innovate achieve lead

BITS Pilani

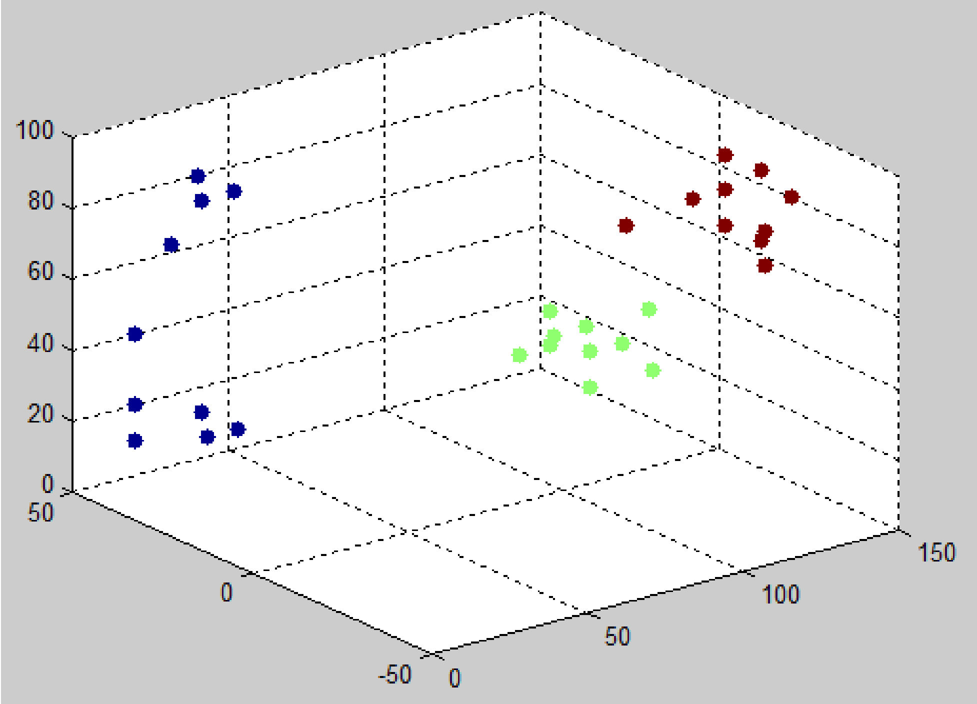
Machine Learning (IS ZC464) Session 3: Classification

Classification

- The goal of classification is to take an input vector x and to assign it to one of K discrete classes C_k where k = 1, 2, 3, …, K
- Examples
  - Email: Spam / Not Spam?
  - Online Transactions: Fraudulent (Yes / No)?
  - Tumor: Malignant / Benign ?

Decision Regions

- Training data is viewed to be plotted in a d-dimensional space where d is the number of features used.
- A test data is also viewed to be mapped in the same space.
- Similarity (or closeness) of the test data from the cluster of training classes is obtained.
- The nearest class is assigned to the test data



