

Supervised

Classification

KNN

DT

Random Forest

AdaBoost

Bayes

Naïve Bayes
SVM

Logistic
Regression

MLP

Regression

Linear

poly

Ridge

MLP

FS → C_v
→ C_{genetic}

DR → PCA
→ LDA

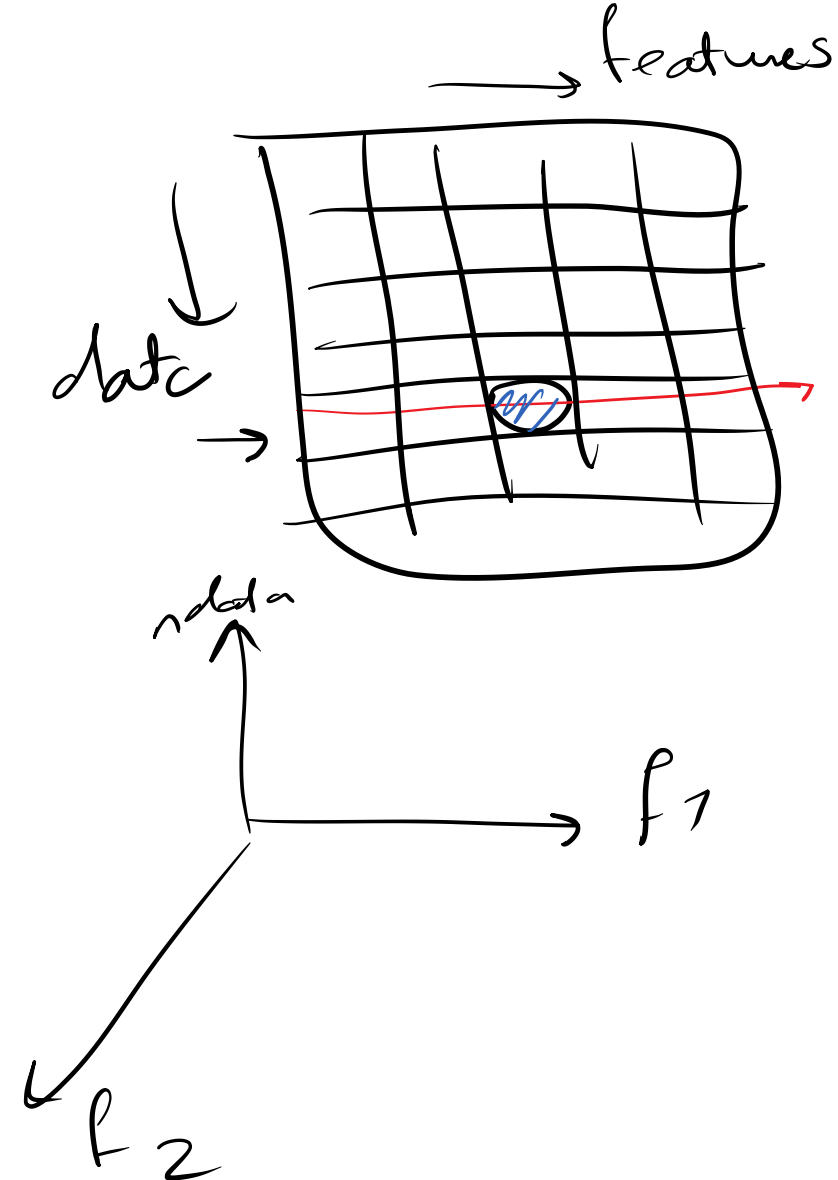
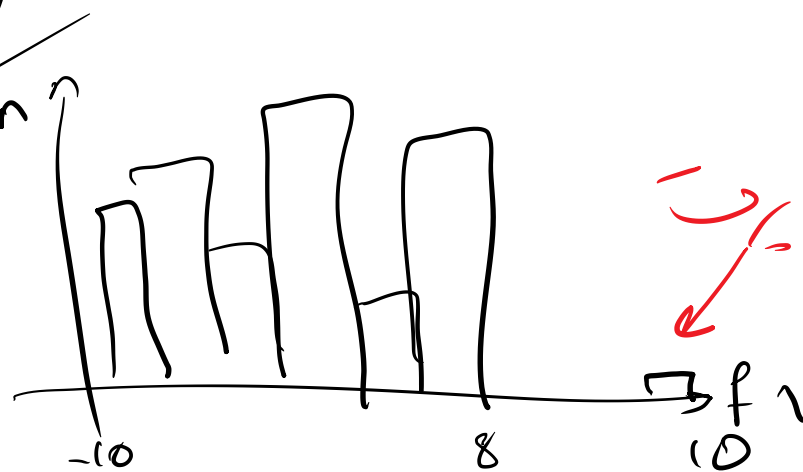
Outlier Analysis

