

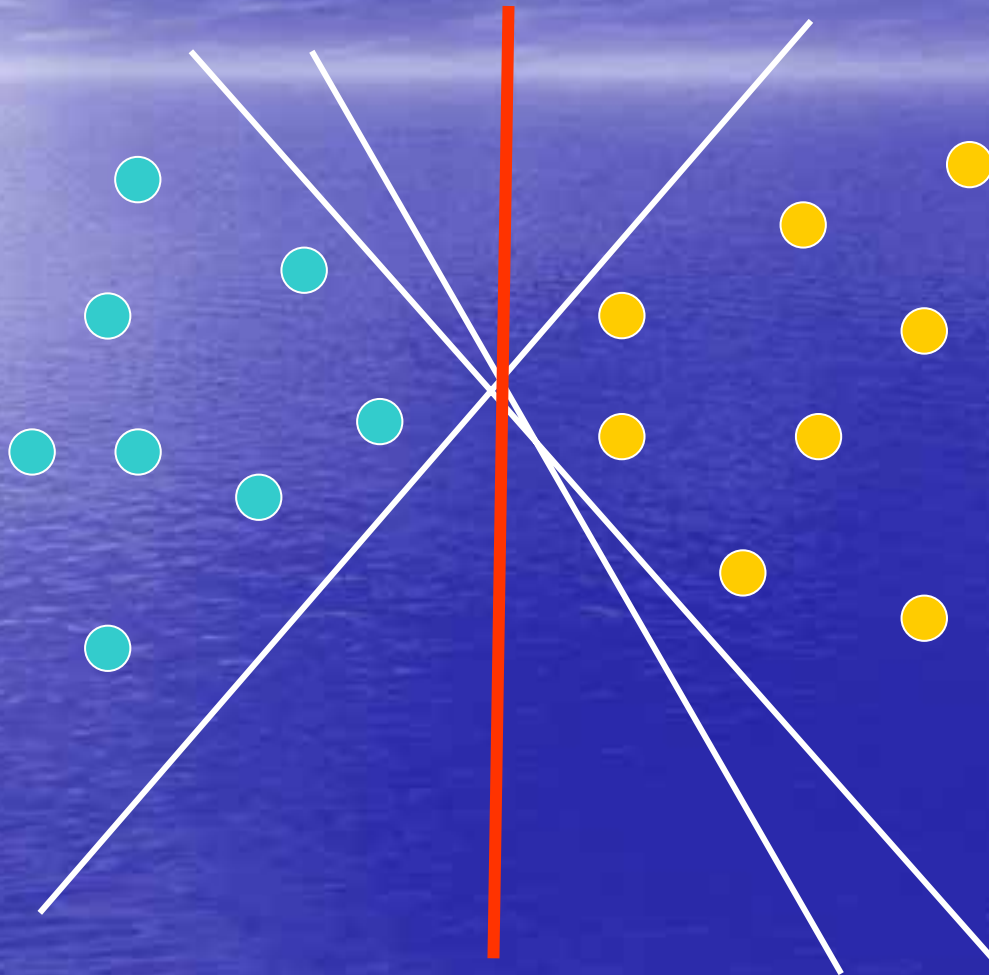
Évaluation d'un algorithme de classification automatique des galaxies fondé sur les SVMs

