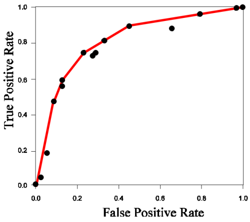
ROC curves

March 3, 2017

Table of Contents

Table of Contents

What are ROC curves?



- A simple way to evaluate classifiers
- Originated in electronic signal detection theory (e.g., radar to detect airplanes)
- Popular in many domains including radiology and computer vision
- ROC = Receiver Operating Characteristic

Table of Contents

Classifiers

A classifier assigns one of a predefined set of classes to an object.

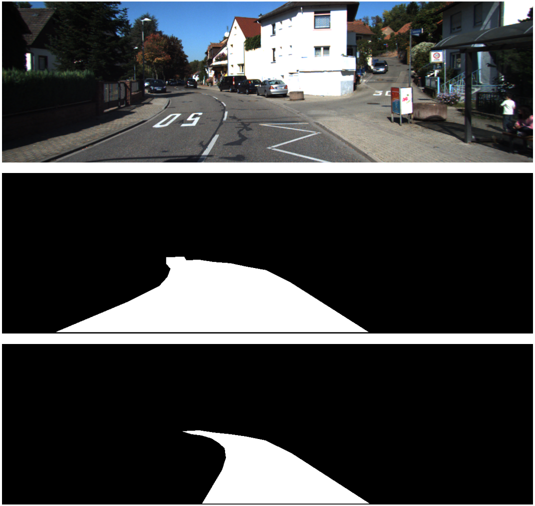
Example

Assign one of the classes

- on my lane
- not on my lane, but on the road
- not on the road

to each pixel in an image.

Classifiers



(i) Original image, (ii) road pixels, (iii) lane pixels.

Binary classifiers

In a binary classifier we only have two classes.
For example: **road** / **not road.**

(i) Original image, (ii) road pixels

Table of Contents

How to assess the performance of a binary classifier?

False Positive (FP)

A pixel is classified as belonging to the road, while it is not

False Negative (FN)

A pixel is classified as not on the road, while it is

2-class Confusion Matrix

True class	Predicted class	
	positive	negative
positive (#P)	#TP	#FN
negative (#N)	#FP	#TN

Confusion matrix is also called a contingency matrix.

2-class Confusion Matrix

True class	Predicted class	
	positive	negative
positive (#P)	#TP	#P - #TP
negative (#N)	#FP	#N - #FP

Classifier performance rates:

- true positive rate = TPR = (#TP)/(#P)
- false positive rate = FPR = (#FP)/(#N)

Which classifier is best?