Binary Classification

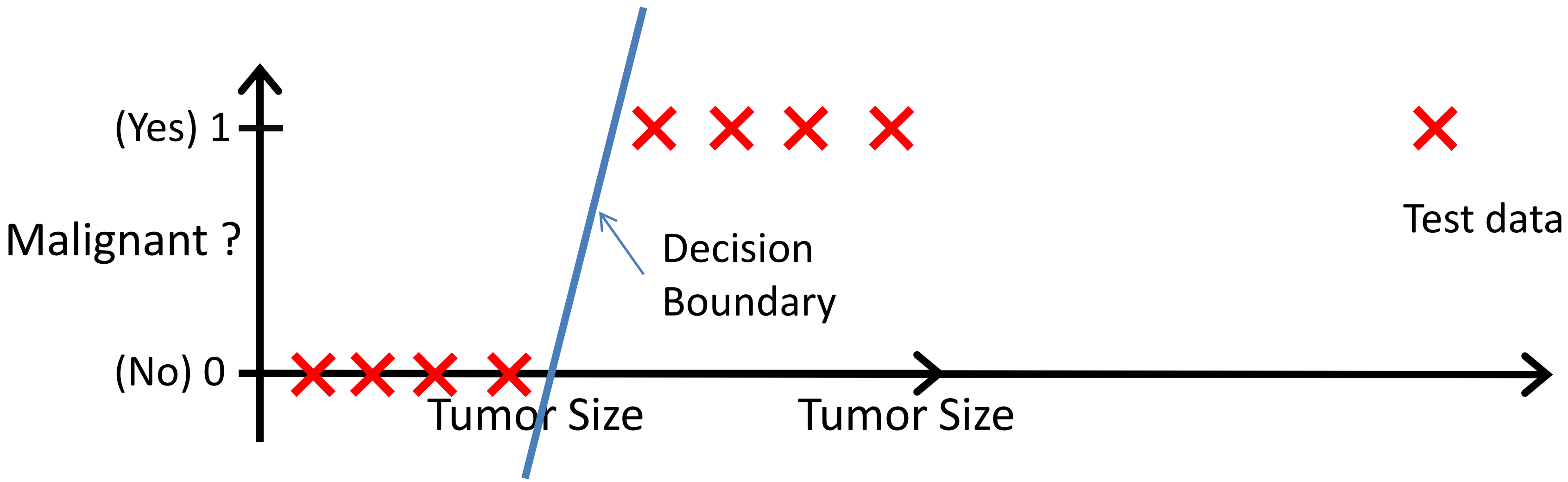
- Only two classes

$$y \in \{0, 1\}$$

0: “Negative Class” (e.g., benign tumor)

1: “Positive Class” (e.g., malignant tumor)

Example of a Decision Boundary



Threshold classifier output $h_\theta(x)$ at 0.5:

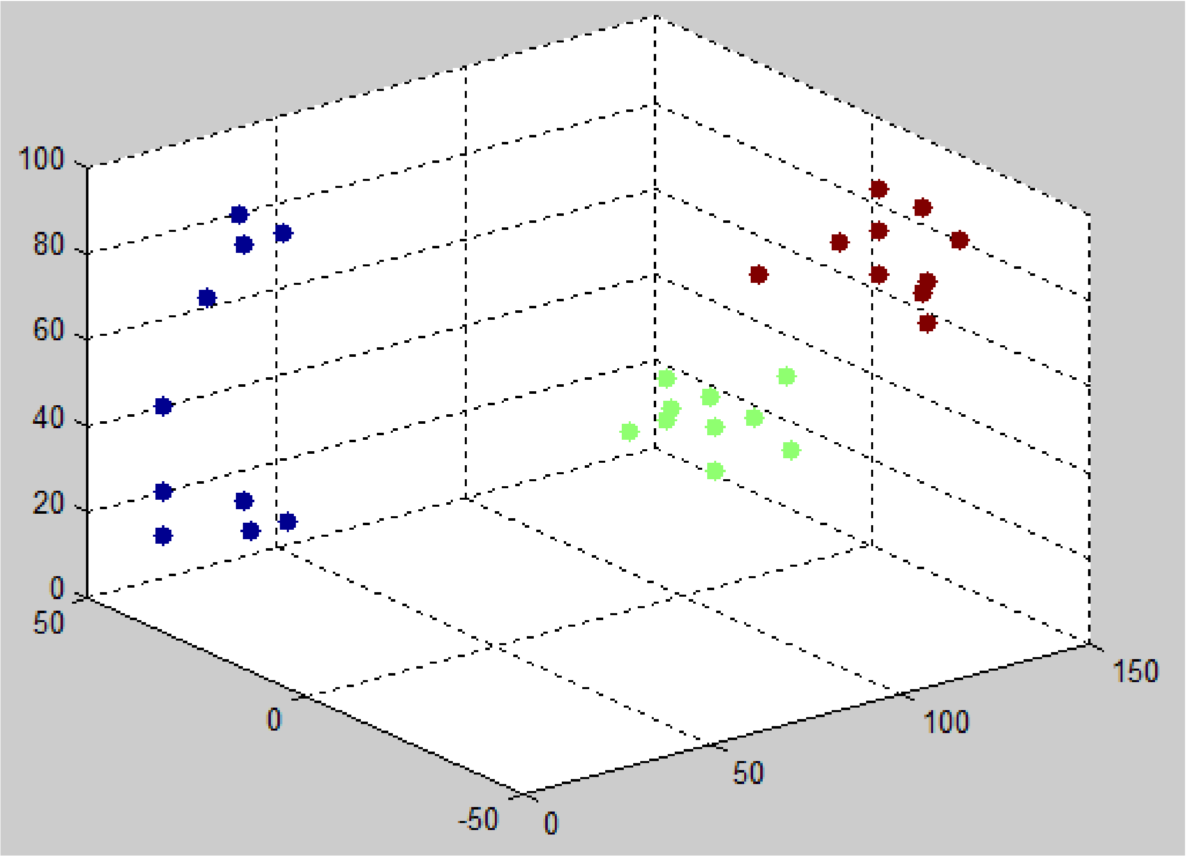
If $h_\theta(x) \geq 0.5$, predict "y = 1"