Réda DEHAK
EFIGI : 11/10/2005

Plan

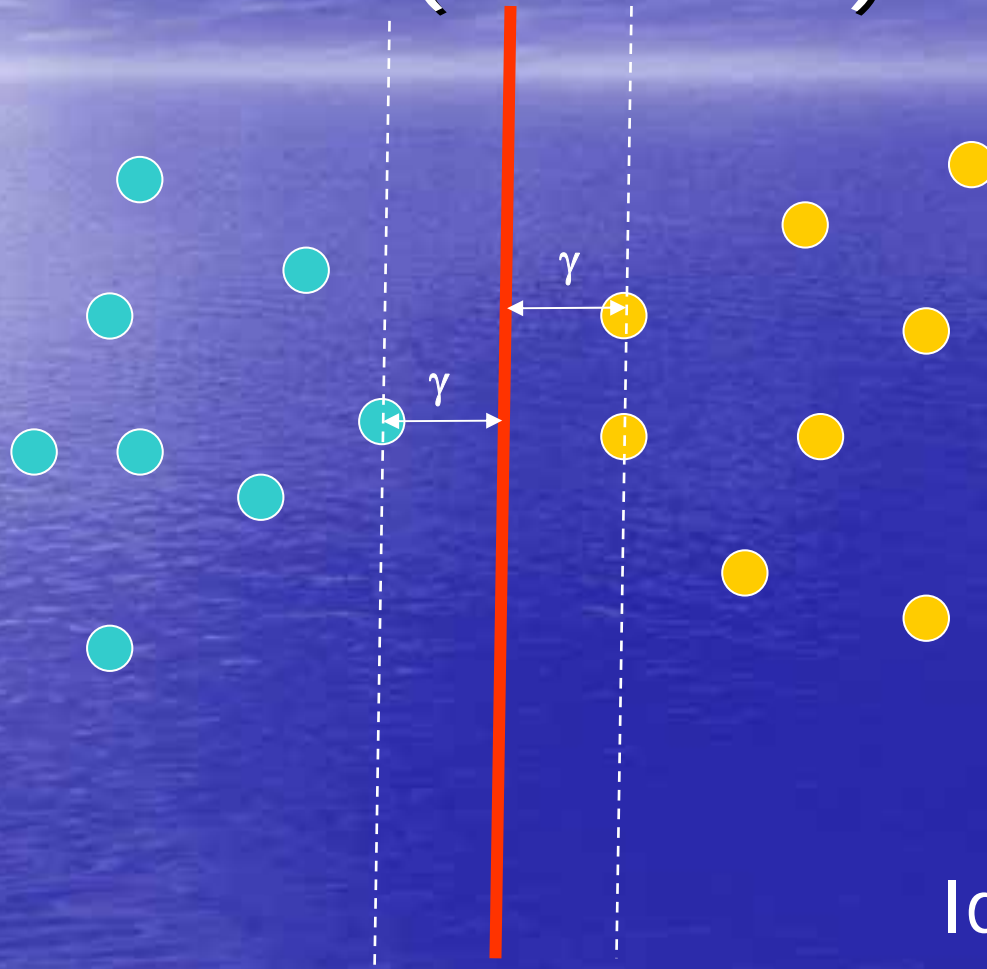
- MLP vs SVM.
- Classification Précoces / Avancées
- Type de Hubble
- Conclusions & Perspectives