• Concept

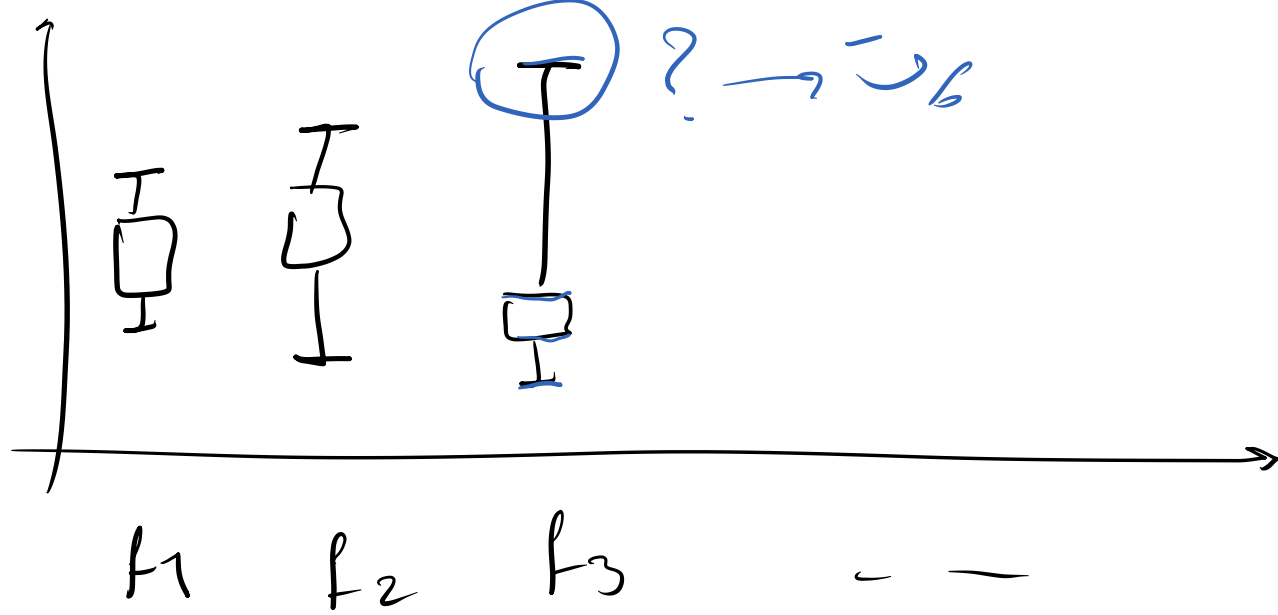
$$Z_i = \frac{x_i - \mu}{\sigma}$$

$-3 \uparrow$
 $\downarrow 3$

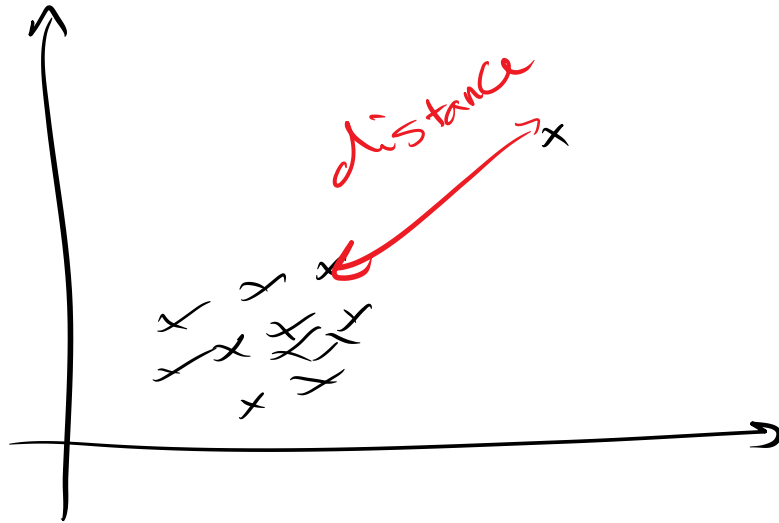
تعداد n



