

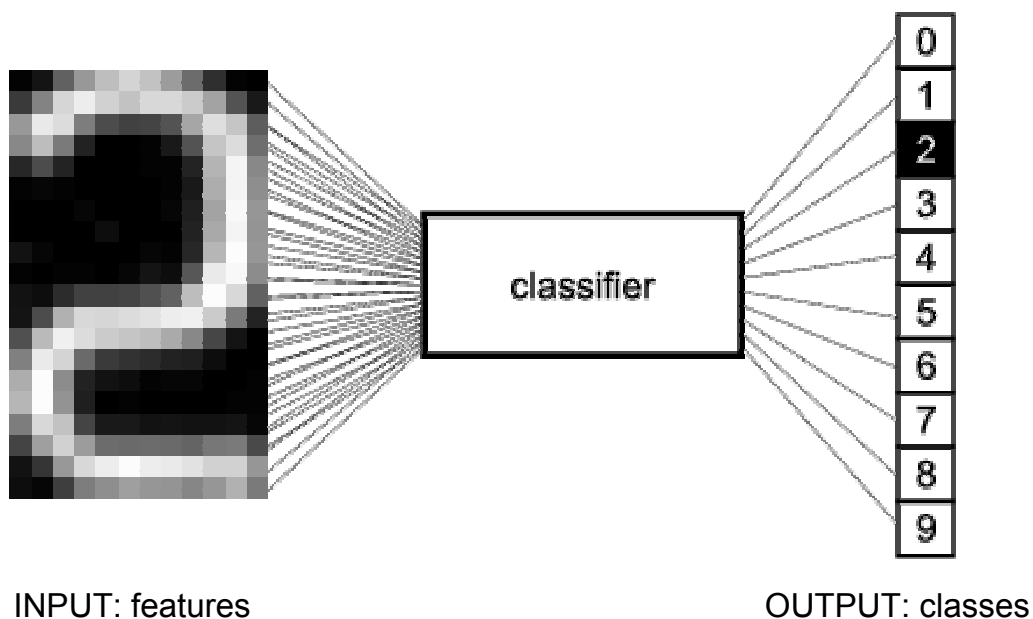
4.3 Classificadors

- 1 Introducció: Què és un classificador
- 2 Feature selection: dos exemples típics (i feature extraction?)
- 3 Funcions discriminants i fronteres de decisió
- 4 Classificadors estadístics, classificadors basats en distàncies, template matching
- 5 Arbres de decisió
- 6 Altres classificadors: sintàctics, CBR, X.Neuronals
- 7 Aprenentatge no-supervisat: clustering

1

Què és un classificador?

SITUEM-NOS EN EL CONTEXTE: segmentar, caracteritzar, **classificar**



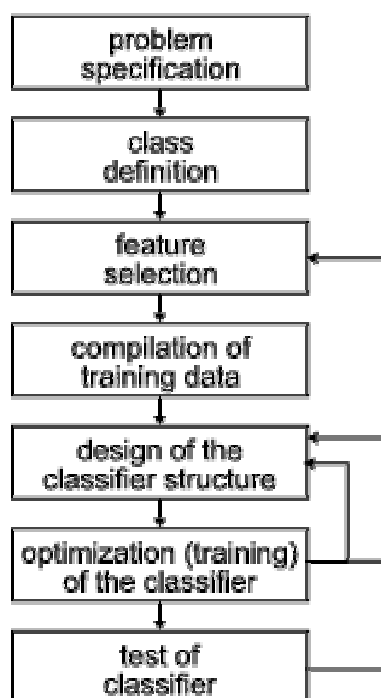
2

Què és un classificador? Què ha de fer?

1. *Classification is the central operation for the recognition of patterns, attempt to give a designation to observed phenomena in order to move from the purely descriptive level of physical, geometrical or other features to a symbolic level*
2. *Classification is the assignment of an object characterized by a set of features to one of a number of predefined classes*
3. *But in image processing, the typical task of classification is to imitate the an assignment prescribed by a human inspector*

3

Com es construeix un classificador?



1. Com sempre, primer cal tenir clar quin és el problema.
2. Les classes a tenir en consideració: la sortida del classificador (no sempre es coneixen)
3. Seleccionar les característiques més apropiades sol ser sempre més complexe que trivial.
4. Cal que les dades utilitzades per el disseny del classificador (entrenament) siguin prou representatives
5. and 6. Construir el classificador i testear-lo amb el training set
7. Després d'un entrenament amb èxit arriba l'hora de testar el classificador amb noves dades

4

Aprendre amb i sense training-data:

Com descriure les diferències entre una **E** i una **F** de forma que un computador pugui entendre-ho?

1. Si no hi ha training-data, s'aprèn explicant que la **E** té una pota més que la **F** no n'hi ha prou. Cal descriure com és una i com és l'altre: ...després al arribar a la part inferior esquerra la lletra **E** pateix un canvi brusc de direcció (de 90°) cap a la dreta....
2. Si hi ha training-data, l'aprenentatge es basa en presentar una col·lecció de mostres de les quals es coneix la “correcta” correspondència entre la seva descripció (features) i la seva classe. Es diu que el classificador aprèn imitant.