Comparaison MLP - SVM



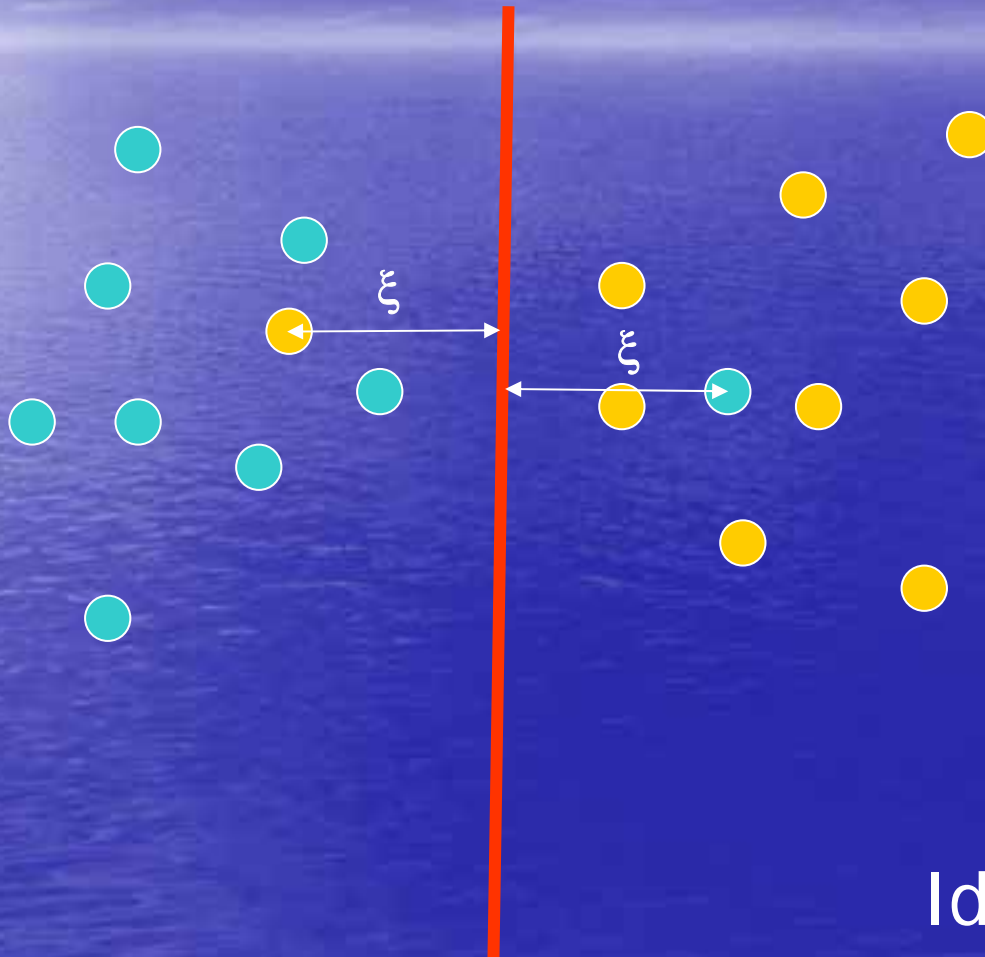
$$\langle w, x \rangle + b = 0$$

SVM : Séparateur à vaste marge (cas idéal)



Idée : Maximiser γ

SVM : Séparateur à vaste marge (cas général)



Idée : Minimiser ξ

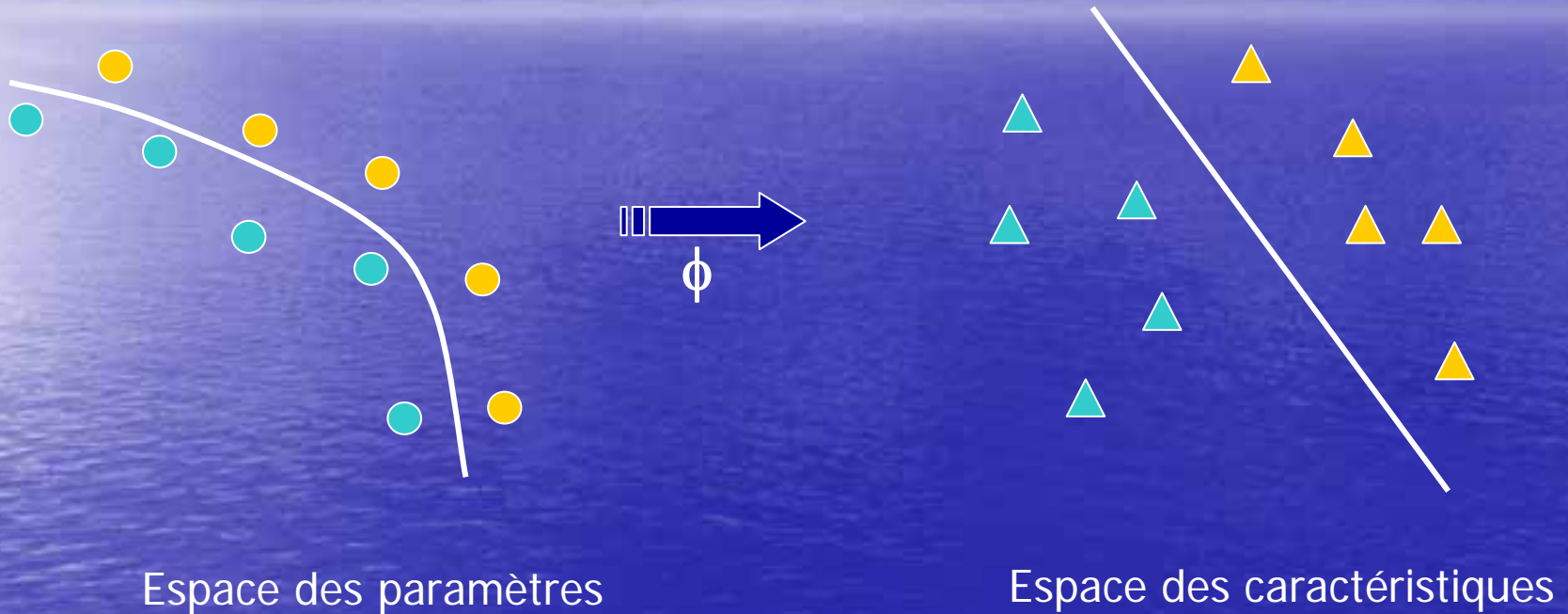
Classification [Cortes & Vapnik]

