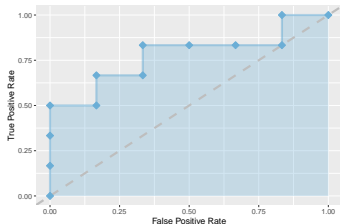


Introduction to Machine Learning

Evaluation: Measures for Binary Classification: ROC Visualization

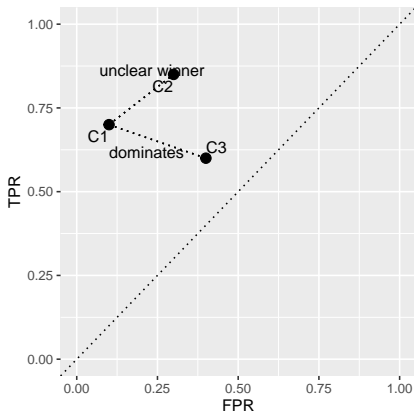


Learning goals

- Understand the ROC curve
- Be able to compute a ROC curve manually
- Understand the definition of AUC and what a certain value of AUC means (and what not!)

LABELS: ROC SPACE

Plot True Positive Rate and False Positive Rate:



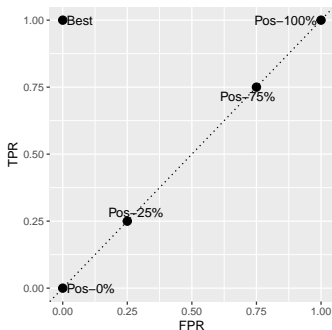
		True Class y	
		+	-
Pred. \hat{y}	+	TP	FP
	-	FN	TN

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

LABELS: ROC SPACE

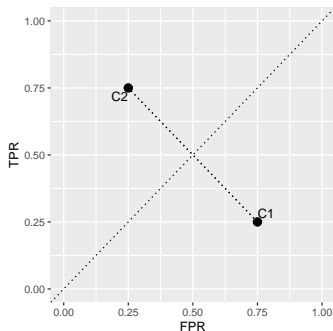
- The best classifier lies on the top-left corner
- The diagonal \approx random labels (with different proportions).
Assign positive x as "pos" with 25% probability $\rightarrow TPR = 0.25$.
Assign negative x as "pos" with 25% probability $\rightarrow FPR = 0.25$.



LABELS: ROC SPACE

In practice, we should never obtain a classifier below the diagonal.

Inverting the predicted labels ($0 \rightarrow 1$ and $1 \rightarrow 0$) will result in a reflection at the diagonal.



LABEL DISTRIBUTION IN TPR AND FPR

TPR and FPR are insensitive to the class distribution:

Not affected by changes in the ratio n_+/n_- (at prediction).

Example 1:

Proportion $n_+/n_- = 1$

	Actual Positive	Actual Negative
Pred. Positive	40	25
Pred. Negative	10	25

$$\text{MCE} = 35/100$$

$$\text{TPR} = 0.8$$

$$\text{FPR} = 0.5$$

Example 2:

Proportion $n_+/n_- = 2$

	Actual Positive	Actual Negative
Pred. Positive	80	25
Pred. Negative	20	25

$$\text{MCE} = 45/150 = 30/100$$

$$\text{TPR} = 0.8$$

$$\text{FPR} = 0.5$$

Note: If class proportions differ during training, the above is not true.
Estimated posterior probabilities can change!

FROM PROBABILITIES TO LABELS: ROC CURVE

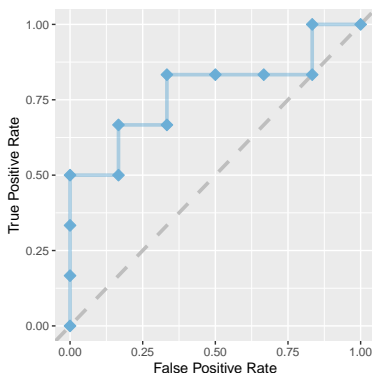
Remember: Both probabilistic and scoring classifiers can output classes by thresholding.

$$h(\mathbf{x}) := [\pi(\mathbf{x}) \geq c] \quad \text{or} \quad h(\mathbf{x}) = [f(\mathbf{x}) \geq c]$$

