Machine Learning in Biomedical Engineering

Ahmet Sacan

Wisconsin Breast Cancer Dataset

- http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original)
- Samples arrive periodically as Dr. Wolberg reports his clinical cases.
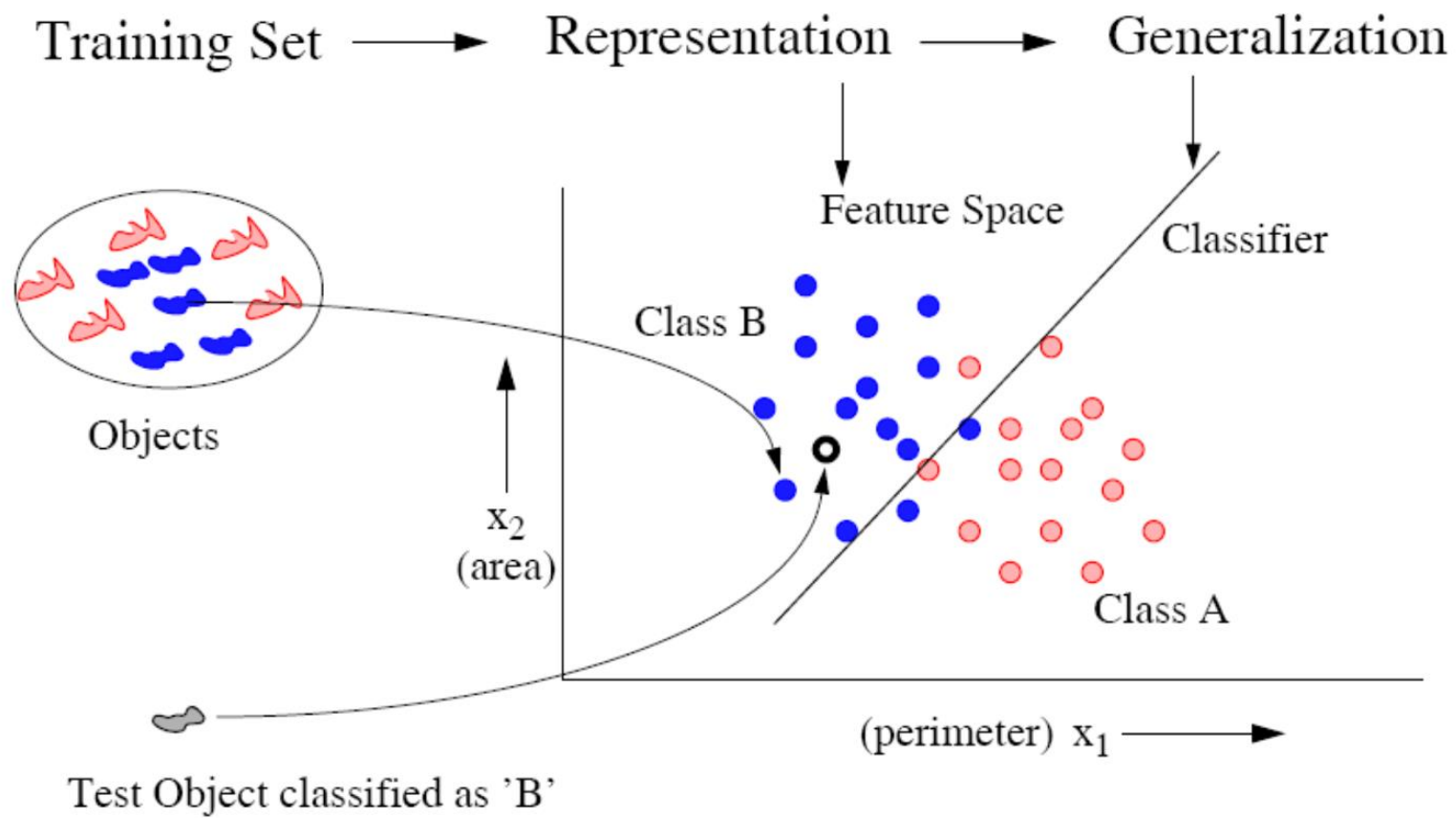
- **Attribute Information:**
  1. Sample code number: id number
  2. Clump Thickness: 1 - 10
  3. Uniformity of Cell Size: 1 - 10
  4. Uniformity of Cell Shape: 1 – 10
  5. Marginal Adhesion: 1 - 10
  6. Single Epithelial Cell Size: 1 - 10
  7. Bare Nuclei: 1 - 10
  8. Bland Chromatin: 1 - 10
  9. Normal Nucleoli: 1 - 10
  10. Mitoses: 1 - 10
  11. Class: (2 for benign, 4 for malignant)

