1 ROC

Given the table below please complete the table below and plot the ROC curve for 0, 0.2, 0.4, 0.6, 0.8, and 1 on the graph in the answer booklet. Please explain whether you would choose to use Classifier 1 or Classifier 2 and why.

Class	$Pr(P\|A)$ Classifier1	$Pr(P\|A)$ for Classifier2
P	0.83	0.92
N	0.78	0.62
P	0.62	0.52
N	0.48	0.49
N	0.32	0.38
N	0.22	0.28

2 Evaluation

Suppose you do 2-fold cross validation, 10-fold cross validation or leave-one-out, or use a 70%/30% train/validation single split.

- What effect will this have on your results?
- Which will give you the best representative value to the "unseen test set"?
- Does it matter how large your original dataset is? Will you get a different answer for a very big or a very small dataset?
- Which will be the fastest in terms of computation?

3 Statistical test

Suppose that we want to select between two prediction models, M_1 and M_2. We have performed 10 rounds of 10-fold cross-validation on each model, where the same data partitioning in round i is used for both M_1 and M_2. The error rates obtained from M_1 are 30.5, 32.2, 20.7, 20.6, 31.0, 41.0, 27.7, 26.0, 21.5, 26.0. The error rates for M_2 are 22.4, 14.5, 22.4, 19.6, 20.7, 20.4, 22.1, 19.4, 16.2, 35.0.

Comment on whether one model is significantly better than the other considering a significance level of 1%.

Hint: Make yourself familiar with the *Student's t-test*.

4 Normalisation

Suppose we have some 8-bit grayscale image data encoded as a matrix:

$$\mathbf{X} = \begin{bmatrix} 0 & 0 & 255 & 120 \\ 0 & 56 & 56 & 25 \\ 40 & 200 & 186 & 15 \\ 78 & 160 & 55 & 0 \end{bmatrix}$$

Use min-max scaling to normalise this data to the interval $[0, 1]$ (represented using real numbers).