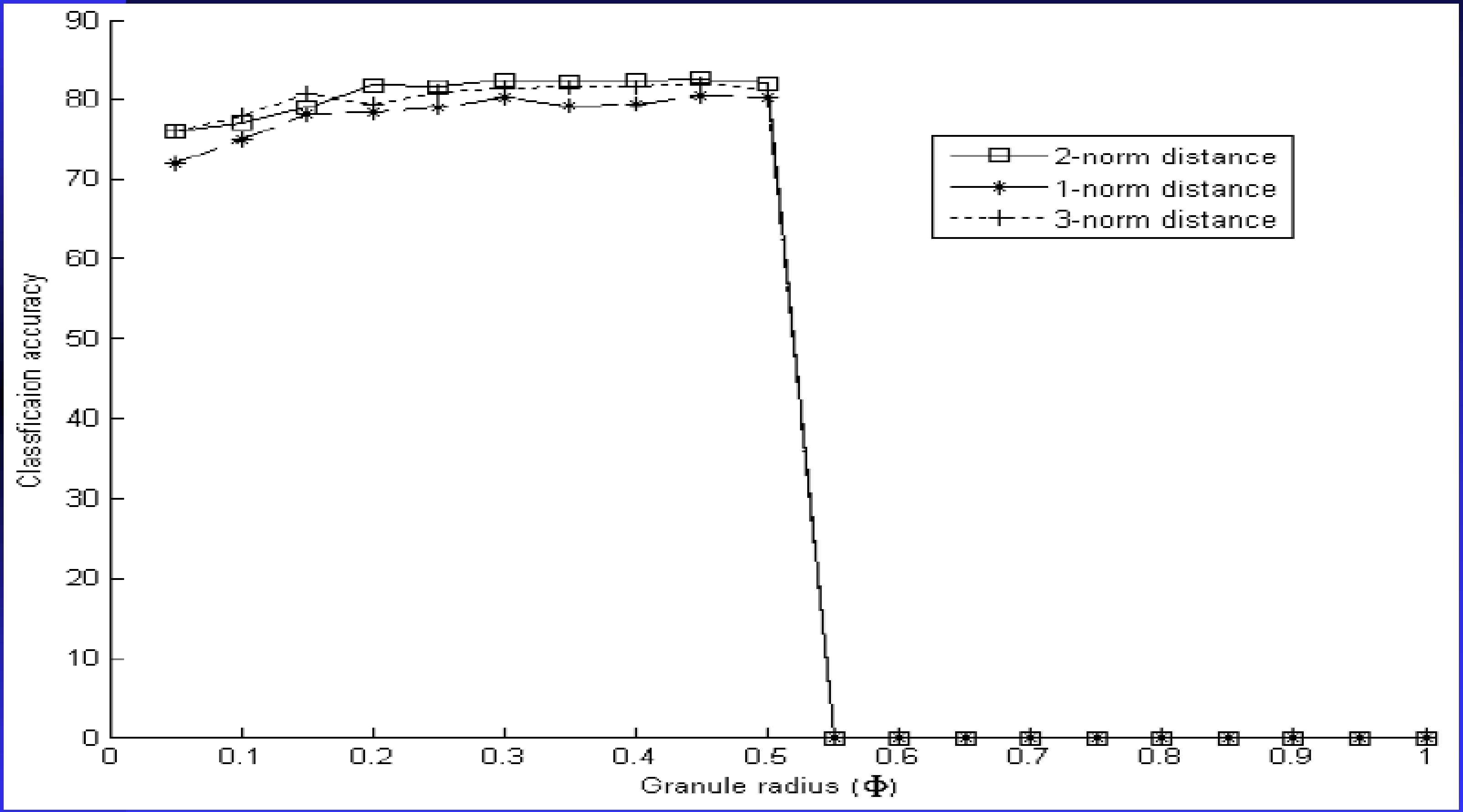


F_2

F_1

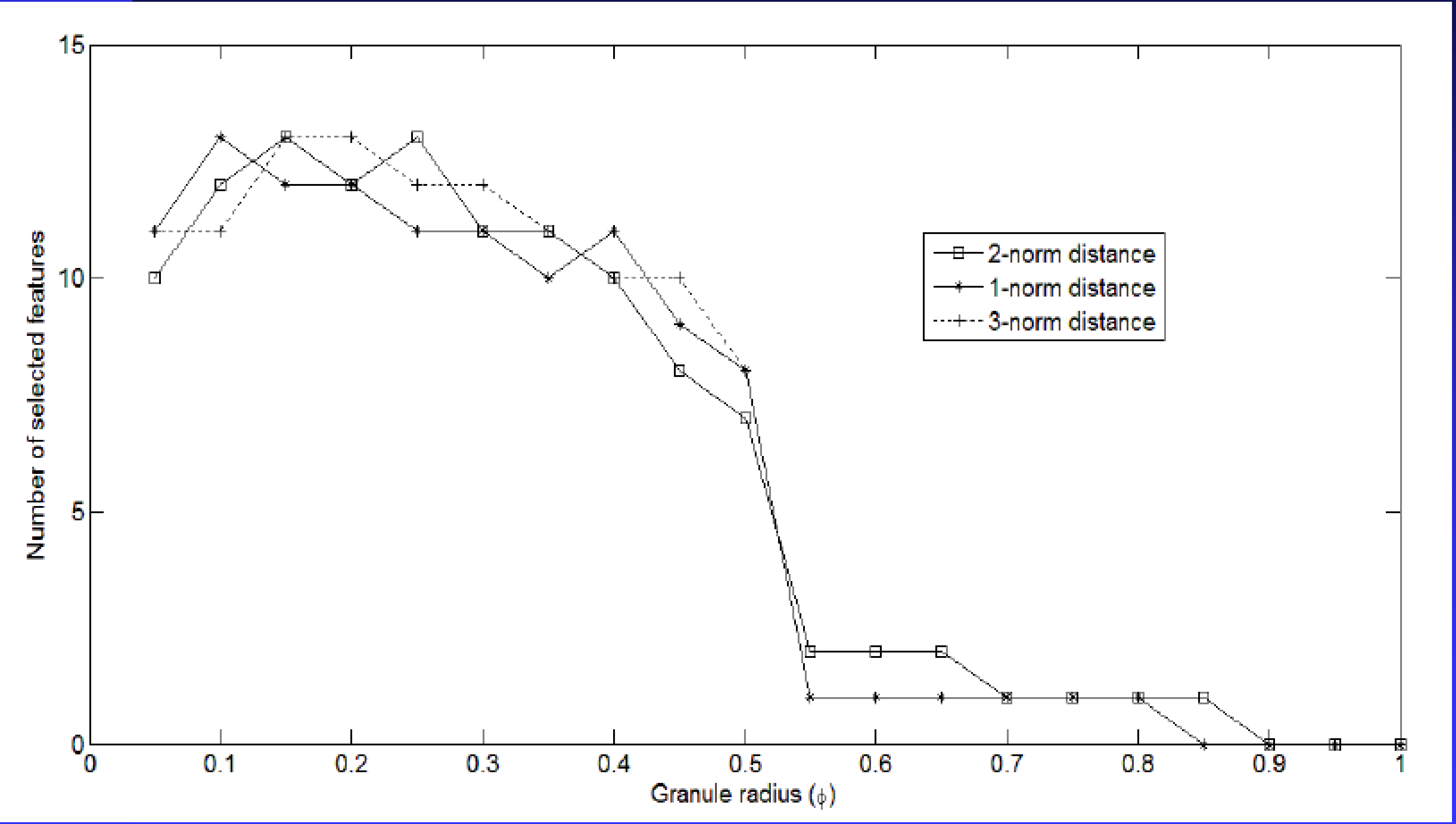
Two neighborhood granules centered at samples x_1 and x_2 in **F1- F2** feature space. ϕ is the radius of the granules and $\Delta(xi, xj) \leq \phi$. Granules' shape & size are determined by p norm distance function (Δ) and threshold ϕ.

Variation of classification accuracy with granule radius ϕ for three p-norm distances for model 5 and VOWEL data (Train set = 20%)