True	Predicted pos	Predicted neg
pos	40	60
neg	30	70

TPR = 0.4, FPR = 0.3

True	Predicted pos	Predicted neg
pos	70	30
neg	50	50

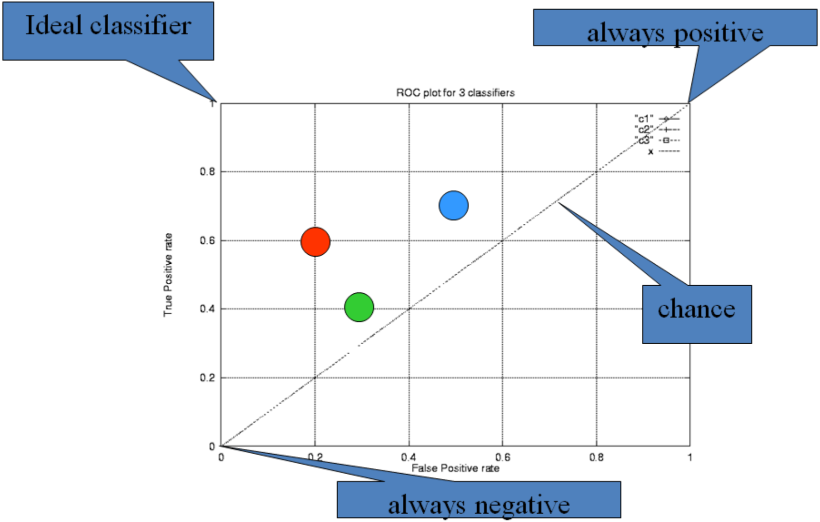
TPR = 0.7, FPR = 0.5

True	Predicted pos	Predicted neg
pos	60	40
neg	20	80

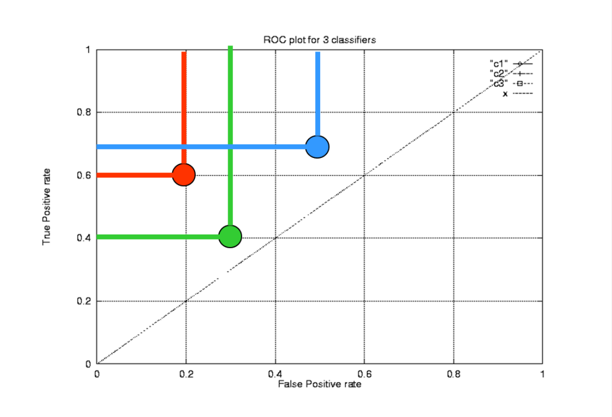
TPR = 0.6, FPR = 0.2

Table of Contents

A ROC plot will tell us more

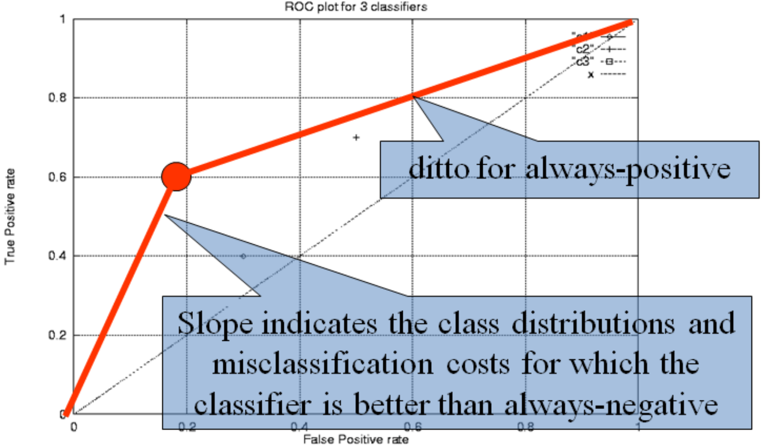


Dominance

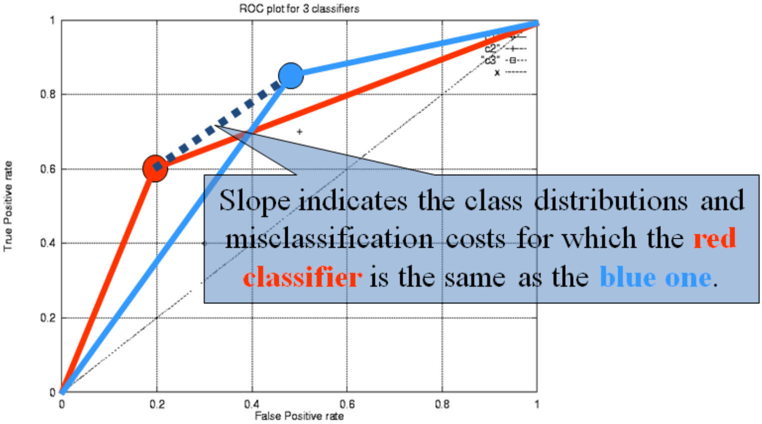


Red classifier is always better than green classifier.
TPR of red classifier is not better than TPR of blue classifier, however, its TPR/FPR ratio is better.