Feature selection:

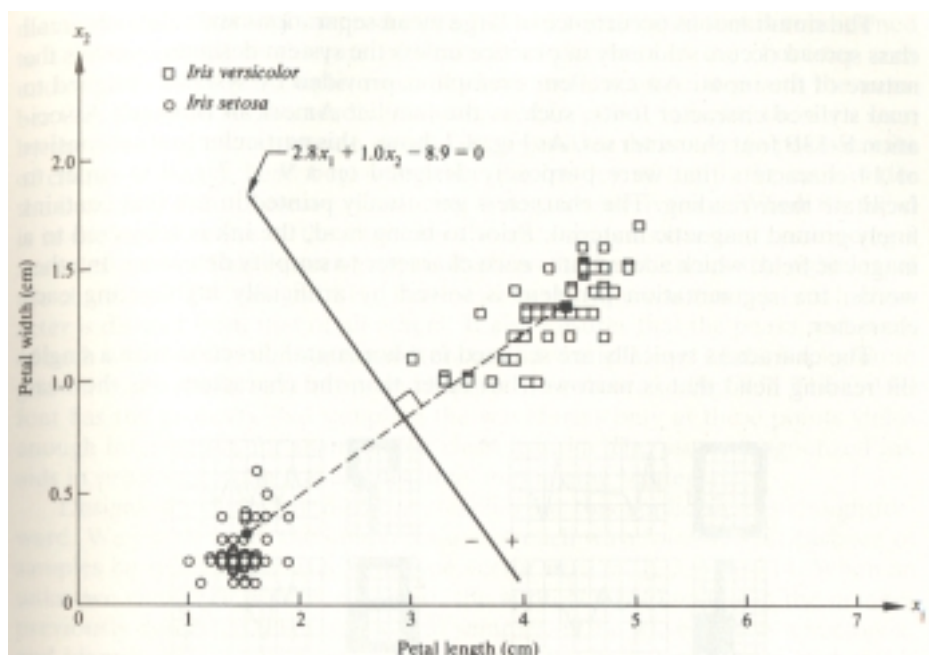
Funcions discriminants i fronteres de decisió:

Donat un $x=(x_1,x_2,\dots,x_n)^T$ que representa un pattern vector. Per M classes $w_1, w_2, \dots, w_M$, el problema bàsic en decision-theoretic pattern recognition és trobar M funcions de decisió (o discriminants) $d_1(x), d_2(x), \dots, d_M(x)$, amb la propietat que si el pattern x pertany a la classe w_i , llavors

$$d_i(x) > d_j(x) \quad j=1,2,\dots,M; j \neq i$$

La frontera de decisió separa la classe w_i de w_j està formada per tots els valors de x tals que $d_i(x) = d_j(x)$

Funcions discriminants i fronteres de decisió:



Classificació estadística:

Bayes classifier:

$$d_i(\mathbf{x}) = \mathbf{x}^T \mathbf{C}^{-1} \mathbf{m}_i - 1/2(\mathbf{m}_i^T \mathbf{C}^{-1} \mathbf{m}_i)$$

....has theoretically minimal rate of error. Since it requires full knowledge of the statistical properties of the problem, it is usually impossible to construct in practice

....thus, the minimum distance classifier is optimum in the bayes sense if,

- (1) the pattern classes are gaussian
- (2) all covariance matrix are equal to the identity matrix,
- (3) ...

9

Classificadors basats en distàncies:

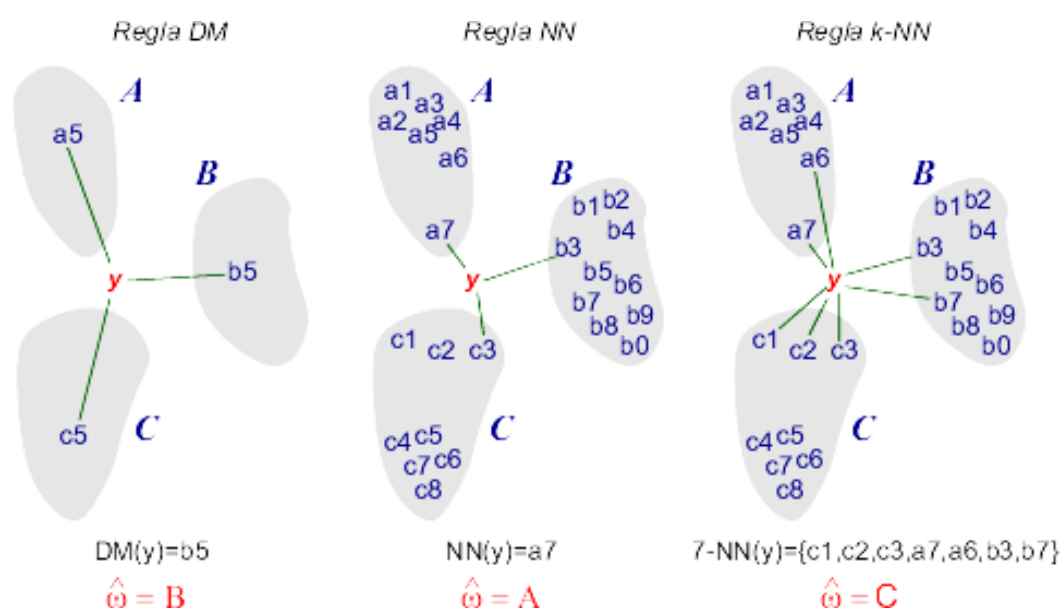
1. Distància mínima
2. Veí més proper
3. K- veïns més propers
4. Nearest neighbor classifier with rejection threshold