- Optimum $\phi = 0.45$
- Beyond 0.5, NRS based model can't select relevant features to distinguish patterns, since possibility of possessing irrelevant/ contradictory feature information by granules increases

Variation of number of selected features for model 5 with granule radius Φ for three p-norm distances and VOWEL data (Train set =20%)



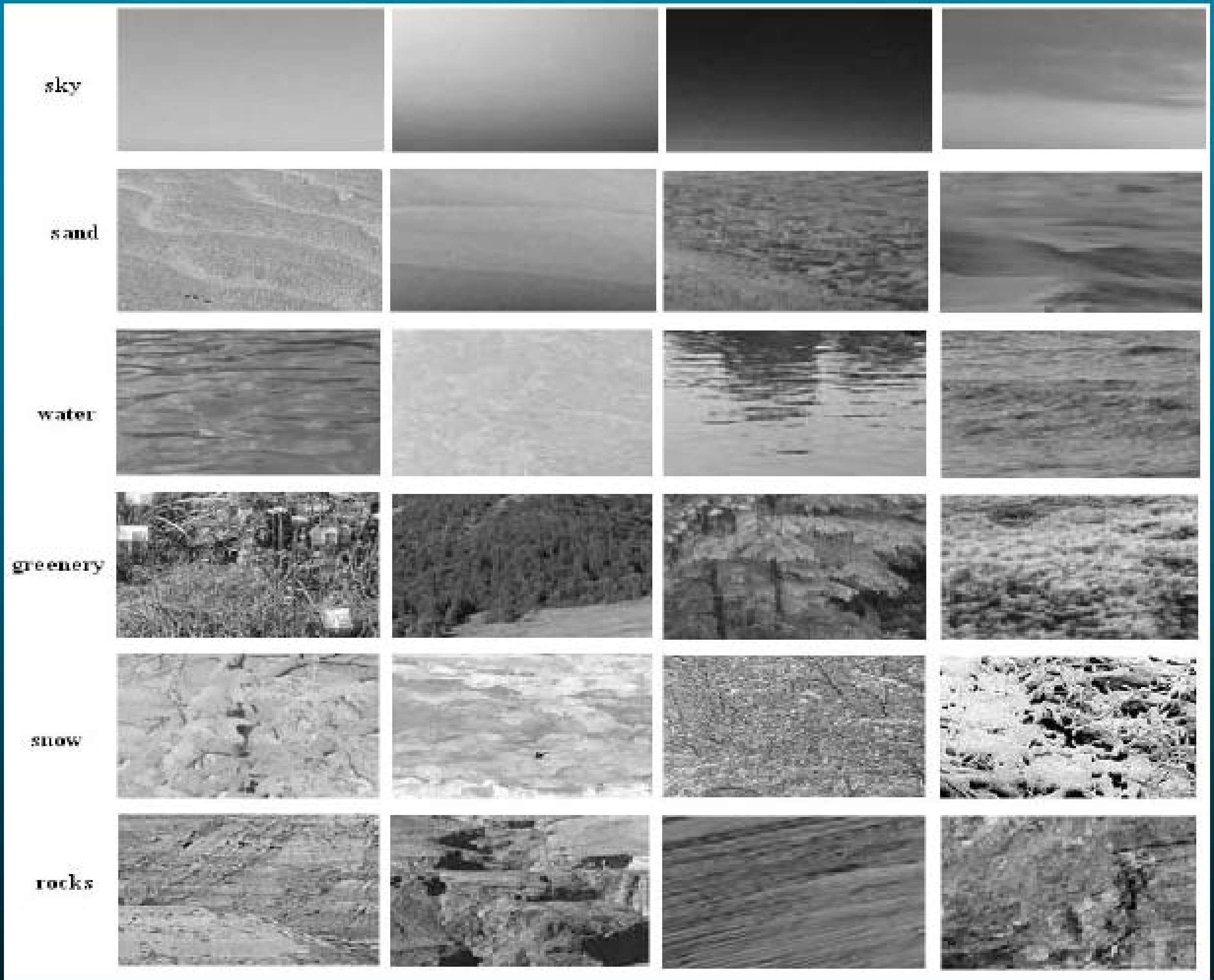
After $\phi = 0.5$, minimum reducts are not available