If $h_\theta(x) < 0.5$, predict "y = 0"

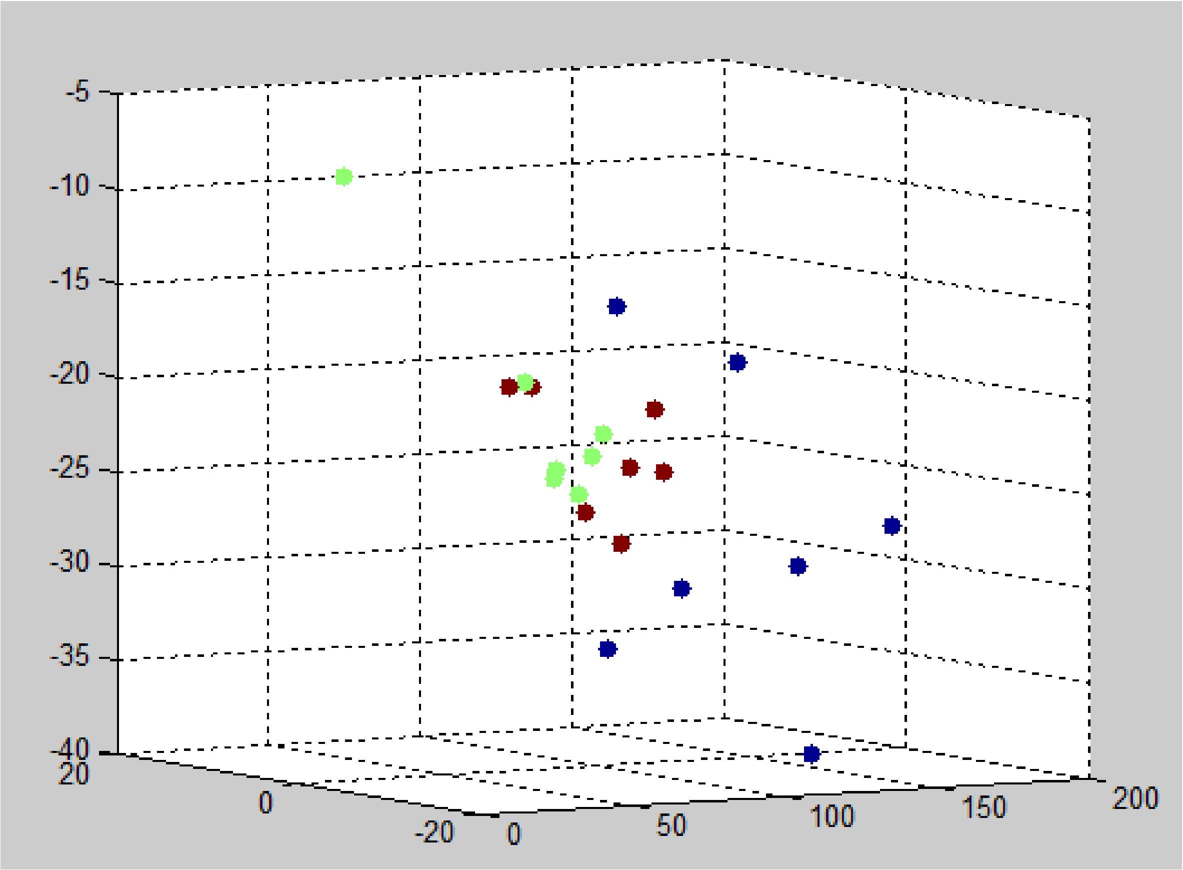
Solving Classification Problems

- Require the decision boundaries (or surfaces in hyper dimensional space) to be identified based on the training data.
- The decision boundary may be a line, a polynomial curve or a surface.
- The decision boundary can be represented as a hypothesis $h_\theta(x)$

