

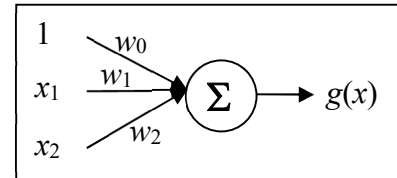
Ch04 – Classifiers

Tutorial

1. For a two-class problem:

$$g(x) = w_0 + w_1x_1 + w_2x_2$$

where x_1 and x_2 is the feature 1 and 2 of a sample.
If $g(x) > 0$, sample x belongs to class 1;
otherwise, it is class 2. Given that w_0 , w_1 and w_2
are 3.5, 5.6 and 2.5 respectively. What is the
value of $g(x)$ and the class when



$$g(x) = X'W$$

where $X = (1 \ x_1 \ x_2)^t$ and
 $W = (w_0 \ w_1 \ w_2)^t$

- a) $x = (-3.0 \ 2.1)$ b) $x = (1.5 \ 3.6)$ c) $x = (-2.0 \ 3.08)$

2. Mary wants to solve a 4-class problem:

	x_1	x_2	ω
1	1	5	1
2	3	4	1
3	4	6	2
4	4	8	2

	x_1	x_2	ω
5	6	2	3
6	8	3	3
7	7	5	4
8	8	5	4

However, she only can use binary classifiers.

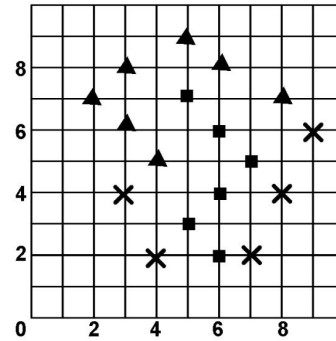
- Can you suggest two methods to help Mary?
- How many binary classifiers should be trained for each method? What are they?
How do these methods work?
- Please discuss the pros and cons of these two methods.

3.

a) According to K-nn, which class the followings samples should belong to when $K = 1$?

i) (5, 5) ii) (7.5, 4) iii) (6, 7)

b) If $K = 3$, what is the answer of part (a)?



4. Bob wants to try a Decision Tree algorithm on a simple dataset listed as follows. The dataset has 8 instances each of which has 3 attributes and a class label. The attributes are abstracted as x_1 , x_2 , and x_3 and the class label as Y . All of them are binary variables.

Sample	x_1	x_2	x_3	Y
1	0	0	0	0
2	0	0	1	0
3	0	1	0	1
4	0	1	1	1
5	1	0	1	1
6	1	0	1	1
7	1	1	0	0
8	1	1	0	0

- What is the value of the Entropy of class labels, $H(Y)$, in the dataset?
- What are the Information Gains of Y with respect to each of the three attributes, $IG(Y, x_1)$, $IG(Y, x_2)$ and $IG(Y, x_3)$?
- Which attribute should be chosen for the root according to the algorithm?
- Build the rest of the tree according to the Information Gains.
- Bob wants to prune the tree as he thinks the tree is too complicated. Which node should be deleted and why?