Description of the data set

Scene-region data

- Name of the dataset: scene-regions of images (collected from Google and Flickr image data bases)
- Number of classes: 6
- Number of features: 39 (13 from each of the red-green-blue colors)
  - 6 texture features such as mean, standard deviation, smoothness, skewness, uniformity and entropy, and
  - 7 invariant moments features. These features are insensitive to translation, scale, change, mirroring and rotation.
- Number of patterns: 700

Scene-region data set



Four examples of scene-regions of natural images.

Description of the data set

Image based *spam-ham* data

- Name of the dataset: image based *spam-ham*
- Number of classes: 2
- Number of original features: 15
  - number of unique colors over the whole image,
  - number of text region pixels and color saturation over the whole image, and
  - color heterogeneity, color smoothness, mean and standard deviation of the images in each of the red-green-blue bands
- Number of patterns: 1800

Spam-Ham data set

Two examples of image based *spam-ham* data set.

Results (Rough-fuzzy granulation)

Performance comparison of models to justify the use of

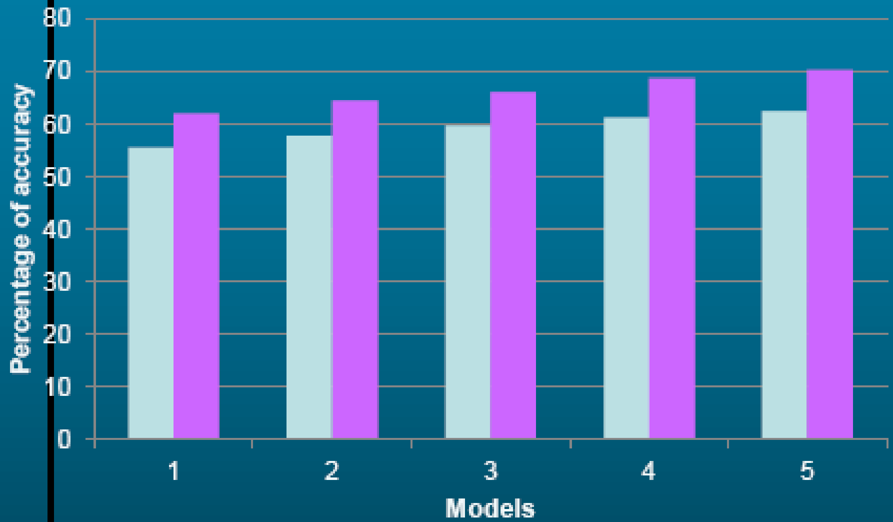
- **Granulated feature space,**
- **Class-dependent (CD) fuzzy granulation,**
- **Neighborhood rough sets (NRS) based feature selection, and**
- **Synergistic integration of the merits of both fuzzy granulation and the theory of NRS.**

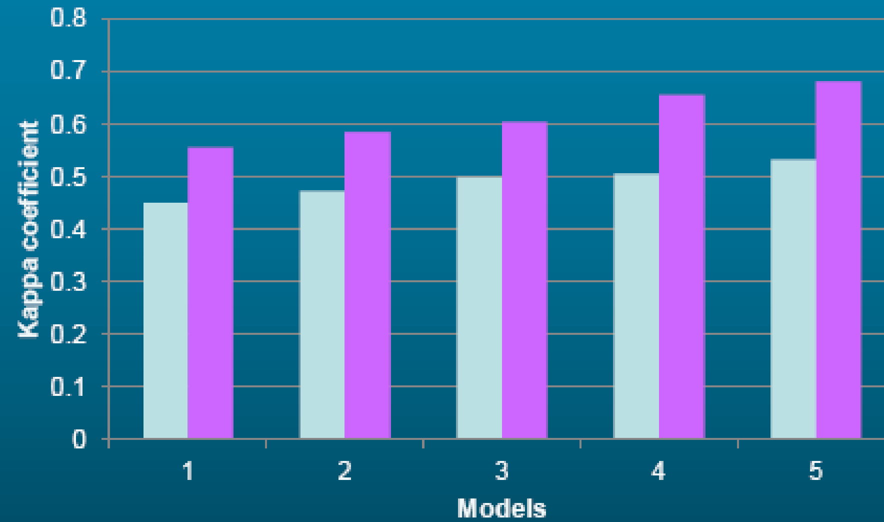
Five different combinations of classification models using rough-fuzzy granular feature space and feature selection methods

- model 1 (F1) : k-nearest neighbor (k-NN with k=1) classifier,
- model 2 (F2): CI fuzzy granulation + k-NN (with k=1) classifier,
- model 3 (F3): CD fuzzy granulation + k-NN (with k=1) classifier,
- model 4 (F4): CD fuzzy granulation + PaRS based feature selection + k-NN (with k=1) classifier,
- model 5 (F5): CD fuzzy granulation + NRS based feature selection + k-NN (with k=1) classifier.

