Some notes on HW #2

How do we evaluate and compare classifiers?

What’s an ROC curve?

Quantifying Quality of a Classifier

Every instance has an unknown *actual* +/- label, and also a *predicted* +/- label

- *Sensitivity,* aka *True Positive Rate:* what fraction of the actual +'s are found among the predicted +'s, independent of how actual -'s are classified
- *Specificity,* aka *False Negative Rate:* what fraction of the actual -'s are found among the predicted -'s, independent of how actual +'s are classified

"just say yes" has 100% sensitivity, but (likely) poor specificity; "just say no", the opposite.

EXAMPLE

“A diagnostic test with sensitivity 67% and specificity 91% is applied to 2030 people to look for a disorder with a population prevalence of 1.48%”

		The patient’s “true” status		
		Condition positive	Condition negative	
blood test outcome	Test outcome positive	**True positive** (TP) = 20	**False positive** (FP) = 180	Positive predictive value = TP / (TP + FP) = 20 / (20 + 180) = **10%**
	Test outcome negative	**False negative** (FN) = 10	**True negative** (TN) = 1820	Negative predictive value = TN / (FN + TN) = 1820 / (10 + 1820) ≈ **99.5%**
		Sensitivity = TP / (TP + FN) = 20 / (20 + 10) ≈ **67%**	**Specificity** = TN / (FP + TN) = 1820 / (180 + 1820) = **91%**	

https://en.wikipedia.org/wiki/Sensitivity_and_specificity

“All the Jargon That’s Fit To Print”

Many quantitative aspects of “Accuracy”

true positive (TP)
eqv. with hit

true negative (TN)
eqv. with correct rejection

false positive (FP)
eqv. with false alarm, Type I error

false negative (FN)
eqv. with miss, Type II error

---

sensitivity or true positive rate (TPR)
eqv. with hit rate, recall

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

specificity (SPC) or true negative rate

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

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 negative rate (FNR)

$$FNR = FN/(TP + FN) = 1 - TPR$$

false discovery rate (FDR)

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

---

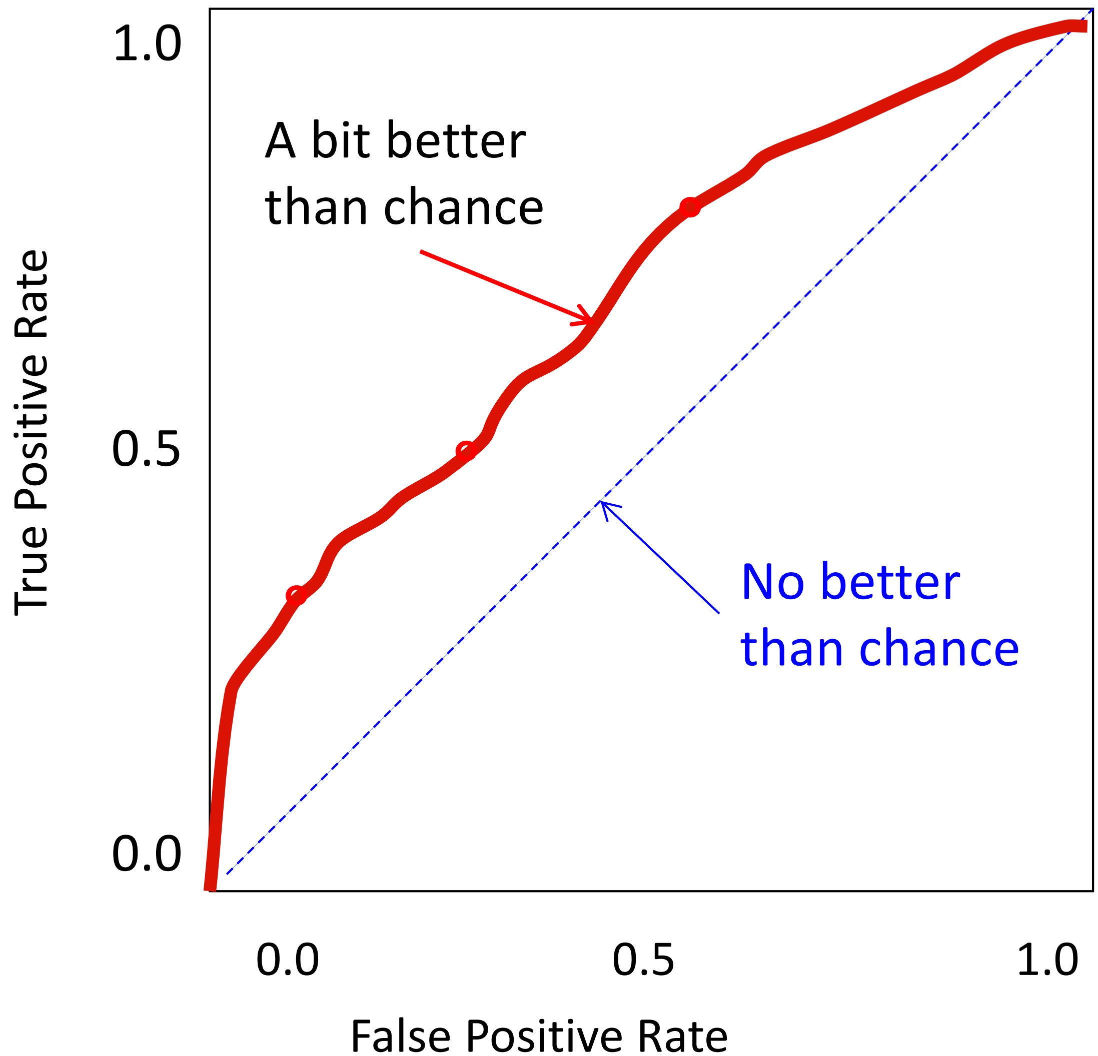
accuracy (ACC)

$$ACC = (TP + TN)/(TP + FP + FN + TN)$$

https://en.wikipedia.org/wiki/Sensitivity_and_specificity

ROC Curves

One View of a 2-parameter trade-off (true/false positives)



TPR = True Pos. Rate
= Sensitivity
= Recall
= TP/(TP+FN)

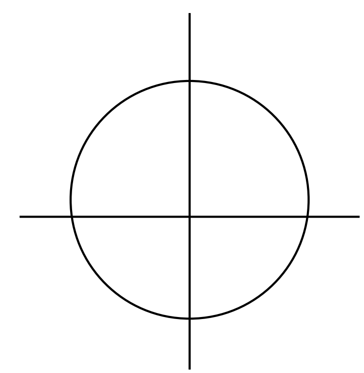
FPR = False Pos. Rate
= 1 – Specificity
= FP/(FP+TN)

Precision, aka PPV
= TP/(TP+FP)

Parametric Curves

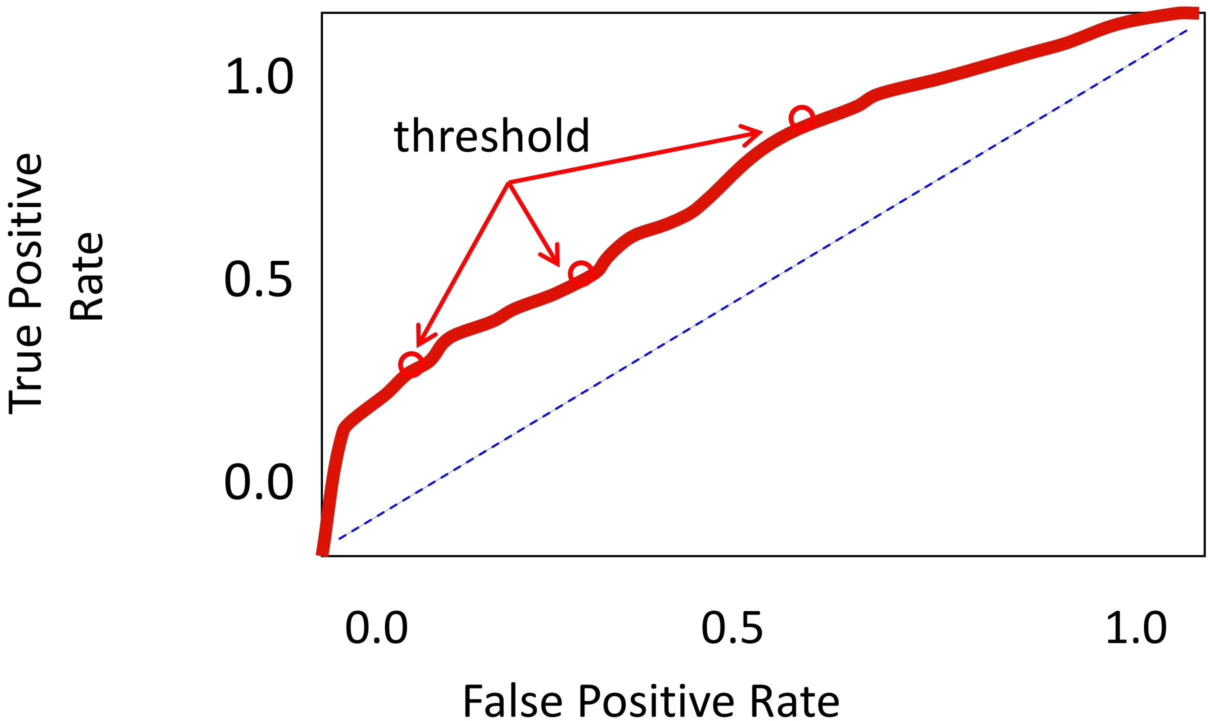
- Cartesian circle:
  - $x^2 + y^2 = 1$ (and good luck finding (x,y) pairs!)

- Parametric circle:
  - $x = \cos t$, $y = \sin t$, $0 \leq t \leq 2\pi$
  - t is "hidden" parameter in plot



ROC Curves are *Parametric*

- There is a "hidden" (in plot) threshold parameter defining TPR/FPR



NB: lowering thresh can*not* give *fewer* positives, hence neither TPR nor FPR lowered, so ROC curve *MUST* be non-decreasing!

How to plot a ROC curve

- Sort all instances by decreasing “score”
- Label each as to “actual” +/- status
- Calculate running totals of # of +/- above each row of table
- For a given score threshold, τ, find last row with score $\geq \tau$; all items at or above that are “predicted positives” if you use threshold τ, so that row index & corresponding totals give TPR & FPR, hence a point on the ROC curve.
- Things change *AT* τ = some row’s score, not *between* rows, so plotting those points gives complete ROC “curve”, and precise AUC (Area Under Curve).