Introduction to Machine Learning

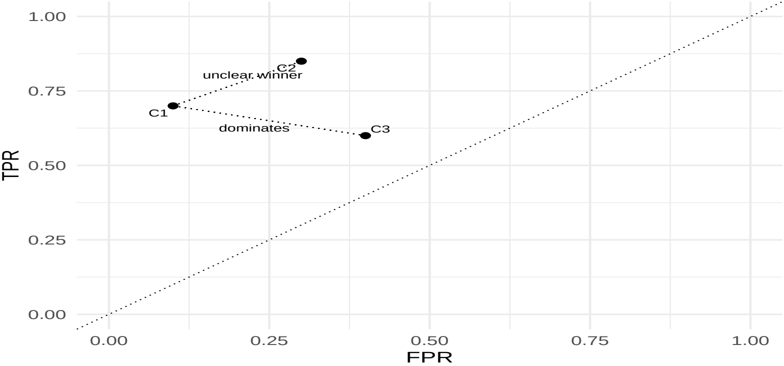
Evaluation: ROC Analysis 2

Bernd Bischl, Christoph Molnar, Daniel Schalk, Fabian Scheipl

Department of Statistics – LMU Munich

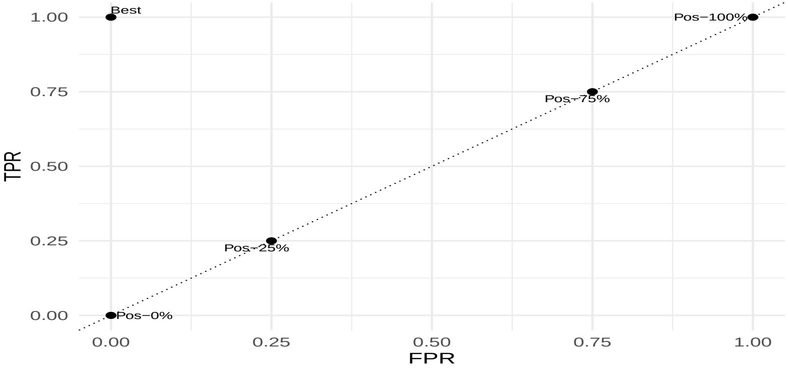
ROC SPACE

- We characterize a classifier by its TPR and FPR values and plot them in a coordinate system
- We could also use 2 different ROC metrics which define a trade-off, like TPR and PPV!



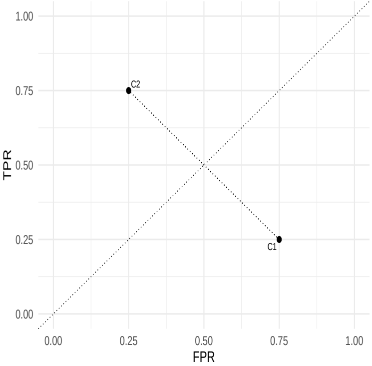
ROC SPACE

- The best classifier lies on the top-left corner
- The diagonal is worst, where classifiers produce random labels (with different proportions). If each positive x will be randomly classified with 25% as "pos", $TPR = 0.25$. If we assign each negative x randomly to "pos", $FPR = 0.25$.



ROC SPACE

In practice, we should never obtain a classifier below the diagonal, as inverting the predicted labels – $0 \rightarrow 1$ and $1 \rightarrow 0$ – will result in a reflection at the diagonal. Because this inverting results in $TPR2 = 1 - TPR1$ and $FPR2 = 1 - FPR1$



ROC AND LABEL DISTRIBUTION

ROC curves are insensitive to class distributions ROC curves are insensitive to the class distribution in the sense that they are not affected by changes in the ratio n_+/n_- (at prediction).

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

Here we have we have a proportion $n_+/n_- = 1$. MCE = 35/100. Now we double the size of the positive class and change the proportion to $n_+/n_- = 2$. MCE = 45/150 = 30/100.

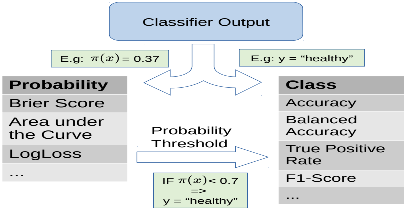
	Actual Positive	Actual Negative
Pred. Positive	80	25
Pred. Negative	20	25

$TPR = 0.8$ and $FPR = 0.5$ do not change.

NB: If we mess around with class proportions during training, the above is not true, as estimated posterior probabilities can drastically change!

SCORING CLASSIFIERS

- A scoring classifier is a model which outputs scores or probabilities, instead of discrete labels, and nearly all modern classifiers can do that.
- Thresholding flexibly converts measured probabilities to labels. Predict 1 (positive class) if $\hat{f}(x) > \tau$ else predict 0.
- Normally we could use $\tau = 0.5$ to convert, but for imbalanced or cost-sensitive situations another threshold could be much better.
- After thresholding, any metric defined on labels can be used.