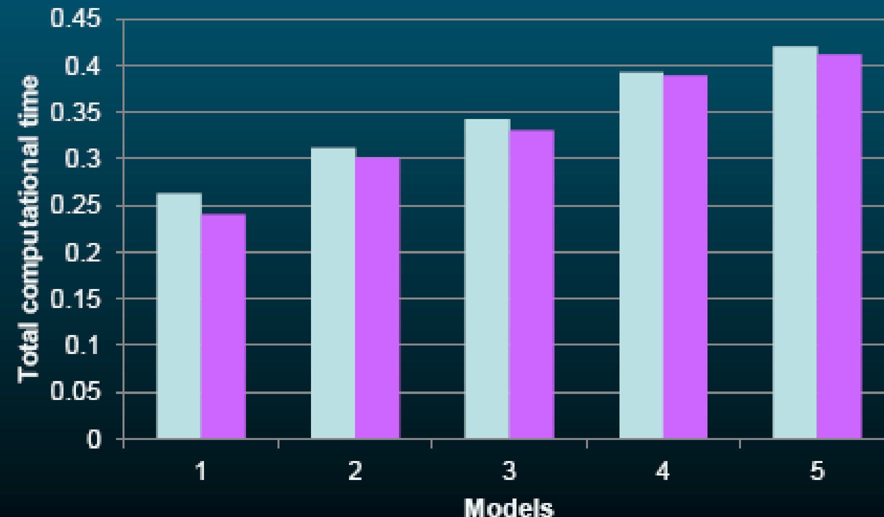
Performance comparison of rough-fuzzy granulated models using 1-nn classifier with scene-region data for 20% and 50% training sets.

