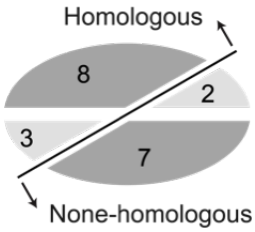
5 Exercise solutions – Model evaluation

1. **Basic measures from confusion matrix**

   The oval representation below shows that a model has classified a test data set with 10 positives and 10 negatives and produced four outcomes.



   (a) Fill each blank cell with one of the four classification outcomes - TP, FP, TN, and FN.

		Test data	
		Homologous	Non-homologous
Model classification	Homologous	TP	FP
	Non-homologous	FN	TN

   (b) Make a confusion matrix for the oval representation.

		Test data	
		Homologous	Non-homologous
Model classification	Homologous	8	3
	Non-homologous	2	7

   (c) Calculate the following basic evaluation measures for the oval representation. Round off the answer to two decimal places if necessary.

$$\text{Accuracy} = \frac{TP + TN}{P + N} = \frac{15}{20} = 0.75$$

$$\text{Error rate} = \frac{FP + FN}{P + N} = \frac{5}{20} = 0.25$$

$$\text{Sensitivity} = \frac{TP}{P} = \frac{8}{10} = 0.8$$

$$\text{Specificity} = \frac{TN}{N} = \frac{7}{10} = 0.7$$

$$\text{Precision} = \frac{TP}{TP + FP} = \frac{8}{11} = 0.73$$

2. **Measures with multiple thresholds**

Create multiple confusion matrices by considering all possible threshold values. Assume that the test data set contains two positives and two negatives. The table below shows the scores given by a model that gives higher scores for the alignments with higher similarities.

Test set label	P	P	N	N
Model score	2.1	3.1	2.3	1.2

(a) Fill the labels that match the sorted scores

Test set label	**P**	**N**	**P**	**N**
Model score	3.1	2.3	2.1	1.2

(b) Fill the labels predicted by different threshold values.

1. 3.5

Classified label	**N**	**N**	**N**	**N**
Model score	3.1	2.3	2.1	1.2

Threshold 3.5

2. 2.7

Classified label	**P**	**N**	**N**	**N**
Model score	3.1	2.3	2.1	1.2

Threshold 2.7

3. 2.2

Classified label	**P**	**P**	**N**	**N**
Model score	3.1	2.3	2.1	1.2

Threshold 2.2

4. 1.6

Classified label	**P**	**P**	**P**	**N**
Model score	3.1	2.3	2.1	1.2

Threshold 1.6

5. 1.0

Classified label	**P**	**P**	**P**	**P**
Model score	3.1	2.3	2.1	1.2

Threshold 1.0

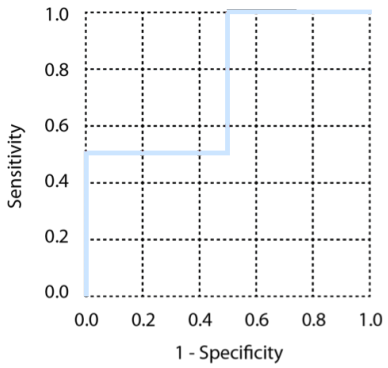
(c) Use the labels in (a) and (b) and calculate TP, FP, TN, and FN for all threshold values.

Threshold	TP	FP	TN	FN
3.5	0	0	2	2
2.7	1	0	2	1
2.2	1	1	1	1
1.6	2	1	1	0
1	2	2	0	0

(d) Use the result in (c) and calculate basic evaluation measures. Round off the answer to one decimal place if necessary.

Threshold	Specificity	1 - specificity	Sensitivity
3.5	1	0	0
2.7	1	0	0.5
2.2	0.5	0.5	0.5
1.6	0.5	0.5	1
1	0	1	1

(e) Draw a ROC curve for the calculated evaluation measures in (d).



(f) Calculate the area under the curve of the curve in (e).

Solution: 0.75

(g) Evaluate the ROC curve in your own words.

Solution: The model performs better than random classifiers