• Boxplot



• ppt

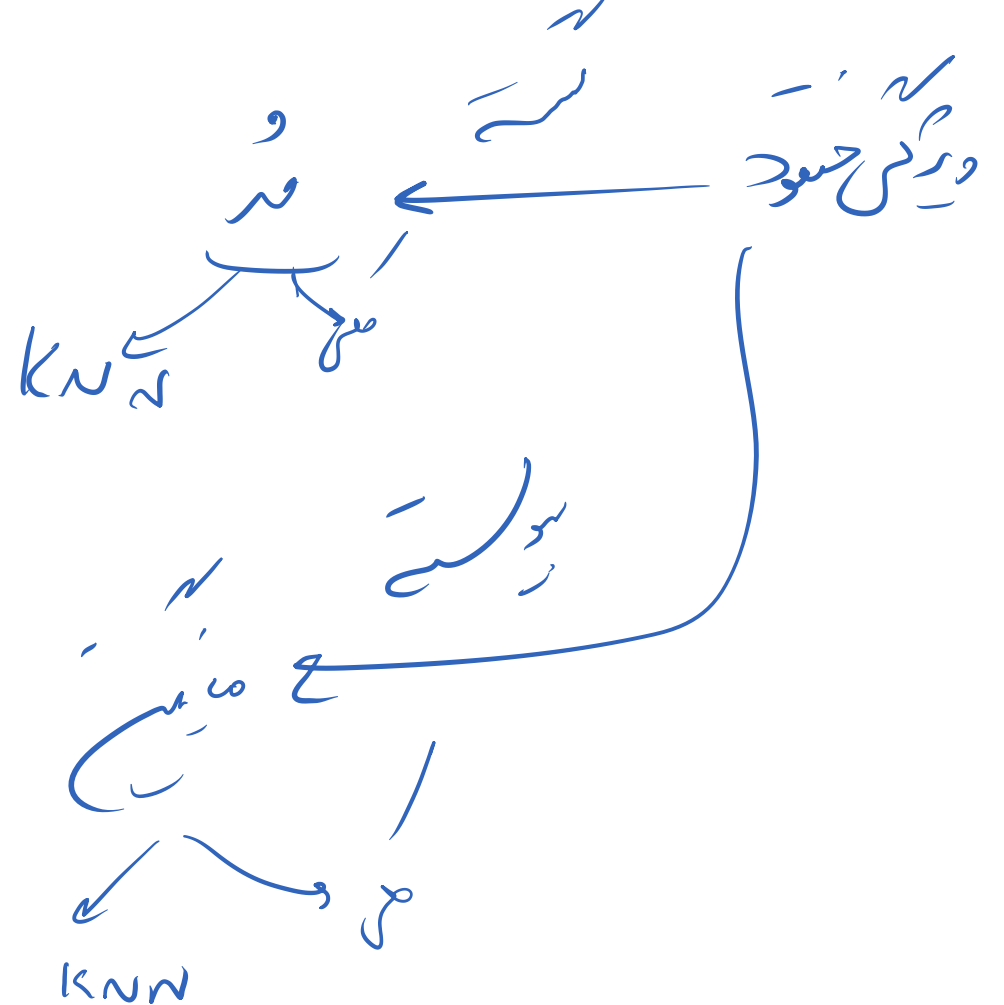


LoF

local outlier factor

Expected
Maximization

اماری



Clustering

K-means

- μ_1, \dots, μ_K Random Select (init)