- Résoudre le système :

$$\frac{1}{2} \|w\|^2 + C \sum_i \zeta_i \rightarrow \text{Min}$$

$$\begin{cases} y_i [\langle w, x_i \rangle + b] \geq 1 - \zeta_i & i = 1, \dots, l \\ \zeta_i \geq 0 \end{cases}$$

Fonctions noyaux



$$K(x, y) = \langle \phi(x), \phi(y) \rangle$$

Applications

- Même protocole de test et de validation que Anthony :
 - L'ensemble des données est décomposé en 5 sous ensembles.
- Noyau : RBF.

Phase 1 : Classification précoces/avancées

- Utilisation des deux premières composantes de l'ACP.

Ensemble de test	Apprentissage	Test
1	79,605%	81,159%
2	80,532%	80,689%
3	81,667%	82,877%
4	80,268%	85,811%
5	81,109%	79,885%

Moyenne : 82,084%

Analyse des résultats

Test 5 : 81,109%, 79,885%

0	1	1	1	2	3	2	5	11	14	17	8	21	7	8	2	5
0	6	2	8	7	2	9	8	3	14	1	0	0	0	0	1	0
-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10

Analyse des résultats

Test 1 : 79,605% , 81,159%

0	1	0	1	3	0	0	7	5	10	10	5	21	0	6	8	7
0	11	3	5	13	1	4	3	4	3	4	2	1	0	0	0	0
-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10

Analyse des résultats

Test 4 : 80,268%, 85,811%

0	0	0	0	3	2	2	4	4	15	8	12	19	10	7	8	11
0	9	0	5	7	1	5	6	0	4	4	2	0	0	0	0	0
-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10