Data Set Characteristics	Multivariate
Number of Instances	699
Attribute Characteristics	Integer
Number of Attributes	10
Missing Values?	Yes

Wisconsin Breast Cancer Dataset

1000025,5,1,1,1,2,1,3,1,1,2
1002945,5,4,4,5,7,10,3,2,1,2
1015425,3,1,1,1,2,2,3,1,1,2
1016277,6,8,8,1,3,4,3,7,1,2
1017023,4,1,1,3,2,1,3,1,1,2
1017122,8,10,10,8,7,10,9,7,1,4
1018099,1,1,1,1,2,10,3,1,1,2
1018561,2,1,2,1,2,1,3,1,1,2
1033078,2,1,1,1,2,1,1,1,5,2
1033078,4,2,1,1,2,1,2,1,1,2
1035283,1,1,1,1,1,1,3,1,1,2
1036172,2,1,1,1,2,1,2,1,1,2
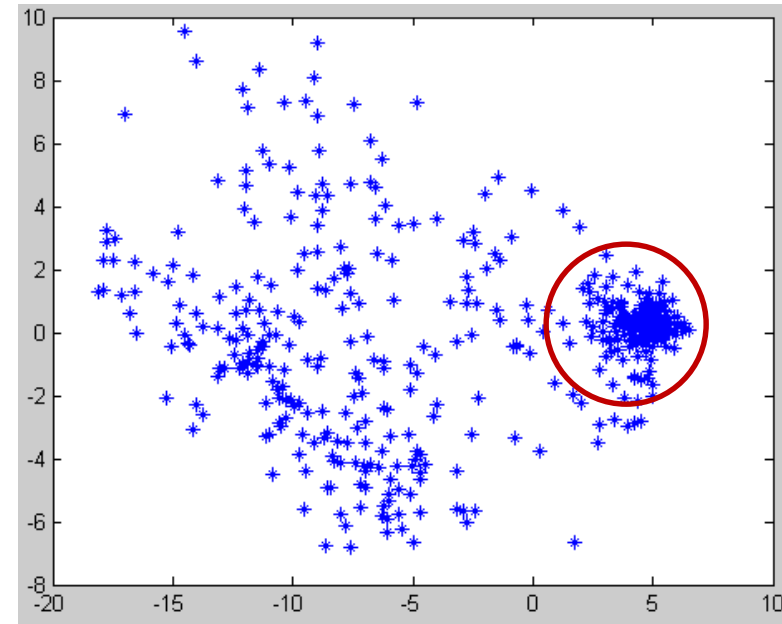1041801,5,3,3,3,2,3,4,4,1,4
1043999,1,1,1,1,2,3,3,1,1,2
...

id	clump thick	unifority of	uniformity	marginal a	single epith	bare nuclei	bare chrom	normal nuc	mitoses	class
1000025	5	1	1	1	2	1	3	1	1	2
1002945	5	4	4	5	7	10	3	2	1	2
1015425	3	1	1	1	2	2	3	1	1	2
1016277	6	8	8	1	3	4	3	7	1	2
1017023	4	1	1	3	2	1	3	1	1	2
1017122	8	10	10	8	7	10	9	7	1	4
1018099	1	1	1	1	2	10	3	1	1	2
1018561	2	1	2	1	2	1	3	1	1	2
1033078	2	1	1	1	2	1	1	1	5	2
1033078	4	2	1	1	2	1	2	1	1	2
1035283	1	1	1	1	1	1	3	1	1	2
1036172	2	1	1	1	2	1	2	1	1	2
1041801	5	3	3	3	2	3	4	4	1	4
1043999	1	1	1	1	2	3	3	1	1	2

Training Set $\longrightarrow$ Representation $\longrightarrow$ Generalization

Feature Space

Classifier

Class B

Objects

x_2
(area)

Class A

Test Object classified as 'B'