while not_Converge:

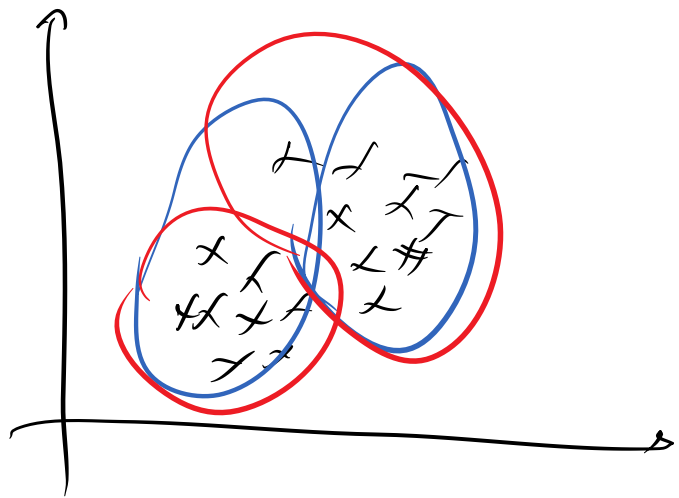
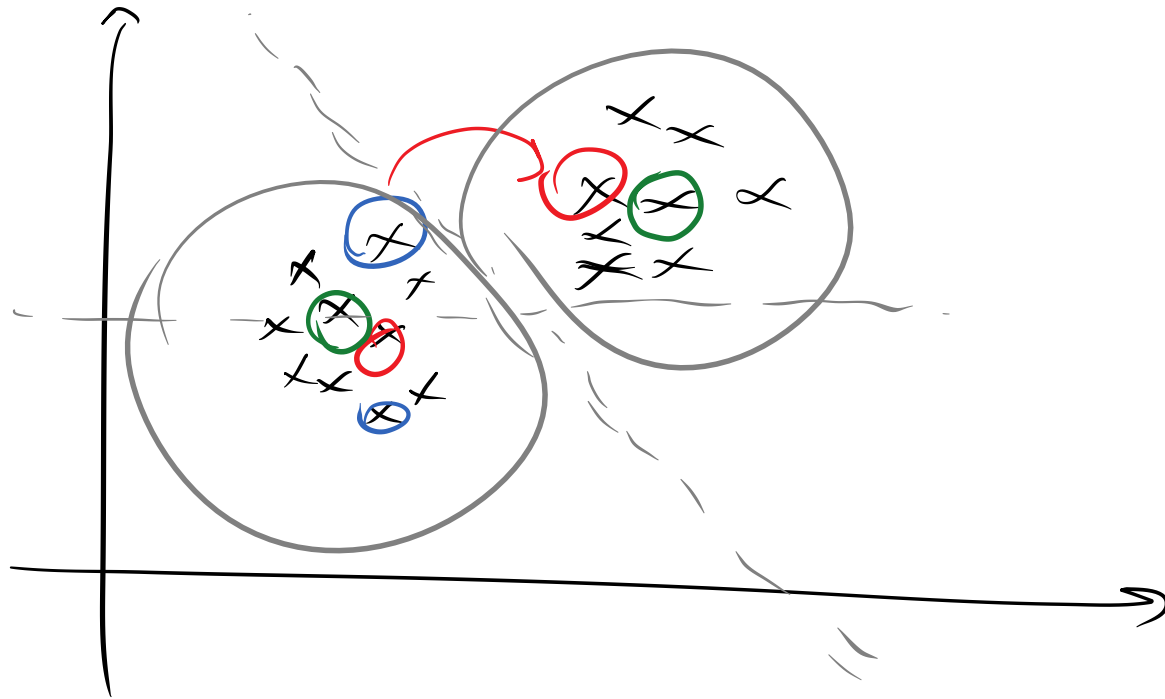
do
for k
 $1 \leq i \leq n$
 $1 \leq j \leq K$