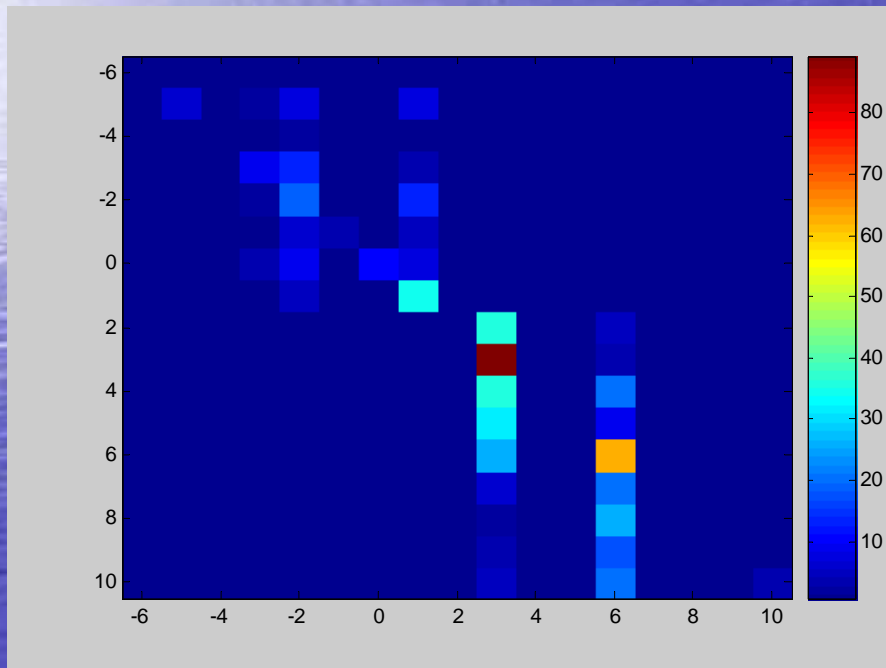
Classification de Hubble

Multi-classes

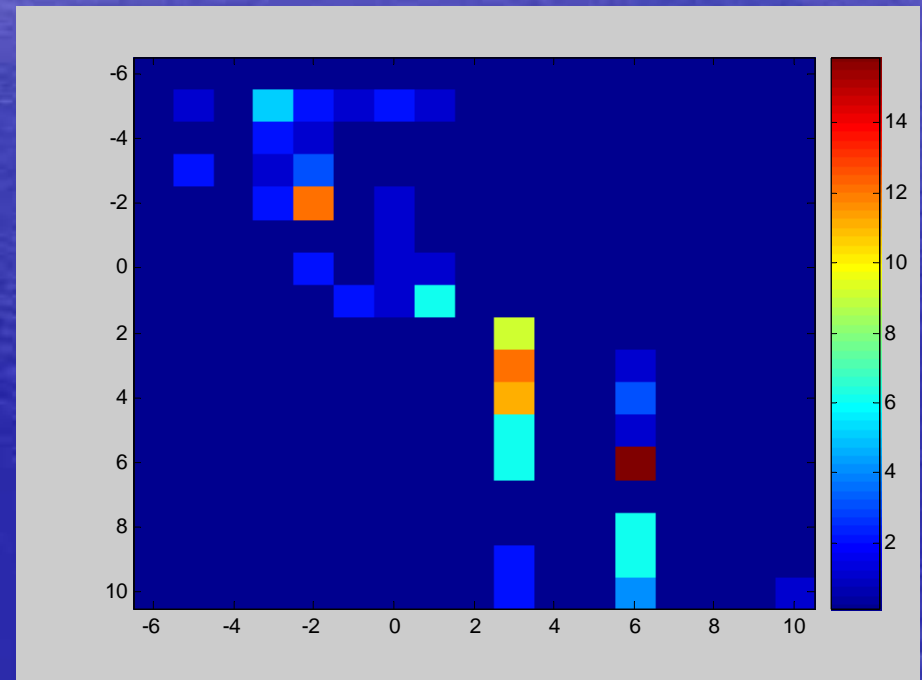
Ensemble de test	Apprentissage	Test	Test ± 1	Test ± 2
1	39,434%	36,764%	59,559%	80,147%
2	49,915%	32,639%	59,722%	81,944%
3	40,909%	45,454%	66,434%	81,818%
4	43,898%	39,496%	57,823%	74,830%
5	61,97%	35,928%	66,467%	87,425%