20% training set   50% training set

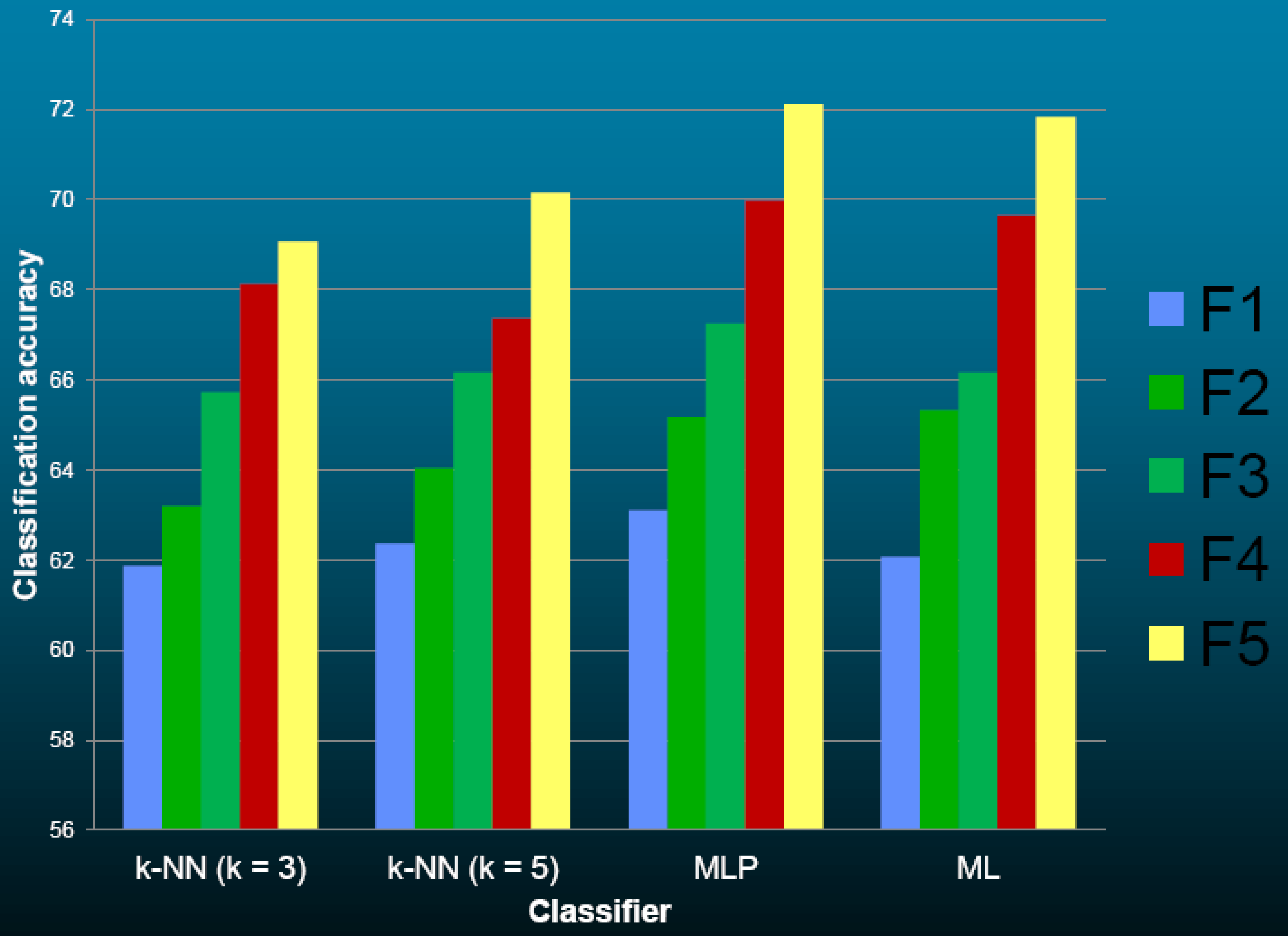




Note: All the simulations are done in MATLAB (Matrix Laboratory) environment in Pentium-IV machine with 3.19 GHz processor speed.

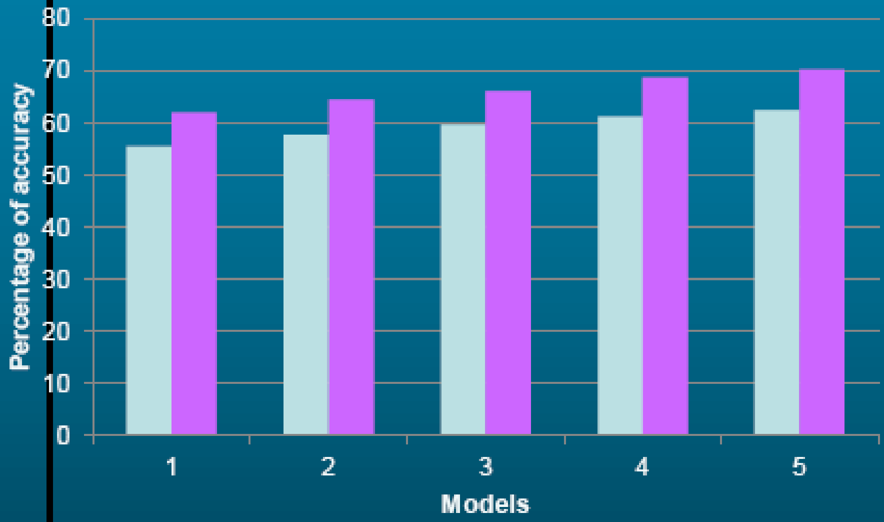


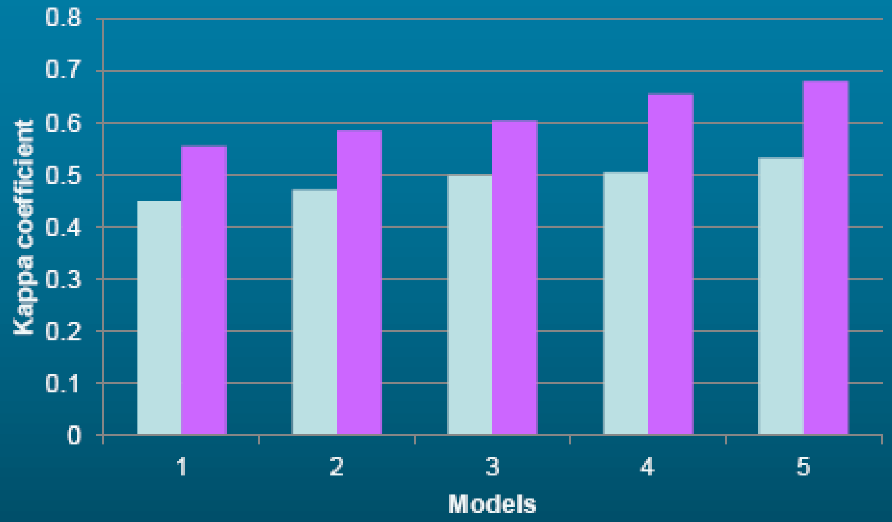
Classification accuracies (PA) of rough-fuzzy models with different classifiers at 50% training set for scene-region data set



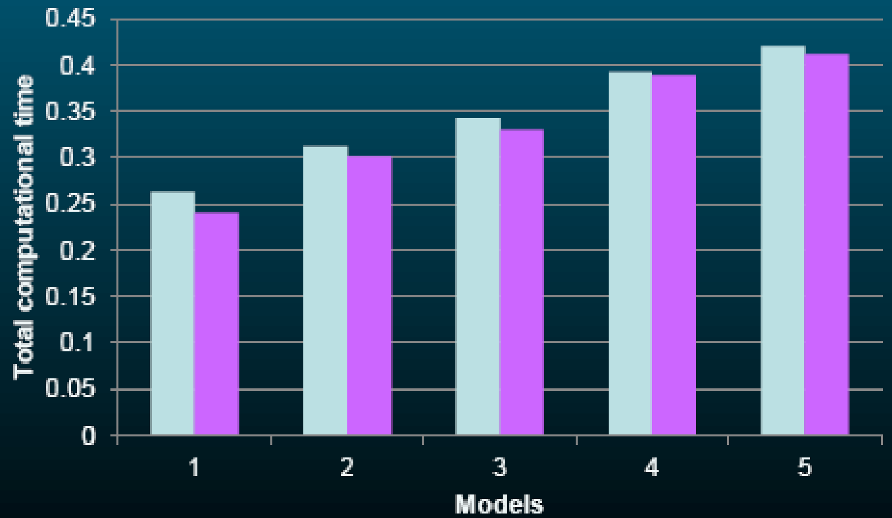
Performance comparison of rough-fuzzy granulated models using 1-nn classifier with *spam-ham* data for 20% and 50% training sets.