(Assignment) : $c^{(i)} : \arg \min_j \|x_i - \mu_j\|^2$

(update Centroid)

$$\mu_j = \frac{\sum_{i=1}^n \mathbf{1}\{c^{(i)} = j\} x_i}{\sum_{i=1}^n \mathbf{1}\{c^{(i)} = j\}}$$

1 2 3
○ ○ ○



عدم دیراری

مساحت داده دیر

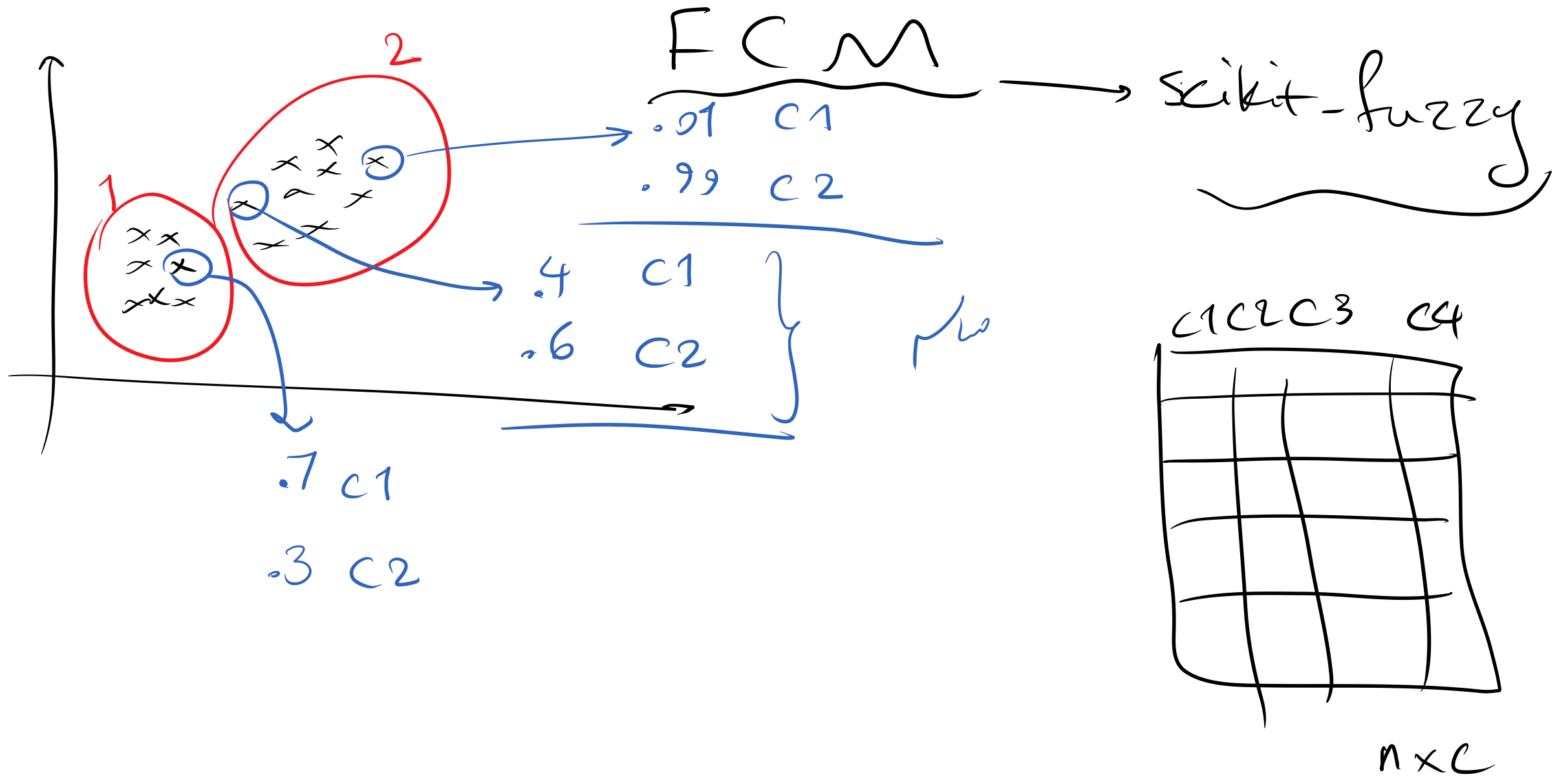
• دیر کار دیر

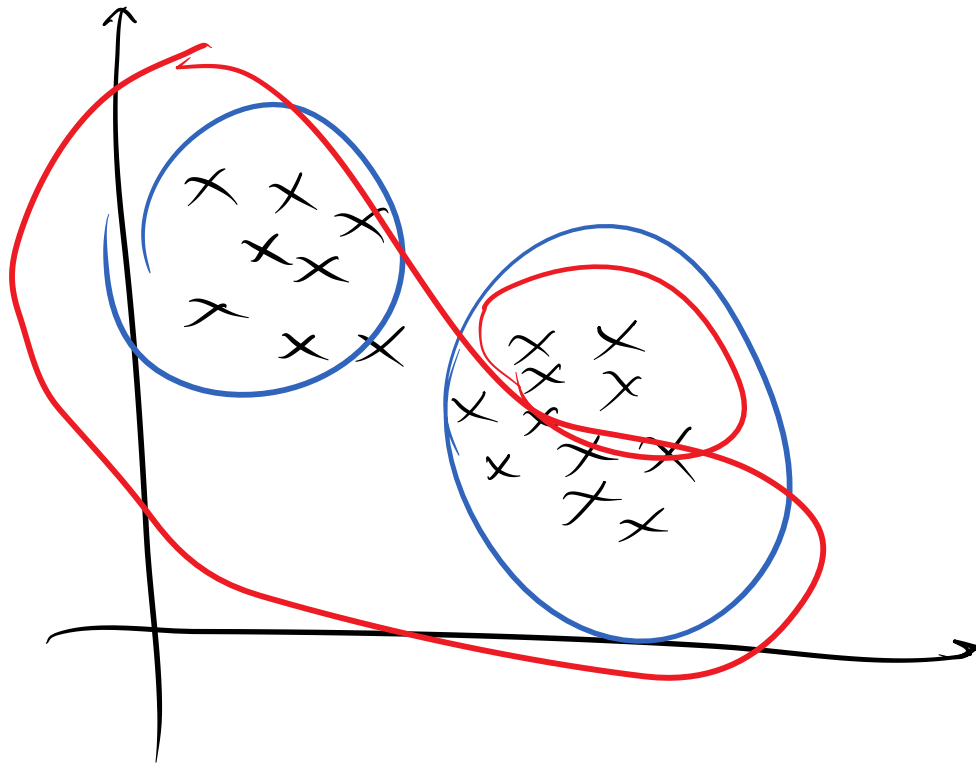
• داده

• سریع

