1. The table below shows a data set **Z** with 10 objects described by 4 features.

Object	x_1	x_2	x_3	x_4	Label
1	5	2	1	4	3
2	6	1	4	0	2
3	-5	1	-4	7	2
4	0	1	3	2	2
5	6	2	1	-5	1
6	4	2	1	7	1
7	-6	3	3	1	2
8	0	2	1	8	3
9	-4	11	5	2	2
10	4	7	2	2	3

(a) Regarding the notations introduced in class, what are n, c and N? (♣3)

$n = 4$ (dimensionality of the feature space)
$c = 3$ (number of classes)
$N = 10$ (number of objects in **Z**)

(b) Consider a classifier D, which consists of the following set of discriminant functions

$$g_1(\mathbf{x}) = x_1 + x_2$$
$$g_2(\mathbf{x}) = x_2 + 2x_3$$
$$g_3(\mathbf{x}) = 3x_4 - x_3$$

Run D on **Z** and give the (guessed) labels of all the objects. (♣3)

Object	$g_1(\mathbf{x})$	$g_2(\mathbf{x})$	$g_3(\mathbf{x})$	Maximum	Guessed label	True label
1	7	4	11	11	3	3
2	7	9	-4	9	2	2
3	-4	-7	25	25	3	2
4	1	7	3	7	2	2
5	8	4	-16	8	1	1
6	6	4	20	20	3	1
7	-3	9	0	9	2	2
8	2	4	23	23	3	3
9	7	21	1	21	2	2
10	11	11	4	11	1	3

(c) Give the confusion matrix of D and calculate the error of D. (♣3)

		Guessed labels		
		1	2	3
True labels	1	1	0	1
	2	0	4	1
	3	1	0	2

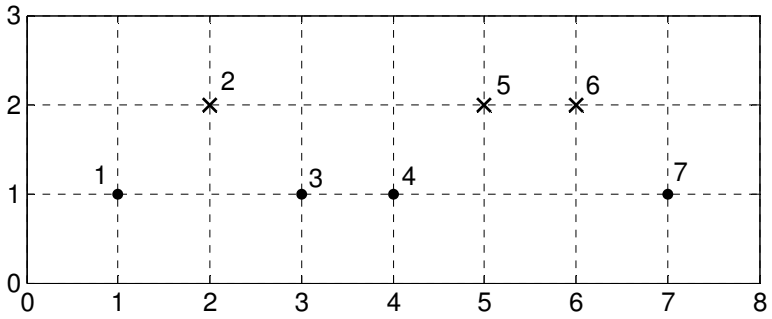
Error (D) = 3/10 = **0.3**.

(d) Estimate the prior probabilities for the classes from **Z**. (♣1)

P(class 1) = 2/10 = **0.2** P(class 2) = 5/10 = **0.5** P(class 3) = 3/10 = **0.3**

2. Consider the classification problem shown in the figure below. The data is depicted as points and the (true) class labels are shown by different markers: class 1 with dots and class 2 with crosses. The object numbers are given next to the markers. Your classifier, D, operates in the following way:

for a point **x** that you want to classify, find the nearest point in **Z**. Assign **x** to the class of its nearest neighbour (this is called *the nearest neighbour classifier*, more on it – later on the modules). In the case of a tie, resolve in favour of the class with the lower number (class 1 here).



(a) Find the leave-one-out error of *D*. (♣3)

Object	1	2	3	4	5	6	7
Nearest neighbour	2	1,3	4	3	6	5	6
Error?	Y	Y	N	N	N	N	Y

Error (*D*) = 3/7 = **0.4286**.

(b) Consider the following experiment. Stage1: Objects 1,2,3 and 4 are used for "training" (they form the data set from which we select the nearest neighbour) and objects 5,6 and 7 are used for testing. Stage 2: Objects 1,3,5,7 are used for training and objects 2,4,6 are used for testing. Stage 3: objects 3,4,5 and 6 are used for training and objects 1,2, and 7 are used for testing. The errors on the three testing sets are averaged to obtain an estimate of the error of *D*. Find this estimate. (♣4)

Stage 1

Object	5	6	7
Nearest neighbour	4	4	4
Error?	Y	Y	N

Error = 2/3

Stage 2

Object	2	4	6
Nearest neighbour	1,3	3	5
Error?	Y	N	N

Error = 1/3

Stage 3

Object	1	2	7
Nearest neighbour	3	3	6
Error?	N	Y	Y

Error = 2/3

Error (*D*) = 1/3 (2/3 + 1/3 + 2/3) = **5/9 = 0.5556**

(c) Which training-testing protocol have we used in **(b)**? (♣1)