Analyse des résultats

Test 1 : 39,434%; 36,764%; 59,559%; 80,147%



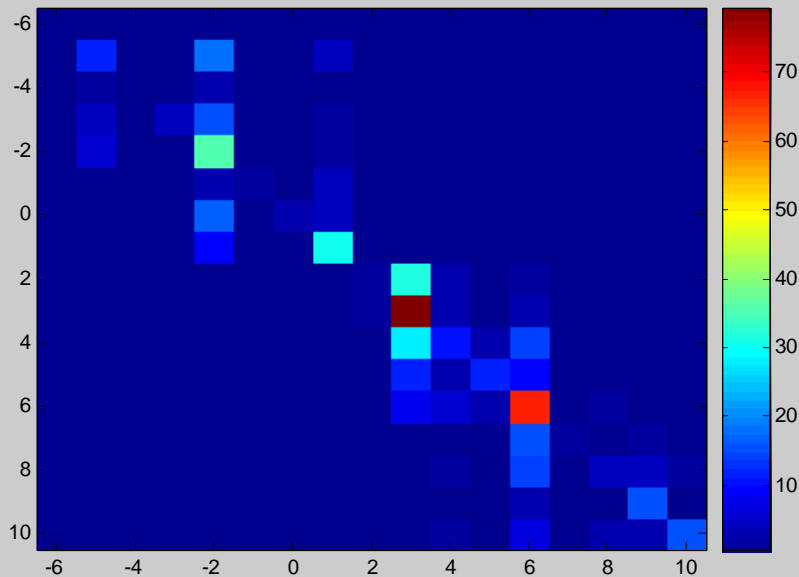
Base d'apprentissage



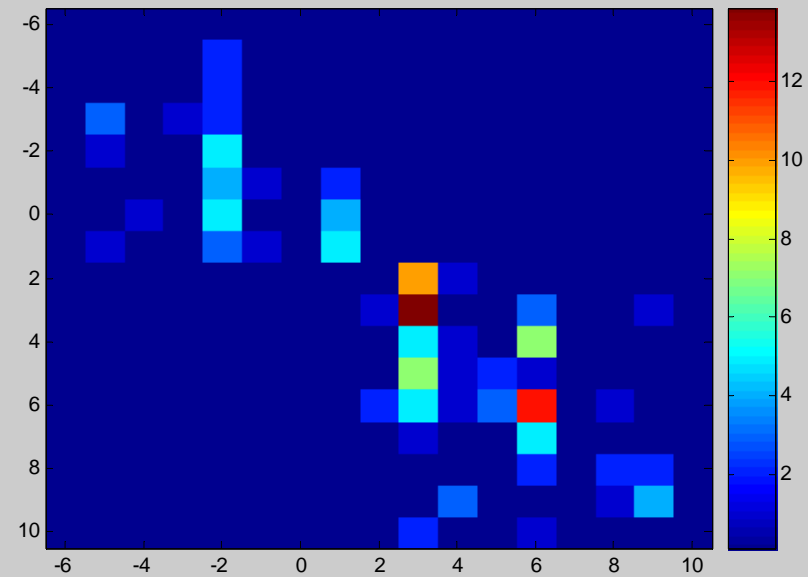
Validation

Analyse des résultats

Test 2 : 49,915%; 32,639%; 59,722%; 81,944%



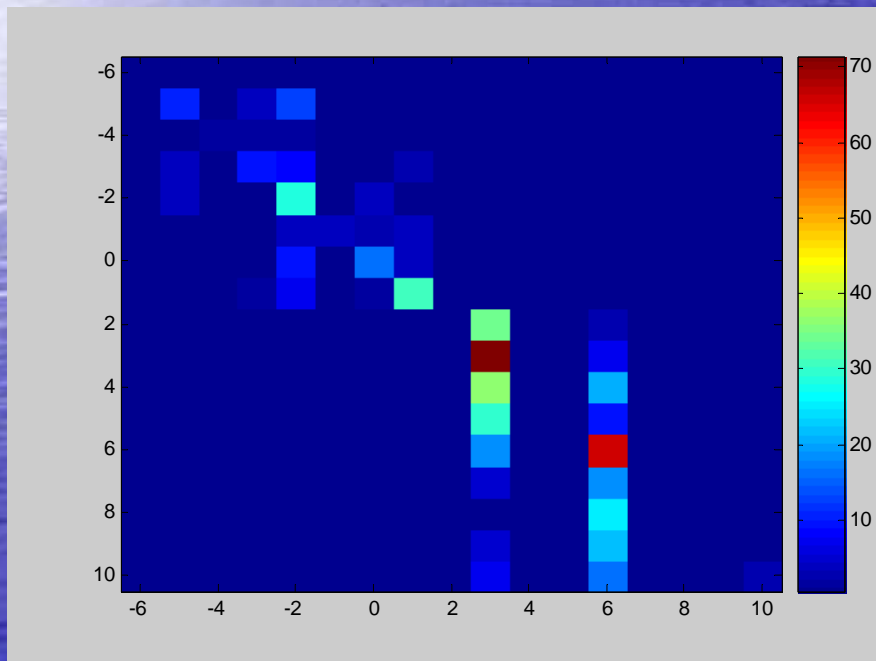
Base d'apprentissage



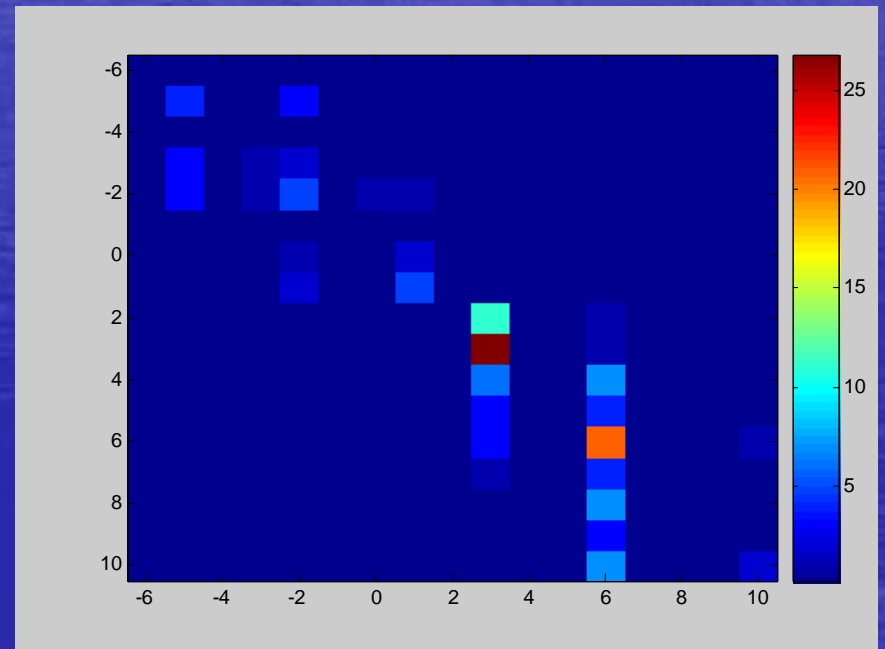
Validation

Analyse des résultats