20% training set   50% training set

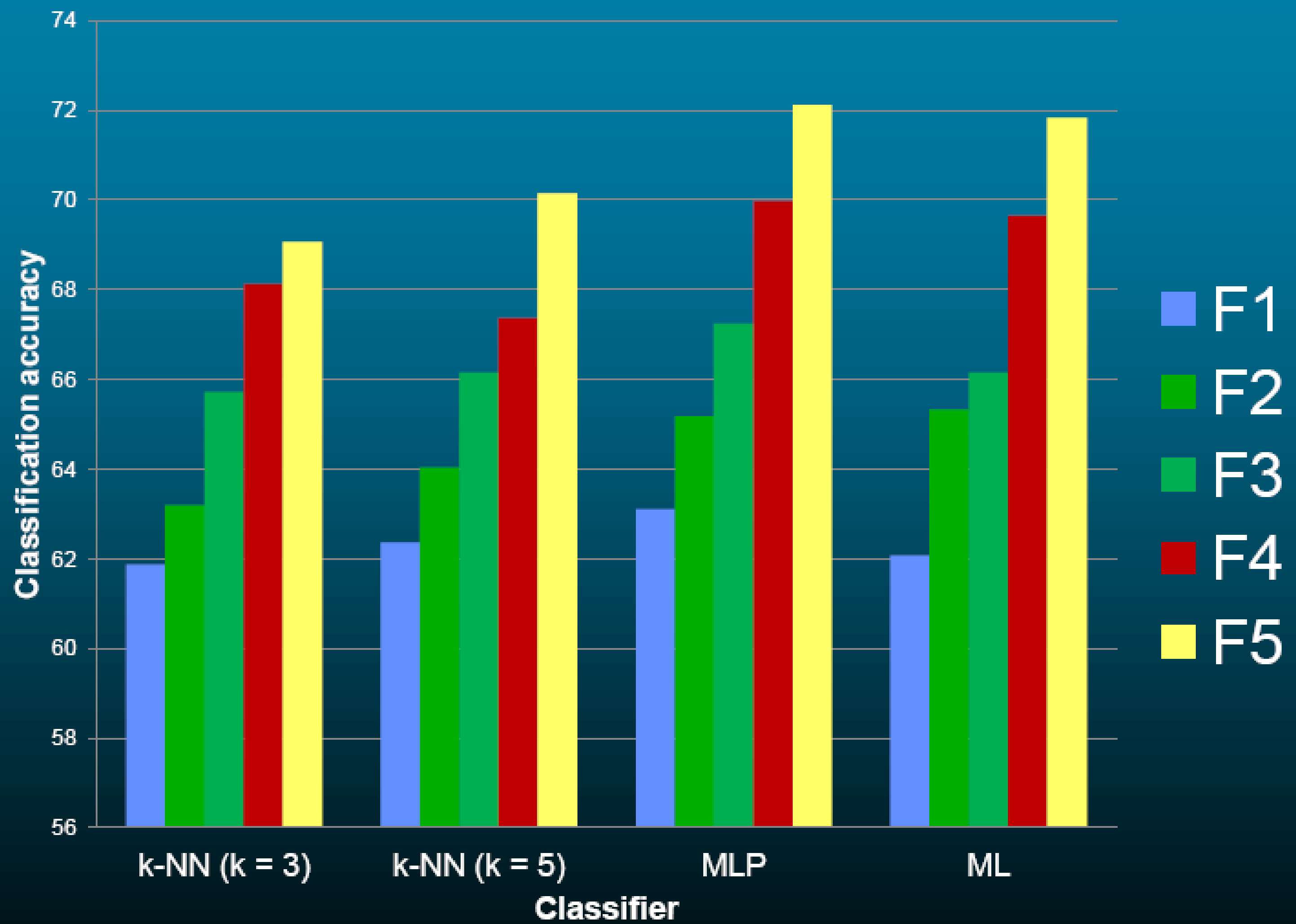




Note: All the simulations are done in MATLAB (Matrix Laboratory) environment in Pentium-IV machine with 3.19 GHz processor speed.



Classification accuracies (PA) of rough-fuzzy models with different classifiers at 50% training set for *spam-ham* data



Remote sensing data set

Six classes for both IRS 1A and SPOT image data :

pure water (PW), turbid water (TW), concrete area (CA), habitation (HAB), vegetation (VEG) and open spaces (OS).

Information of RS images :

IRS 1A

- size 512 x 512
- spatial resolution of 36.25 m x 36.25 m and wavelength range of 0.45-0.86 µm
- No: of bands: 4 (blue, green, red and near infrared)

SPOT

- size 512 x 512
- spatial resolution of 20 m x 20 m and wavelength range of 0.50-0.89 µm.
- No: of bands: 3 (green, red and near infrared)

Remote sensing data set

(a)

(b)

(a) IRS-1A (band-4) enhanced image, and (b) SPOT (band-3) enhanced image

Results with remote sensing data set

(a)

(b)

Classified IRS-1A images with (a) model 1 and (b) model 5 (proposed model).

