1c) Given dataset:

F	label
1	0
2	0
3	1
6	1
6	1
7	0
10	0
11	0

→ Divide the set queally for training & testing data.

F. i) We need to train set using knn for K=3

Lets use Euclidean Distance between the data points:

Formula ⇒ $D = \sqrt{(x_2 - x_1)^2}$

→ Using Eucledian formula for data point to find out the predicted output

→ ① $d = \sqrt{(x_2 - x_1)^2}$

Train

F	label
1	0
2	0
3	1
6	1

① data point = 6

train point = 6

$\sqrt{(6-6)^2} = 0$

② data point = 6

train point = 3

$d = \sqrt{(6-3)^2} = 3$

③ data point = 6

train point = 2

$d = \sqrt{(6-2)^2} = 4$

④ data point = 6

train point = 1

$d = \sqrt{(6-1)^2} = 5$

0 3 4 5

F	label	Predicted label
6	1	
7	0	
10	0	
11	0	

⇒ Take nearest values i.e (1, 1, 0)

→ Predicted test point for 6 is 1
Similary for, 7, 10, 11 is (1, 1, 1)

F	Label	Predicted
6	1	1
7	0	1
10	0	1
11	0	1

10. 6) True Positive (TP): It occurs when actual value and predicted value is positive (1). (TP).

Confusion Matrix:

F	label	Predicted	TP/TN/FP/FN
6	1	1	TP
7	0	1	FP
10	0	1	FP
11	0	1	FP

We have $T_P = 1$ $T_N = 0$

$F_P = 3$ $F_N = 0$

$$\text{Accuracy} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} = \frac{1+0}{1+3+0+0} = \frac{1}{4}$$

$$\text{Sensitivity} = \frac{T_P}{T_P + F_N} = \frac{1}{1+0} = 1$$

$$\text{Specificity} = \frac{T_N}{F_P + T_N} = \frac{0}{3+0} = 0.$$