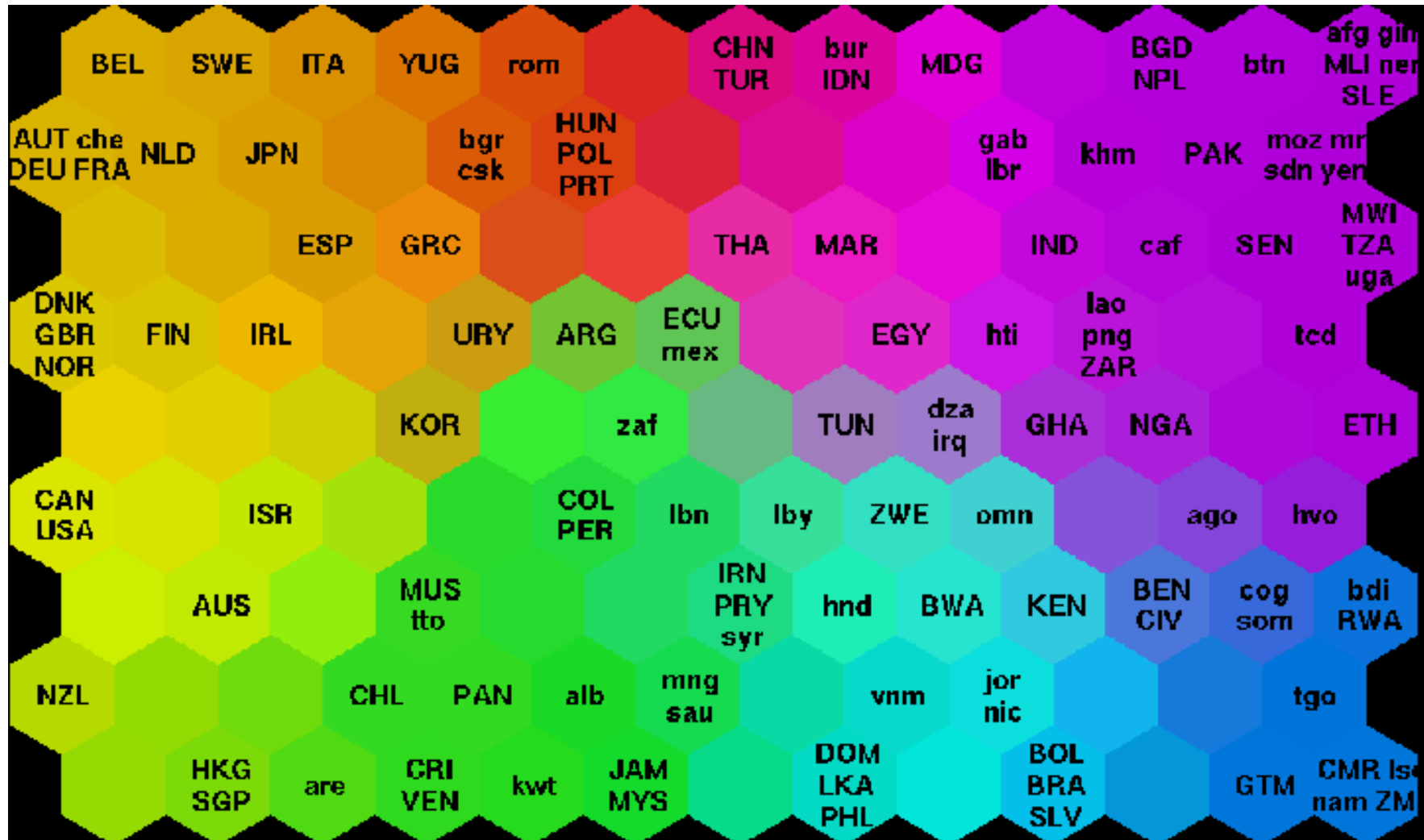(perimeter) x_1 $\longrightarrow$

Duin et al.

Projection

1000025,5,1,1,1,2,1,3,1,1,2
1002945,5,4,4,5,7,10,3,2,1,2
1015425,3,1,1,1,2,2,3,1,1,2
1016277,6,8,8,1,3,4,3,7,1,2
1017023,4,1,1,3,2,1,3,1,1,2
1017122,8,10,10,8,7,10,9,7,1,4
1018099,1,1,1,1,2,10,3,1,1,2
1018561,2,1,2,1,2,1,3,1,1,2
1033078,2,1,1,1,2,1,1,1,5,2
1033078,4,2,1,1,2,1,2,1,1,2
1035283,1,1,1,1,1,1,3,1,1,2
1036172,2,1,1,1,2,1,2,1,1,2
1041801,5,3,3,3,2,3,4,4,1,4
1043999,1,1,1,1,2,3,3,1,1,2
…

World Poverty Map

World Poverty Map