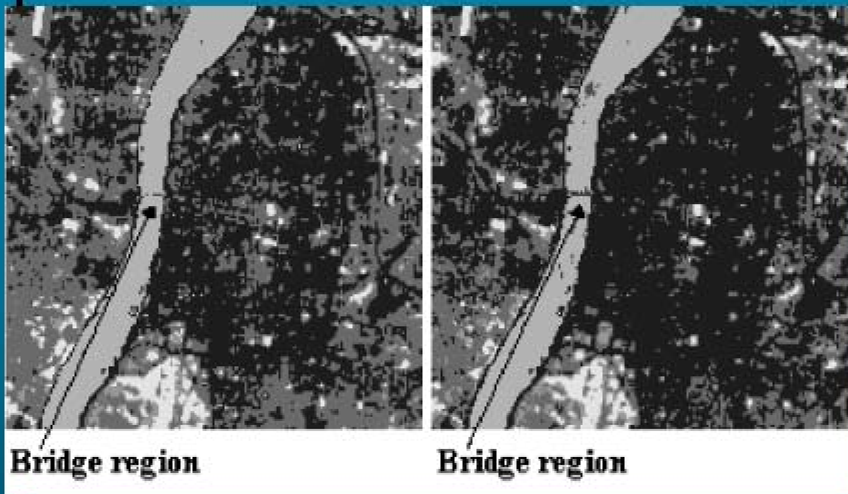
Results with remote sensing data set

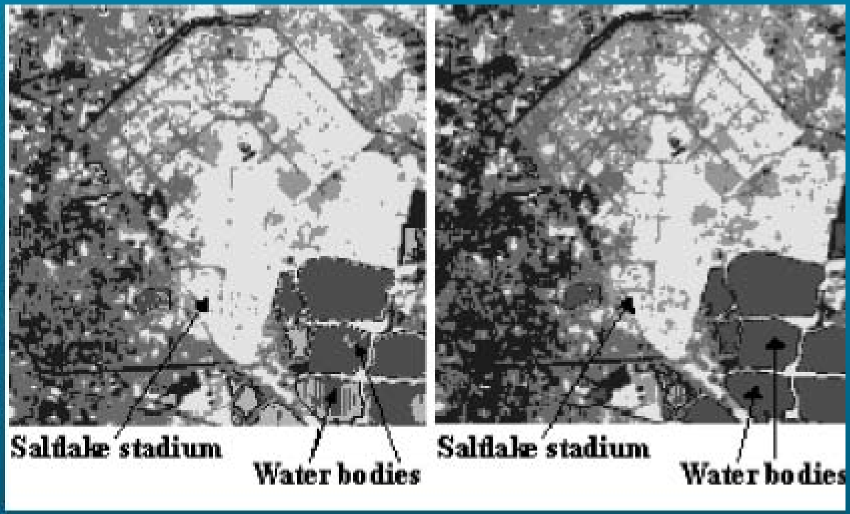
(a)

(b)

Classified SPOT images with (a) model 1 and (b) model 5 (proposed model).