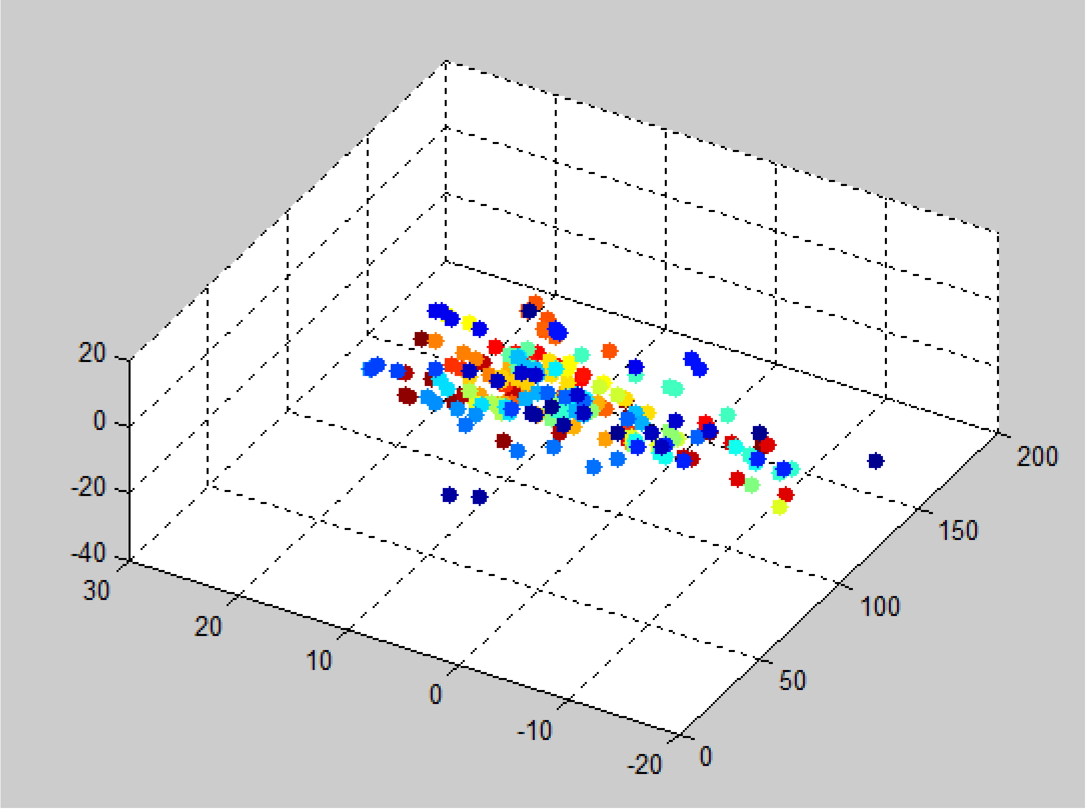
Linearly Separable Non-Face Data



Each face is a point in the n-dimensional space. (ORL face data for three persons)



The points in the n-dimensional space cannot be clustered (colorwise) by hyperplanes.



Discriminant Functions

- Represent the decision boundary
- Discriminant functions are obtained by taking a linear function of the input vector (feature vector).
- Define $y(x) = w_0 + w_1x+w_2x+\ldots w_Dx$
- Take a simple case

$$y(x) = w_0 + w_1x$$

- This is the equation of line.
- How does this behave as a decision boundary

Example

- Consider the following training data
- Class 1: <1,2>, <1,1>, <2,1>
- Class 2: <3,3>, 3,4>, <4,3>
- Can view a decision boundary as a line separating two classes
- The equation of the line is

$$x_2 = -x_1 + 1$$

(not using y deliberately as used for target)