Test 3 : 40,909%; 45,454%; 66,434%; 81,818%



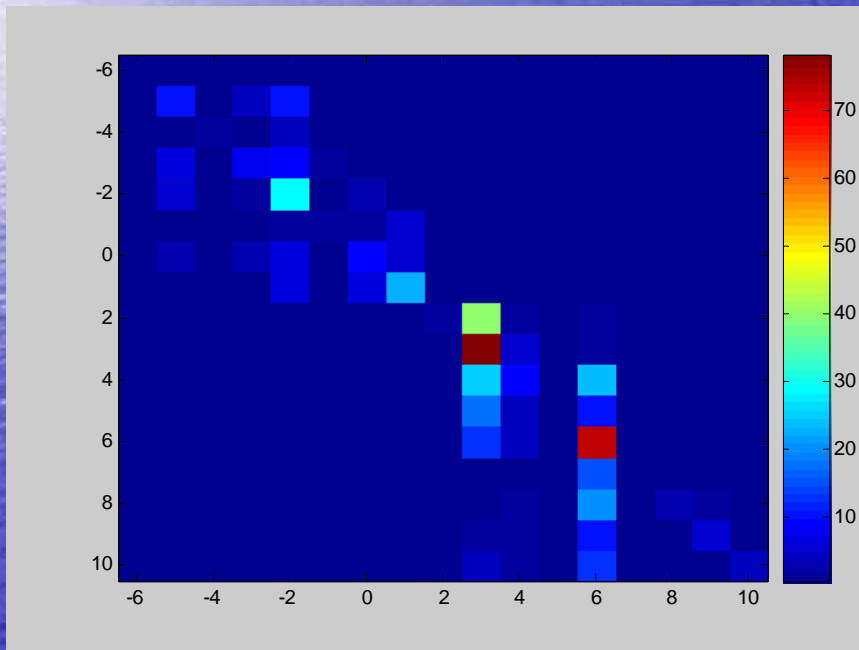
Base d'apprentissage



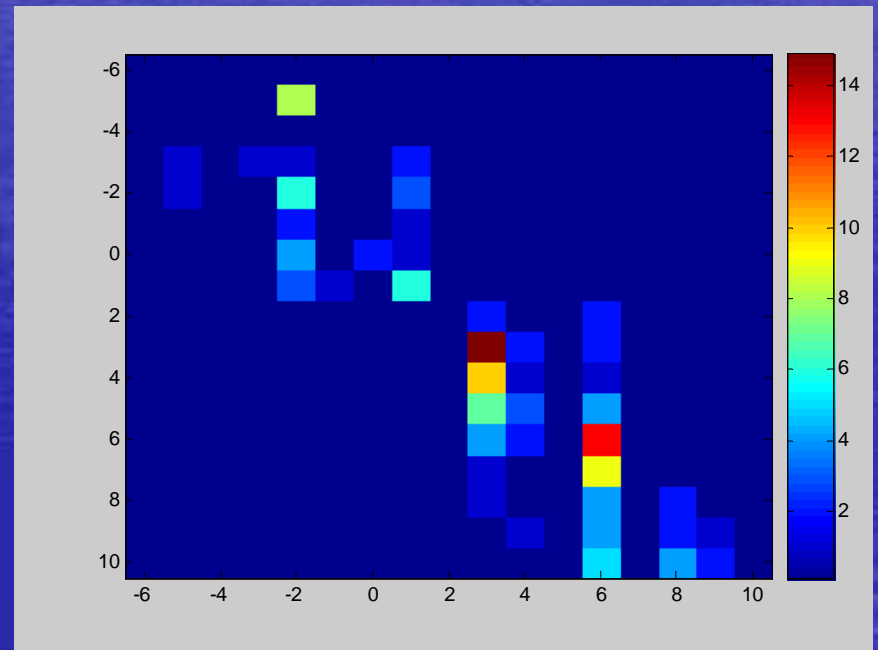
Validation

Analyse des résultats

Test 4 : 43,898%; 39,496%; 57,823%; 74,830%



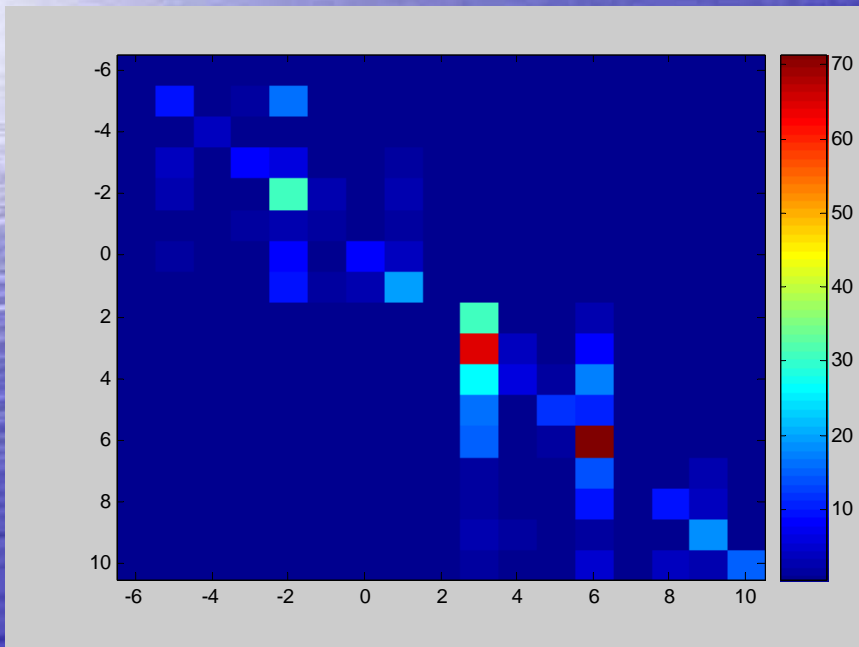
Base d'apprentissage



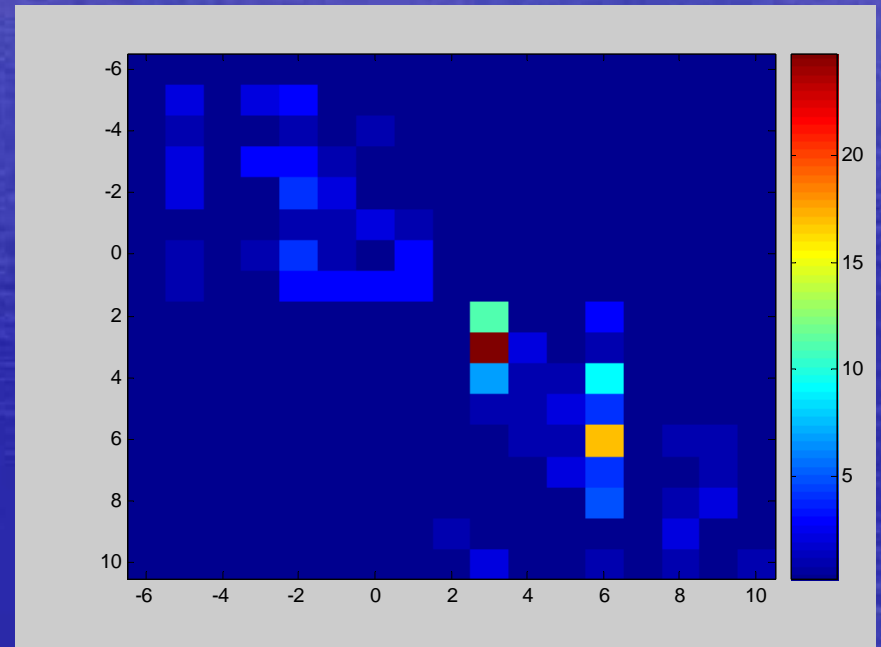
Validation

Analyse des résultats

Test 5 : 61,97%; 35,928%; 66,467%; 87,425%



Base d'apprentissage



Validation

Conclusions & Perspective

- Résultats satisfaisants pour la classification des galaxies précoces/avancés :
 - Valider sur une base plus importante.
- Résultats insuffisants pour la classification de Hubble :
 - Base d'apprentissage trop petite.
 - Classes mal représentées.
 - Espace des paramètres mal adapté.