Results with remote sensing data set

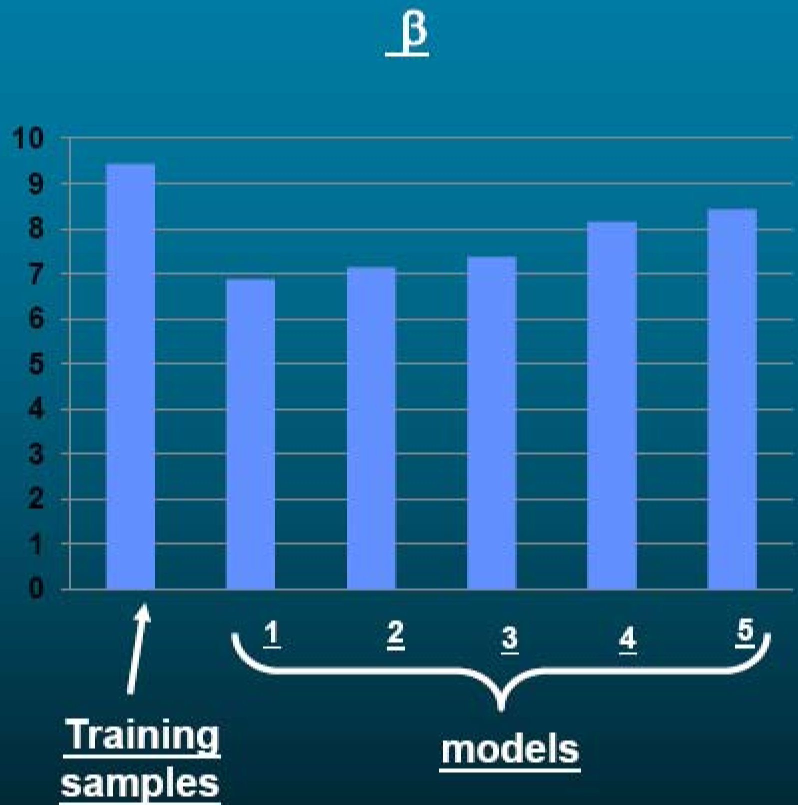




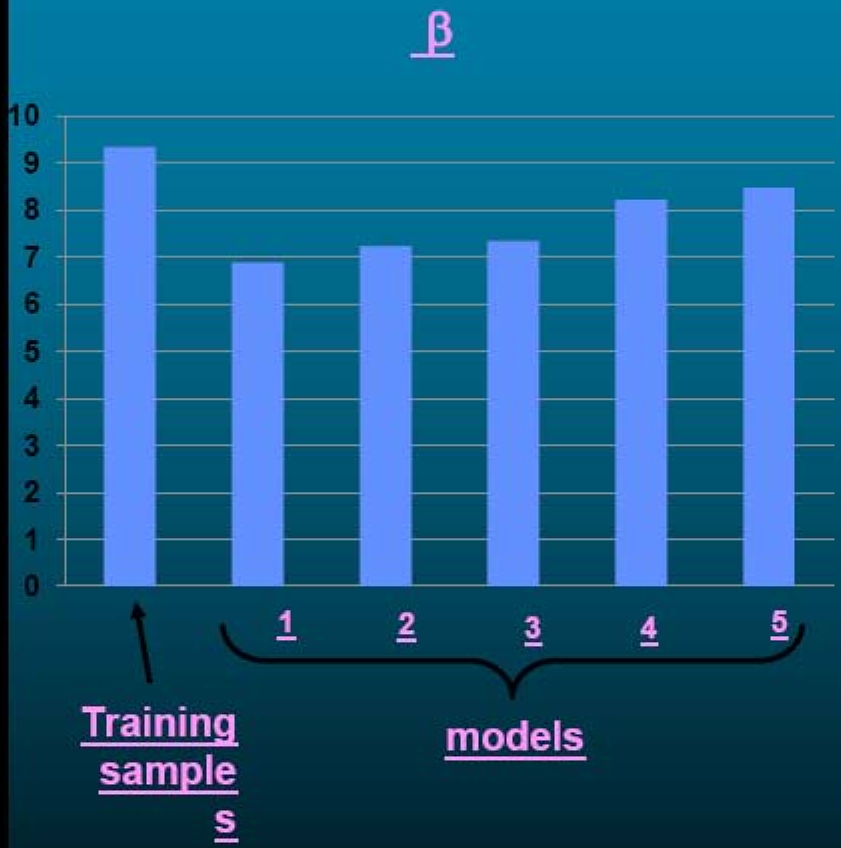
(a) (b) (c) (d)

(Zoomed) Two selected regions of classified IRS-1A image with (a and c) model 1, and (b and d) model 5

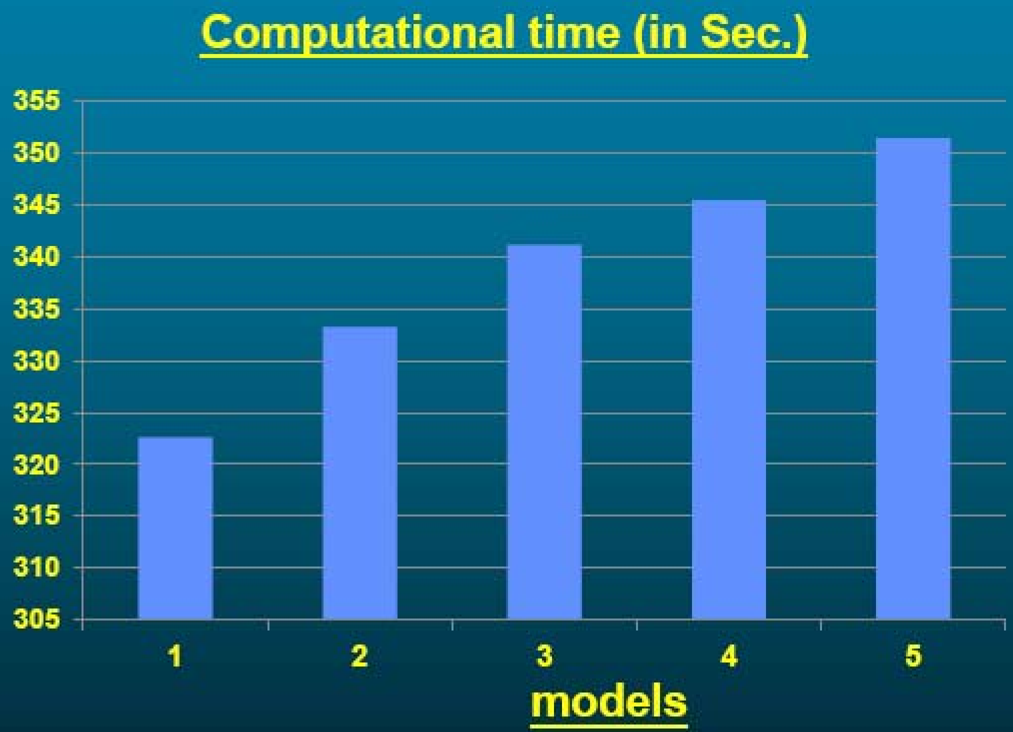
Performance comparison of rough-fuzzy models in terms β and DB indexes using k-NN classifier (k=1) with IRS-1A remote sensing image



Performance comparison of rough-fuzzy models in terms β and DB indexes using k-NN classifier (k=1) with SPOT remote sensing image (P=2, Phi=0.45)



Performance comparison of rough-fuzzy models in terms computational time (Tc) using k-NN classifier (k=1) with IRS-1A remote sensing image



Conclusions

- Rough-fuzzy granulated models for pattern classification are proposed, and encouraging performance was achieved with the synergetic integration of the both the granulation process,
- The advantage of neighborhood rough sets (NRS) that deal with both numerical and categorical data without any discretisation is also realized in the proposed models,
- Models with granulated feature space yielded improved performance compared to models with non-granulated feature space; justifying the use of granular computing based methods,
- NRS based feature selection method performed better than PaRS in both types of granulated models.
- Model F5 (among the rough-fuzzy granulated models) performed the best.
- Class-dependent fuzzy granulated model (F5) was superior to others with the cost of little higher value of *Tc*.