distàncies?

Euclidean, Mahalanobis, Bhattacharyya

10

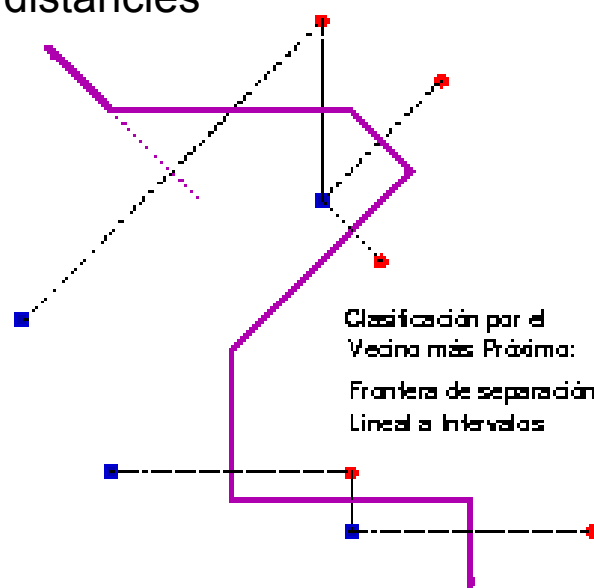
Classificadors basats en distàncies:



11

Classificadors basats en distàncies:

Exemple de fronteres de decisió en classificadors basats en distàncies



12

Template matching:

Definition:

match-based consists on moving an image (sub-image or template) across the current image and computing a similarity measure at each position.

Open questions:

1. *Similarity measure? color, texture?*
2. *Histogram comparison?*
3. *Coarse-to-fine-search and subsampling: increasing the step size for moving the template across the image and subsampling the template*

13

Arbres de decisió (I):

Cancerous Cell Example

Classes: { Cancerous, Healthy }

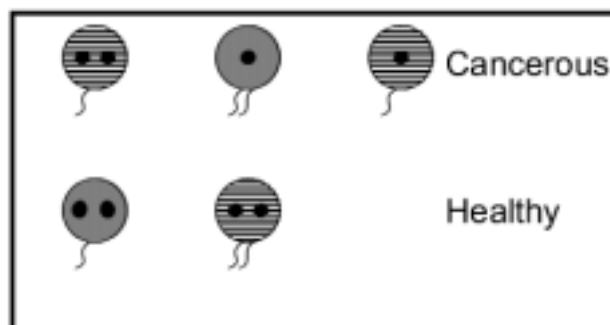
Features:

Cell Body: { Shaded, Striped }

#Nuclei: { 1, 2 }

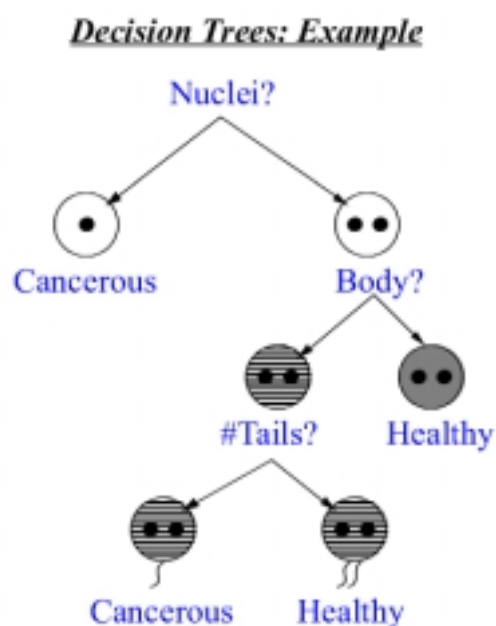
#Tails: { 1, 2 }

Examples:



14

Arbres de decisió (II):



15

Arbres de decisió (III):

Open questions:

1. *n classes?*
2. *Selection function (node): monothetic trees, which use only one attribute at each internal node; polythetic, which can use more than one*
3. *Prunning? Pre or post?*
4. *Stopping criteria?*
5. *REVIEW: Decision tree tesi jordif*

16

Aprenantatge No-Supervisat:

Open questions:

- 1. Number of classes?*
- 2. Labelling the training set*