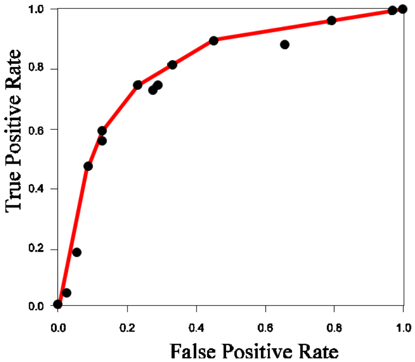
Operating Range



Convex hull

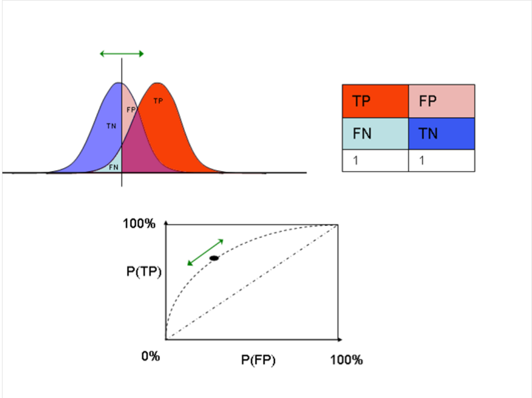

ROC plot for 3 classifiers
True Positive rate
False Positive rate
0
0.2
0.4
0.6
0.8
1
"c2"
"c3"
x
Slope indicates the class distributions and misclassification costs for which the red classifier is the same as the blue one.

ROC curve



ROC curve displays classifier performance when we change the discrimination threshold.
For lower thresholds, more true positives are accepted, but false positive rate increases.

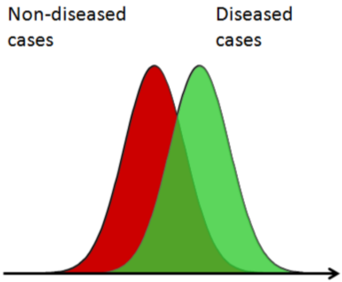
ROC curve for a binary classifier with known class distributions



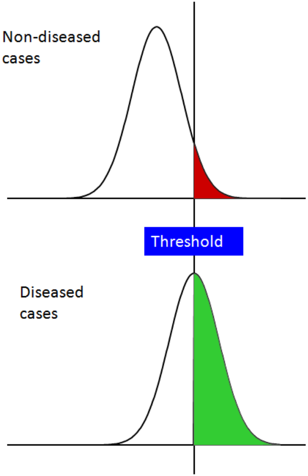
Binary classifier with single threshold. When threshold moves to the left, TPR increases, but also FPR. When threshold moves to the right, FPR decreases, but also TPR.

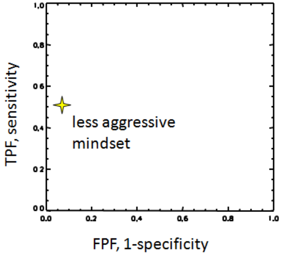
Example

Abnormal liver enzyme values indicate liver malfunctioning (ALT, AST, ALP, Albumin, Bilirubin, GGT, LD, PT). Suppose we have the following distributions for a certain enzyme value.