To draw a ROC curve:

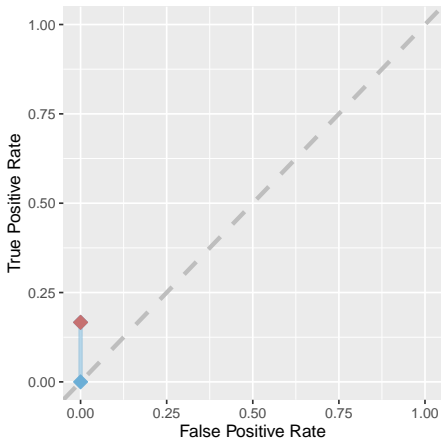
Iterate through all possible thresholds c

→ Visual inspection of all possible thresholds / results



ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



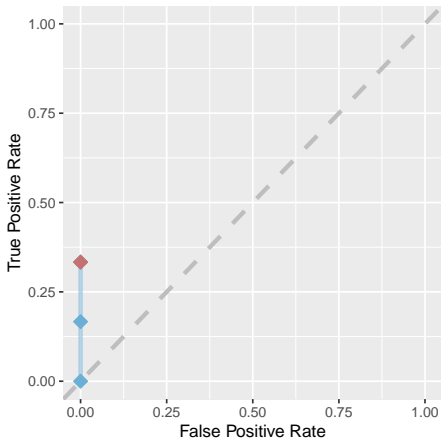
$$c = 0.9$$

$$\rightarrow \text{TPR} = 0.167$$

$$\rightarrow \text{FPR} = 0$$

ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



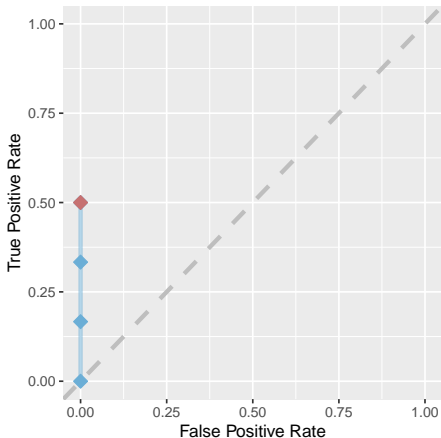
$$c = 0.85$$

$$\rightarrow \text{TPR} = 0.333$$

$$\rightarrow \text{FPR} = 0$$

ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



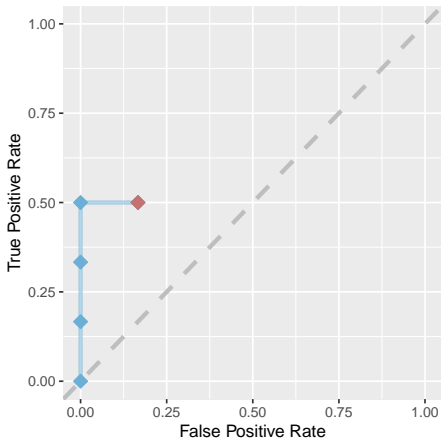
$$c = 0.66$$

$$\rightarrow \text{TPR} = 0.5$$

$$\rightarrow \text{FPR} = 0$$

ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



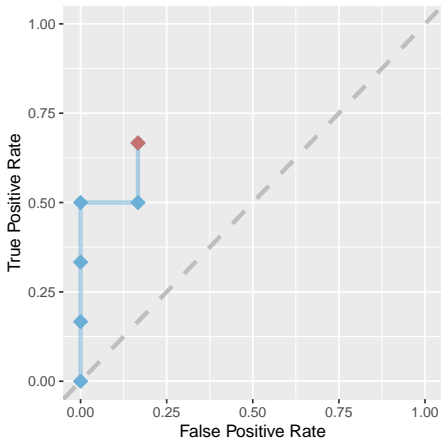
$$c = 0.6$$

$$\rightarrow \text{TPR} = 0.5$$

$$\rightarrow \text{FPR} = 0.167$$

ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



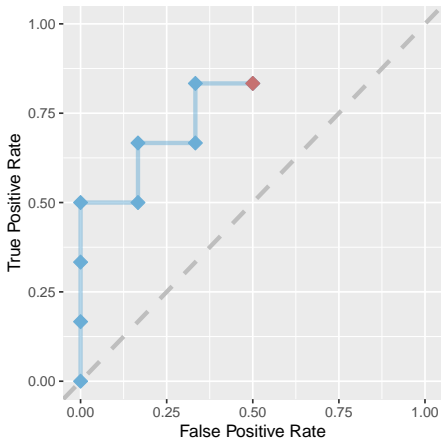
$$c = 0.55$$

$$\rightarrow \text{TPR} = 0.667$$

$$\rightarrow \text{FPR} = 0.167$$

ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



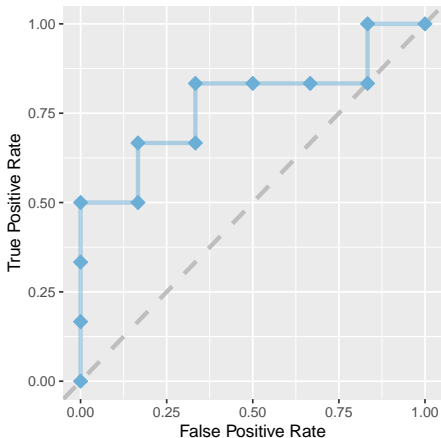
$$c = 0.3$$

$$\rightarrow \text{TPR} = 0.833$$

$$\rightarrow \text{FPR} = 0.5$$

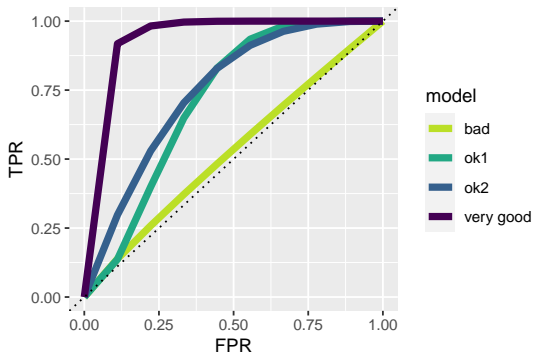
ROC CURVE

#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



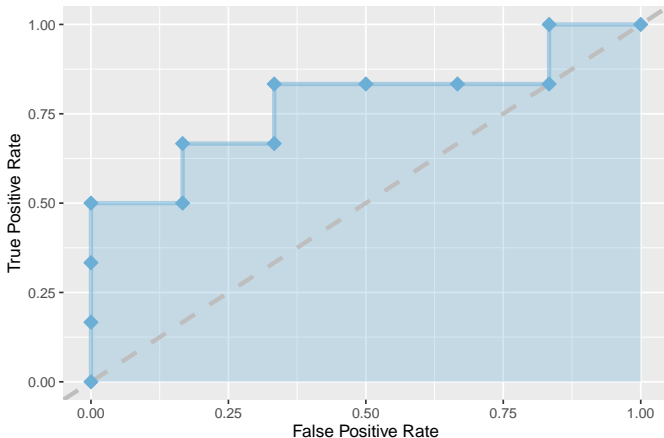
ROC CURVE

- The closer the curve to the top-left corner, the better
- If ROC curves cross, a different model can be better in different parts of the ROC space



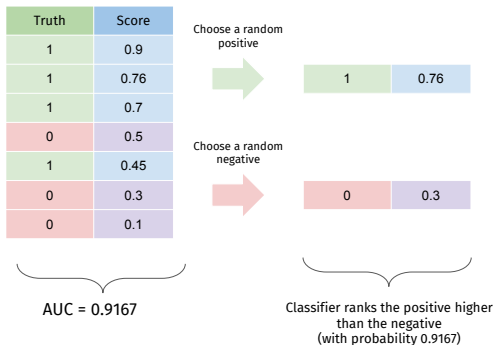
AUC: AREA UNDER ROC CURVE

- The AUC (in $[0,1]$) is a single metric to evaluate scoring classifiers
- AUC = 1: Perfect classifier
- AUC = 0.5: Randomly ordered



AUC: AREA UNDER ROC CURVE

Interpretation: Probability that classifier ranks a random positive higher than a random negative observation



PARTIAL AUC

- Sometimes it can be useful to look at a specific region under the ROC curve \Rightarrow partial AUC (pAUC).
- Examples: focus on a region with low FPR or a region with high TPR:

