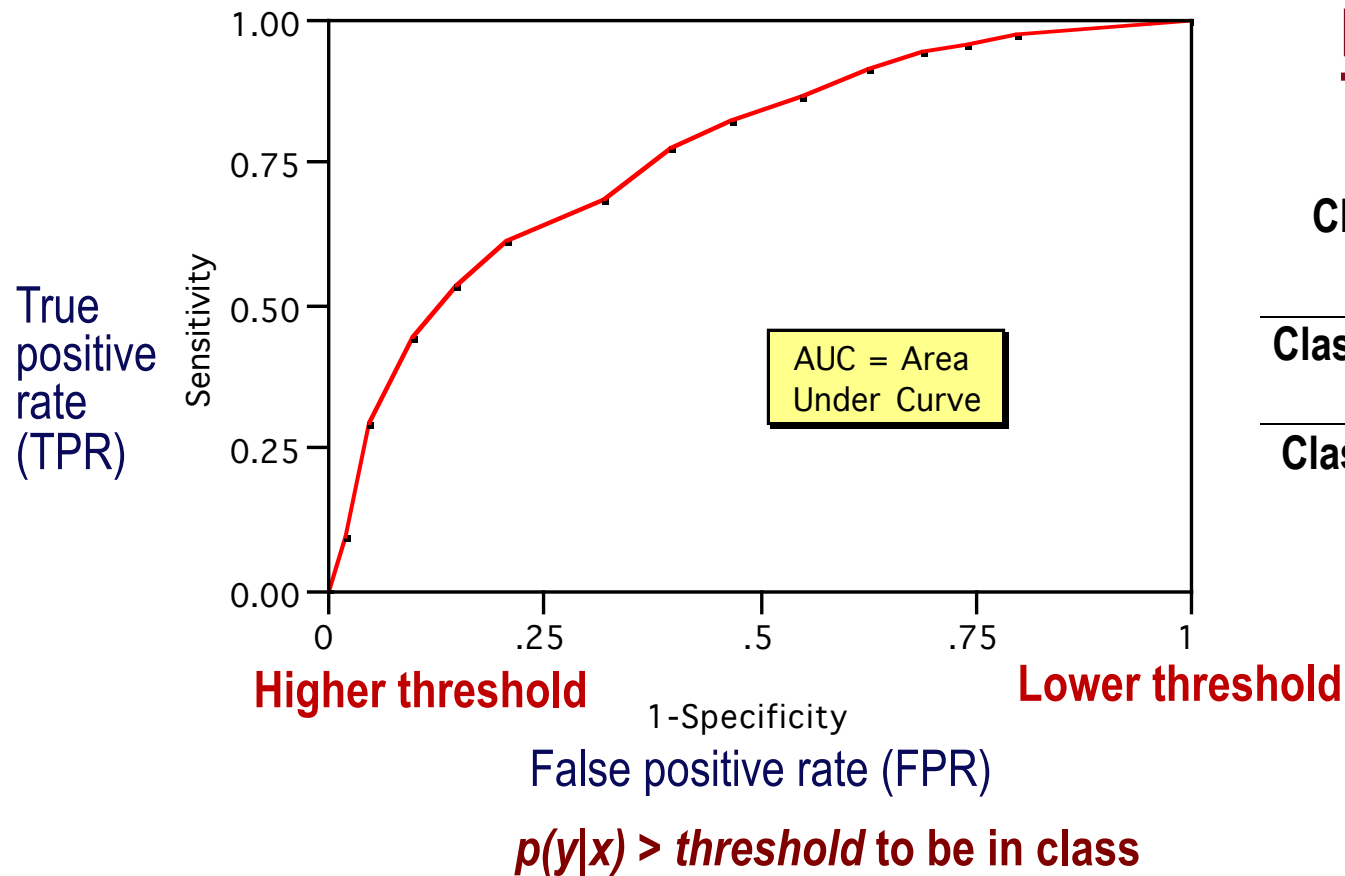


Recitation

Lyle Ungar

ROC Chart Varies Threshold



$$\text{FPR} = \text{FP} / (\text{FP} + \text{TN})$$

$$\text{TPR} = \text{TP} / (\text{FN} + \text{TP})$$

Claim\Is	True Yes	True No
Classify Yes	TP	FP
Classify No	FN	TN

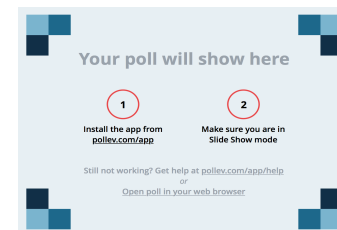
Is F1 symmetric?

- **Precision**
 - $P(\text{yes} \mid \text{predicted as yes}) = TP / (TP + FP)$
- **Recall (or Sensitivity)**
 - $P(\text{predicted as yes} \mid \text{yes}) = TP / (TP + FN)$
- **Specificity**
 - $P(\text{predicted as no} \mid \text{no}) = TN / (TN + FP)$
- $F1 = 2 \text{ precision} * \text{recall} / (\text{precision} + \text{recall})$

Claim/Is	True Yes	True No
Classify Yes	TP	FP
Classify No	FN	TN

If F1 symmetric?

- A) Yes**
- B) No**



Imputation

- ◆ Single column vs full matrix

What method would you use?

- ◆ A doctor has 1,000 patients who had heart surgery; outcome is whether or not they were readmitted to the hospital within 30 days
- ◆ An ophthalmologist has 10,000 images of eyes labeled as “retinopathy” or “normal
- ◆ I have the Facebook posts of 3,000 people and their scores on a questionnaire to assess excess drinking