Practice Exercise 4 – Evaluation Metrics (solution)

Question 1: Is there a disadvantage/limitation of ROC curve for analysing the classification performance?

ROC provides an indication about model performance only for binary classification cases. For multi-class classification, we need to find a way around (e.g. one vs rest) to see at various plots.

Question 2: Given the confusion matrix below as a measure of classifier's performance to predict class labels for a 2-class problem, answer the following:

		Actual class label	
		Diseased	Healthy
Predicted label	Diseased	21	3
	Healthy	4	37

a) Compute the overall classification accuracy.

(21+37)/(21+37+3+4)

c) Compute the true positive, true negative, false positive, false negative.

TP = 21; TN = 37; FP = 3; FN = 4

d) Compute the sensitivity and specificity.

Sen = TP/(TP+FN) = 21/25

Spec = TN/(TN+FP) = 37/40

Question 4: The ROC curve is a way to visualise the combined sensitivity and specificity performance of a classification system. Answer the following:

a) Is there a way to quantitatively measure the classification performance from the ROC curve?

Yes, one can compute area under the ROC (AUROC) to obtain a single quantitative measure of classification performance.

b) By analysis of ROC curve, is there a way to 'perturb' the algorithm such that the sensitivity and/or specificity can be enhanced?

Yes, the probability threshold of assigning class labels can be 'perturbed' to change the class assignments. By default, the threshold is 0.5 but this can be modified. A 'perturbation' will result in different assignments and thus different sensitivity/specificity. This is the case for classifiers which produce probability estimate for each class (e.g. Bayesian classifier).