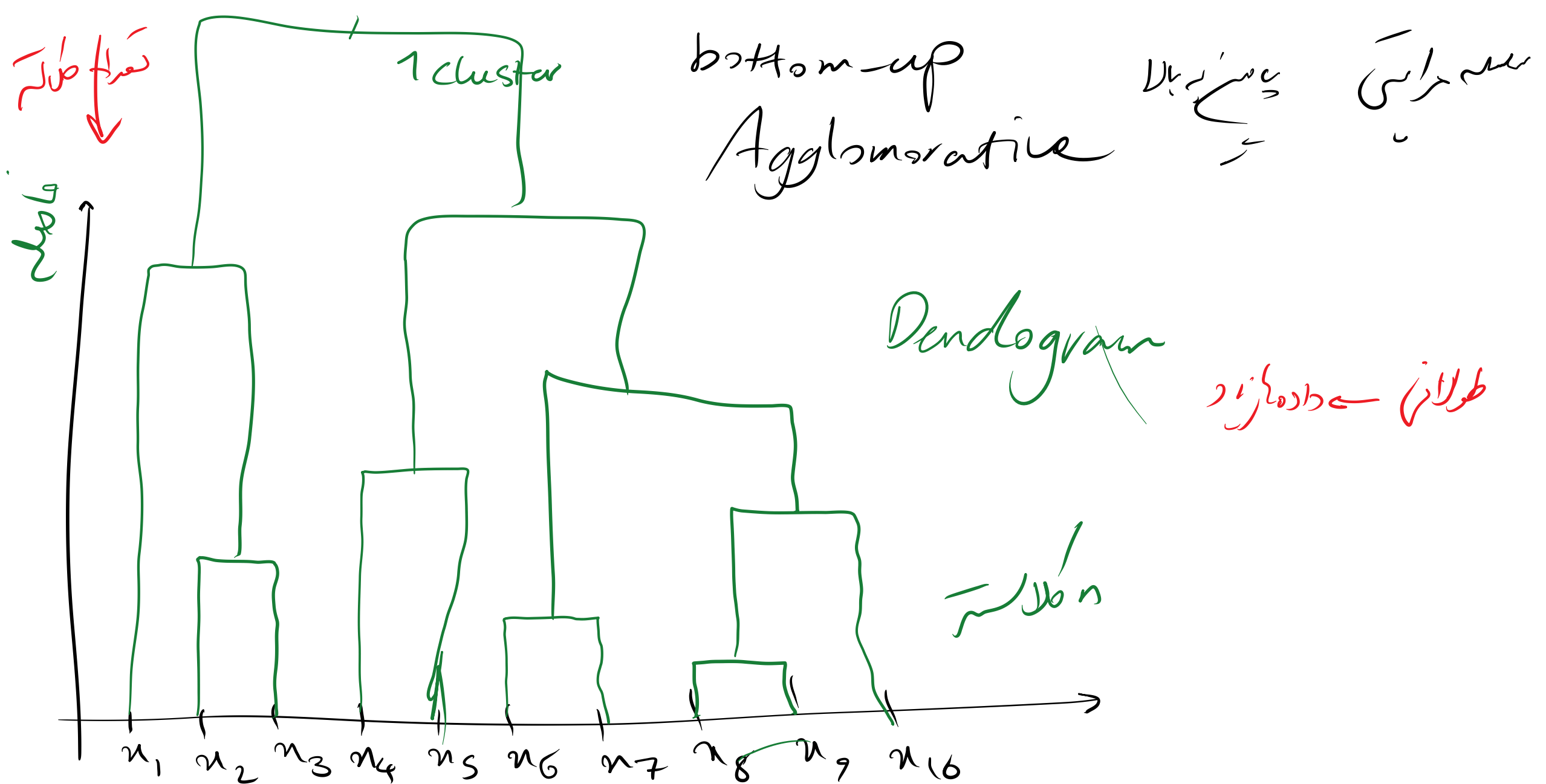
• حرکت دیر

• حرکت دیر





معادله
خطه درون کلاسی ↓
خطه برون کلاسی ↑



Top-down

ساختار از بالا به پایین

division

تقسیم و حاکم