ROC CURVE

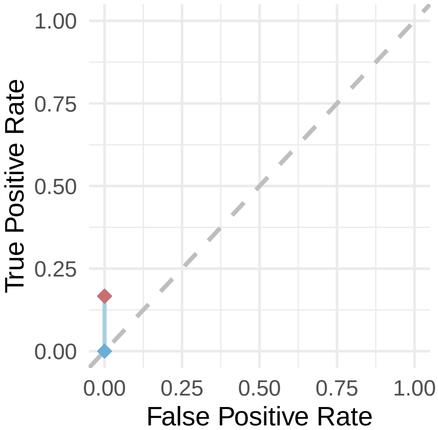
- Are based on thresholding classifiers
- We iterate through all possible threshold, and draw a point in the ROC space (FPR, TPR) for the resulting classifier
- The resulting plot is called an ROC curve
- Small thresholds will very liberally predict class 1, and result in a potentially higher FPR, but also higher TPR
- High thresholds will very conservatively predict class 1, and result in a lower FPR and TPR
- As we have not defined the trade-off between false postives and false negative costs, we cannot easily select the "best" threshold; but a visual inspection of all possible results seems useful

ROC CURVE

- Rank test observations on decreasing score
- Set $\alpha = 1$, so we start in $(0, 0)$; we predict everything as "neg"
- For each observation x (in the decreasing order).
  - Reduce threshold, so prediction for next observation changes
  - If x is "pos", move TPR $1/n_+$ up, as we have one TP more
  - If x is "neg", move FPR $1/n_-$ right, as we have one FP more

ROC CURVE

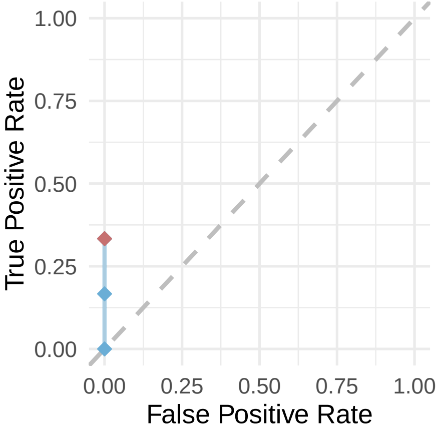
#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



Set threshold $\tau = 0.9$ yields TPR 0.167 and FPR 0.

ROC CURVE

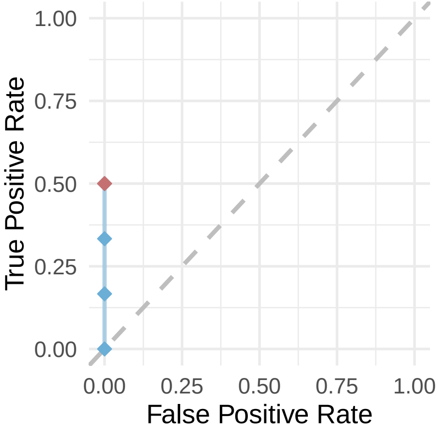
#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



Set threshold $\tau = 0.85$ yields TPR 0.333 and FPR 0.

ROC CURVE

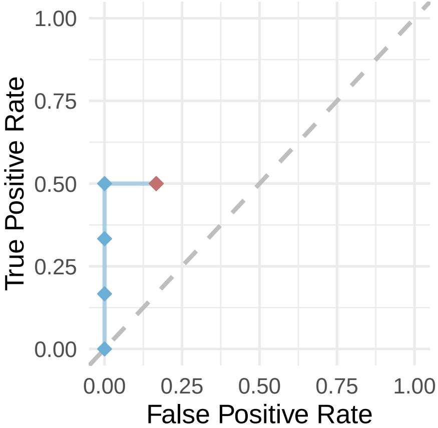
#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



Set threshold $\tau = 0.66$ yields TPR 0.5 and FPR 0.

ROC CURVE

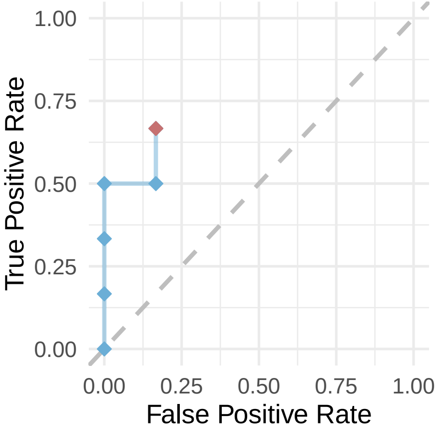
#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



Set threshold $\tau = 0.6$ yields TPR 0.5 and FPR 0.167.