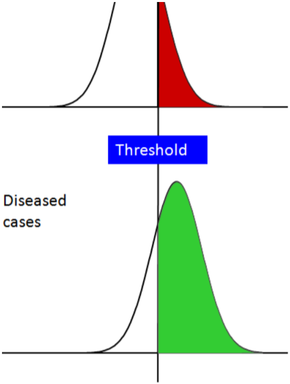
Example

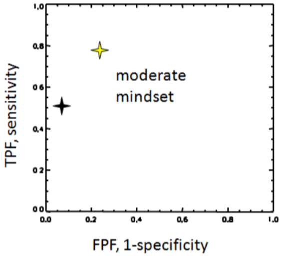




An optimistic doctor

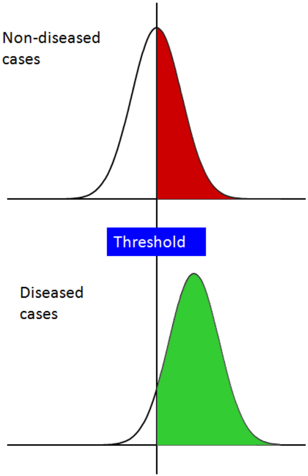
Example

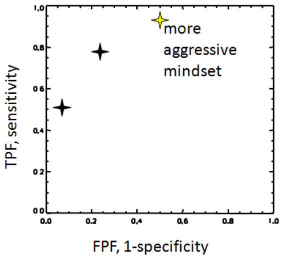




A not so optimistic doctor

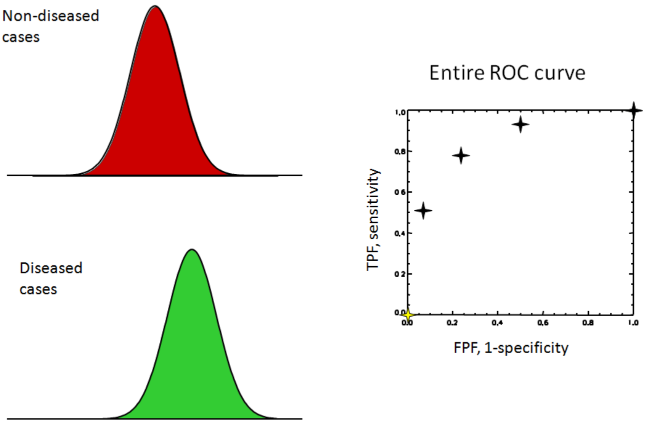
Example





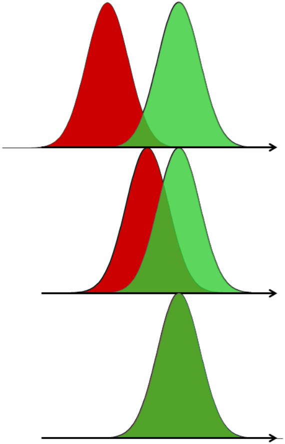
A pessimistic doctor

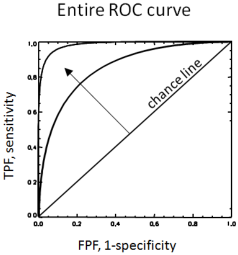
Example



A very pessimistic doctor

Example





Fortunately there may be other indicators, with better discrimination between diseased or non-diseased.

Recall, precision and F1 score

- recall = # TP / (# TP + # FN)
- precision = # TP / (# TP + # FP)
- F1 score = harmonic mean of recall and precision
- F1 score = $2 \frac{1}{\frac{1}{recall} + \frac{1}{precision}} = 2 \frac{precision \cdot recall}{precision + recall}$

Other classifier performance measures

sensitivity or true positive rate (TPR)

eqv. with hit rate, recall

$$TPR = TP/P = TP/(TP + FN)$$

specificity (SPC) or True Negative Rate

$$SPC = TN/N = TN/(FP + TN)$$

precision or positive predictive value (PPV)

$$PPV = TP/(TP + FP)$$

negative predictive value (NPV)

$$NPV = TN/(TN + FN)$$

fall-out or false positive rate (FPR)

$$FPR = FP/N = FP/(FP + TN) = 1 - SPC$$

false discovery rate (FDR)

$$FDR = FP/(FP + TP) = 1 - PPV$$

Miss Rate or False Negative Rate (FNR)

$$FNR = FN/P = FN/(FN + TP)$$

Other classifier performance measures

accuracy (ACC)

$$ACC = (TP + TN)/(P + N)$$

F1 score

is the harmonic mean of precision and sensitivity

$$F1 = 2TP/(2TP + FP + FN)$$

Matthews correlation coefficient (MCC)

$$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Informedness = Sensitivity + Specificity - 1

Markedness = Precision + NPV - 1