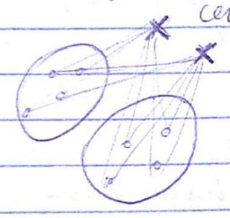
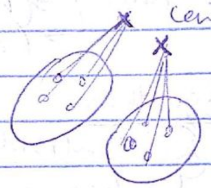
Jan 19th 2017



centroids



centroids

inter class

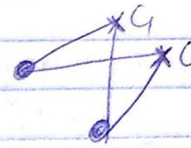
classification

distance of all x to all centroids

$x \in X$

distance of all x to corresponding centroid C_j

$x \in C_j$

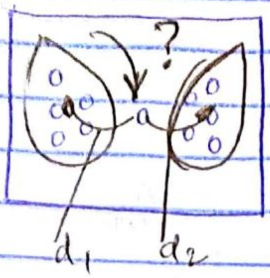


$\underset{C_1}{d(x_1)} < \underset{C_2}{d(x_1)}$

$\Rightarrow x_1 \in C_1$

$\Rightarrow x_2 \in C_2$

Butterfly Problem.



$d_1 = d_2$

if $x_1 \xrightarrow{59} \times C_1$, $x_1 \xrightarrow{60} \times C_2$, x_2

(Crisp) or (Hard clustering)

In clustering, we are making "hard" decisions based on a yes - no scheme. $x \in C_j$ or $x \notin C_j$

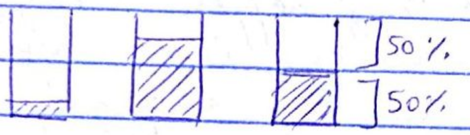
↳ set theory.

Set theory.

$A = \{1, 2, 3\}$ $A \subset X$

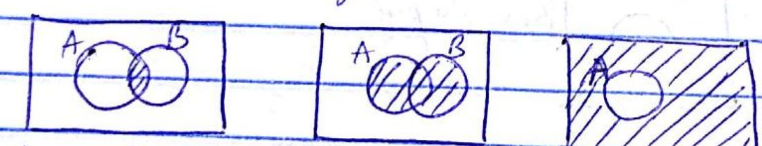
$A = \{x \mid x \text{ has property } \Delta\}$

Description via characteristic function $f_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$



true (1), false (0), dont know (0.5)

(Boolean) Set Theory



$A \cap B$ $A \cup B$ $\neg A$ or $\bar{A}$ or (not A)

Laws of

\#1 Law of non-contradiction

$A \cap \bar{A} = \phi$.

\#2 Law of excluded middle

$A \cup \bar{A} = X$.

(Characteristic func^n is either 1 or 0.)

(nothing ~~int~~ between, its all complete)

But "*Fuzzy Sets*" $X, A, \quad A \subset X, \mu_A$ membership function.

$$A = \{ (x, \mu_A(x)) \,\|\, x \in X, \mu_A(x) \in [0,1] \}$$

Example

$X = \{1, 2, \dots 7\}$

↳ set of neighbors of $4 = A_{classic} = \{3, 4, 5\}$

but Fuzzy says:

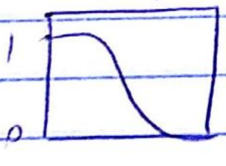
$$A_{Fuzzy} = \left\{ \frac{0.3}{1} \frac{0.7}{2} \frac{1}{3} \frac{1}{4} \frac{1}{5} \frac{0.7}{6} \frac{0.3}{7} \right\}$$

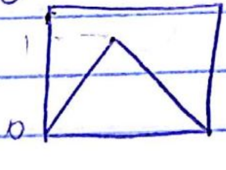
lot more memory,
degree of membership to every element.

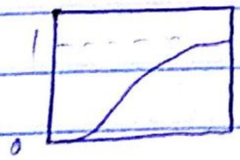
Meaning of membership

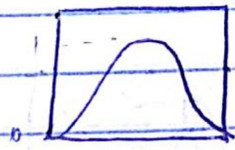
1. Similarity.
2. Probability
3. Intensity
4. Approximation
5. Compatibility.

Before things happen → probability
After things happen → Fuzzy
Eg. shade of T-shirt after a person enters class is fuzzy.
Probability that a person enters class with red shirt → probability.



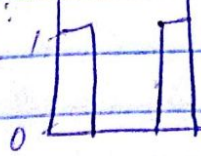


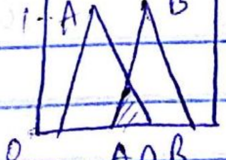




Bell shaped.

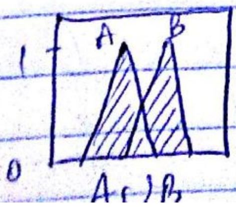
classic:



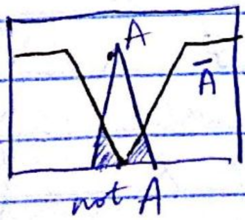


A∩B

at some point, gradually, you lose membership



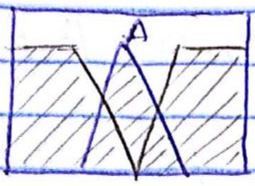
A∪B



not A

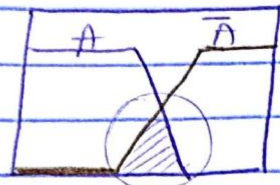
$A \cap \bar{A} \neq \phi$!!

Scanned by CamScanner

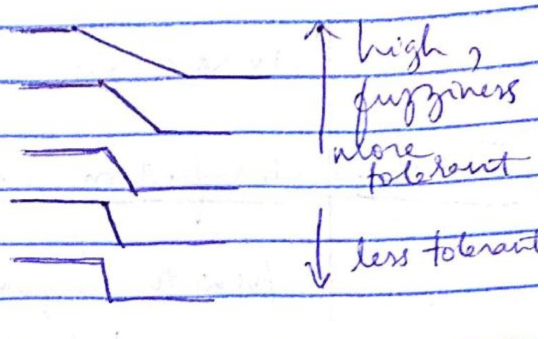


$A \cup \bar{A} \neq X$

Can we measure fuzziness?



$A \cap \bar{A} \neq \phi$



(like binary) (high fuzziness)

index of fuzziness

$$\nu = \frac{2}{N} \sum_{n} \min\left(\mu_n(x), 1 - \mu_n(x)\right)$$

J. Bezdek Fuzzy C-Means. (FCM)

Kmeans → hard clustering
FCM: soft, more tolerant

① Initialization

#clusters c, m, μ

fuzzifier membership

② Cluster

$$C_i = \frac{\sum_{k=1}^{n} (\mu_{ik})^m X_k}{\sum_{k=1}^{n} (\mu_{ik})^m}.$$

weighted sum of data → based on a membership function that is smartly designed.

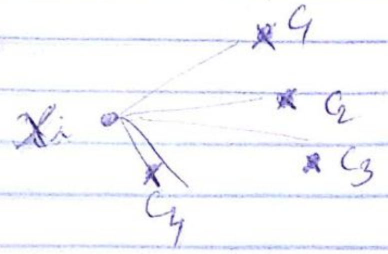
③ Memberships:

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{2/(m-1)}}$$

④ Interruption $\| U^{current} - U^{before} \| \leq \varepsilon$

Scanned by CamScanner

eg.



kmeans gives: $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$

FCM: $\begin{bmatrix} 0.1 & 0.2 & 0.2 & 0.5 \end{bmatrix}$

average