ROC CURVE

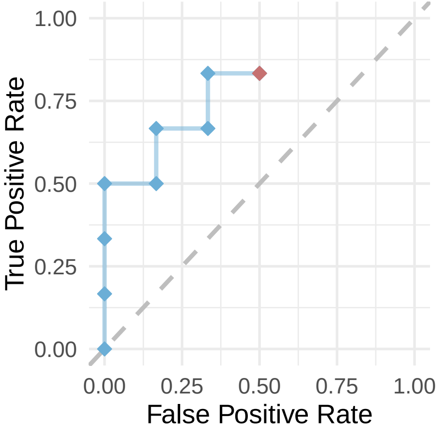
#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



Set threshold $\tau = 0.55$ yields TPR 0.667 and FPR 0.167.

ROC CURVE

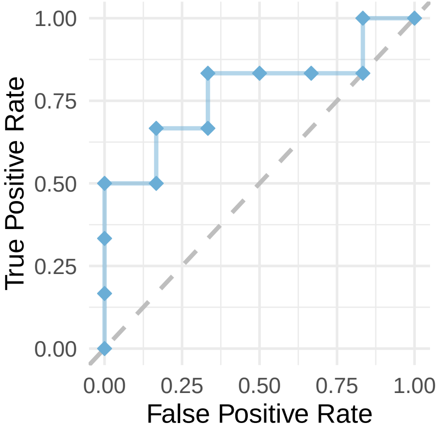
#	Truth	Score
1	Pos	0.95
2	Pos	0.86
3	Pos	0.69
4	Neg	0.65
5	Pos	0.59
6	Neg	0.52
7	Pos	0.51
8	Neg	0.39
9	Neg	0.28
10	Neg	0.18
11	Pos	0.15
12	Neg	0.06



Set threshold $\tau = 0.3$ yields TPR 0.833 and FPR 0.5.

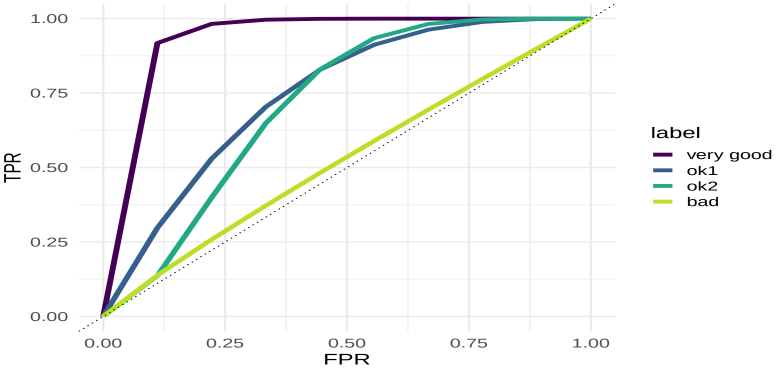
ROC CURVE

#	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



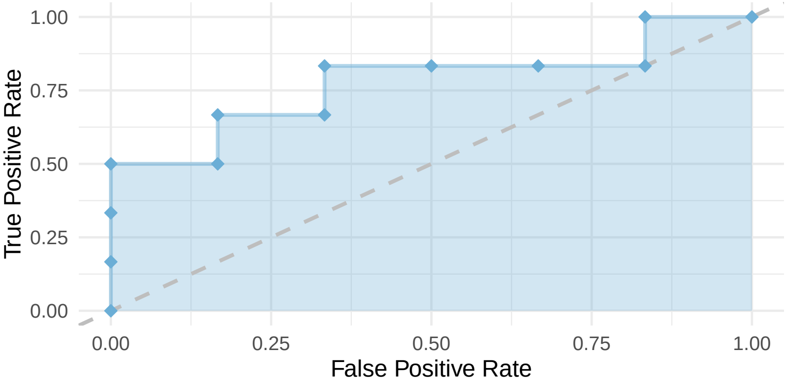
ROC CURVE

- The closer the curve to the top-left corner, the better
- Unfortunately, ROC curves can also cross
- Then, depending on costs and what you want, 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 = 0: Perfect, with inverted labels



AUC: AREA UNDER ROC CURVE

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

Truth	Score
1	0.9
1	0.76
1	0.7
0	0.5
1	0.45
0	0.3
0	0.1

Choose a random positive

1	0.76

Choose a random negative

0	0.3

AUC = 0.9167

Classifier ranks the positive higher than the negative (with probability 0.9167)

PARTIAL AUC

- Sometimes it can be useful to look at a specific region under the ROC curve $\Rightarrow$ partial AUC (pAUC).
- Let $0 \leq c_1 < c_2 \leq 1$ define a region.
- For example, one could focus on a region with low FPR ($c_1 = 0, c_2 = 0.2$) or a region with high TPR ($c_1 = 0.8, c_2 = 1$):