ROC curves
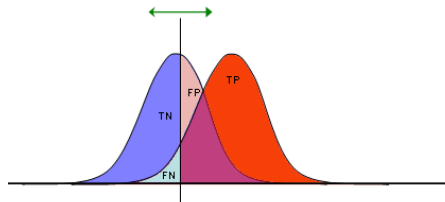
Jeffrey Leek

May 18, 2016
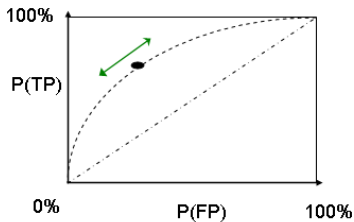
Why a curve?

- In binary classification you are predicting one of two categories
- Alive/dead
- Click on ad/don't click
- But your predictions are often quantitative
- Probability of being alive
- Prediction on a scale from 1 to 10
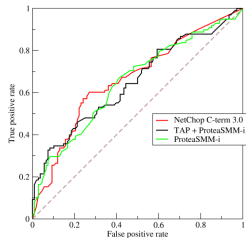- The *cutoff* you choose gives different results

ROC curves

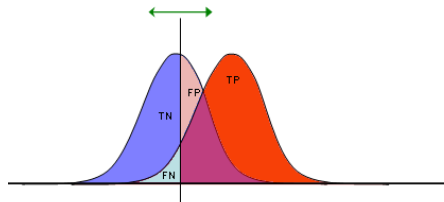| TP	FP
FN	TN
1	1

http://en.wikipedia.org/wiki/Receiver_operating_
characteristic

An example

http://en.wikipedia.org/wiki/Receiver_operating_
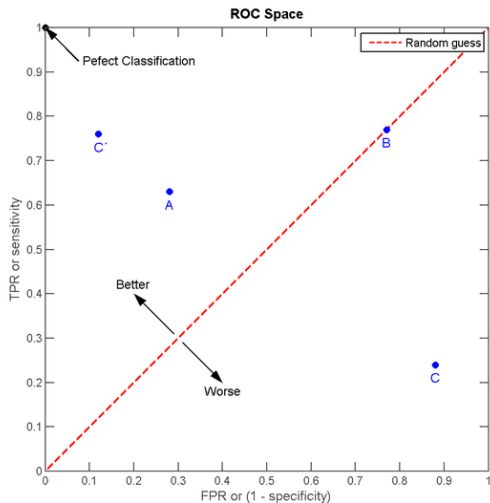characteristic

Area under the curve

TP	FP
FN	TN
1	1
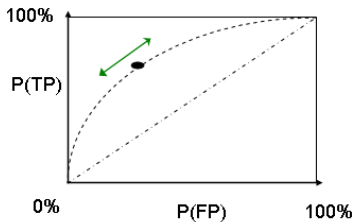
- AUC = 0.5: random guessing
- AUC = 1: perfect classifer

What is good?