CS 66: Machine Learning

Prof. Sara Mathieson

Spring 2019

SWARTHMORE

Outline for March 1

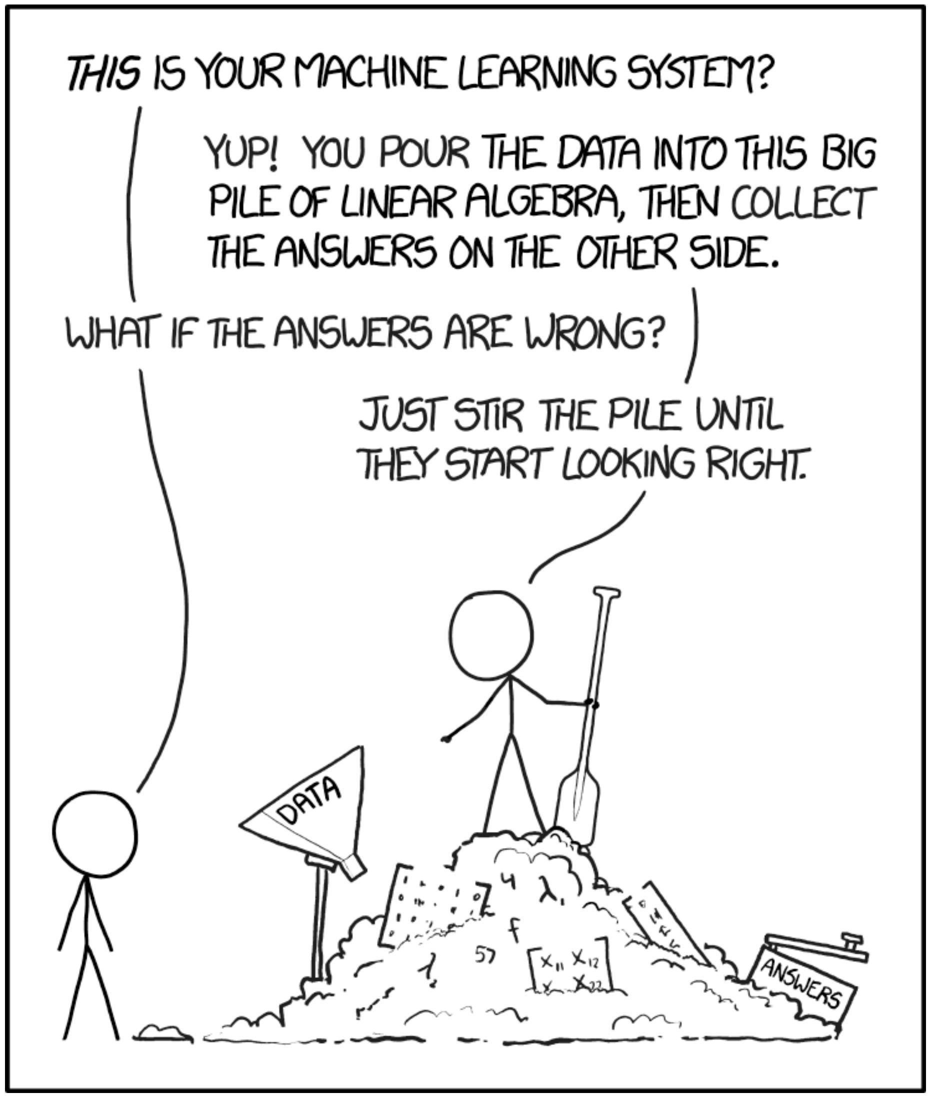
- Evaluation metrics
  - Precision and Recall
  - Confusion matrices revisited
  - ROC curves
  - Relationship to probabilistic methods
  - Cross-validation

- **Lab 4 due March 8** (week from today)
- Office hours **TODAY** 1-3pm

Outline for March 1

- Evaluation metrics
  - Precision and Recall
  - Confusion matrices revisited
  - ROC curves
  - Relationship to probabilistic methods
  - Cross-validation

Thanks for all your work on the exam!



xkcd: "Machine Learning"

Goals of Evaluation

- Think about what metrics are important for the problem at hand
- Compare different methods on the same problem
- Common set of tools that other researchers/users can understand

Precision and Recall

- Precision: of all the "flagged" examples, which ones are actually relevant (i.e. positive)?

- Recall: of all the relevant results, which ones did I actually return?

Precision and Recall

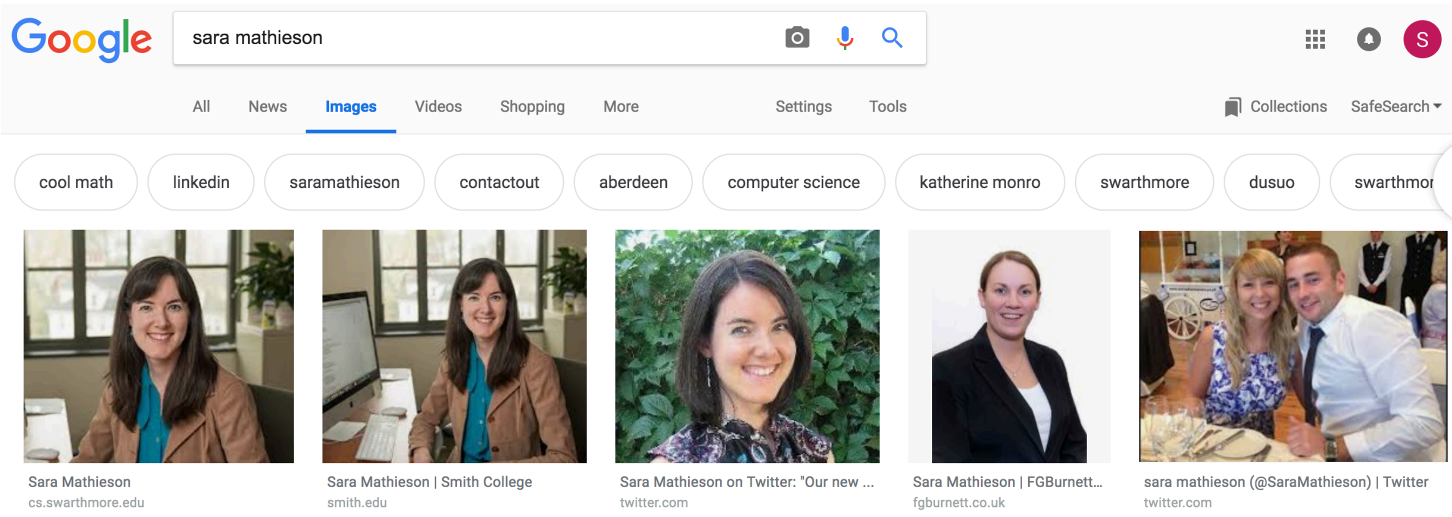
- <u>Precision</u>: of all the “flagged” examples, which ones are actually relevant (i.e. positive)?

(Purity)

- <u>Recall</u>: of all the relevant results, which ones did I actually return?

(Completeness)

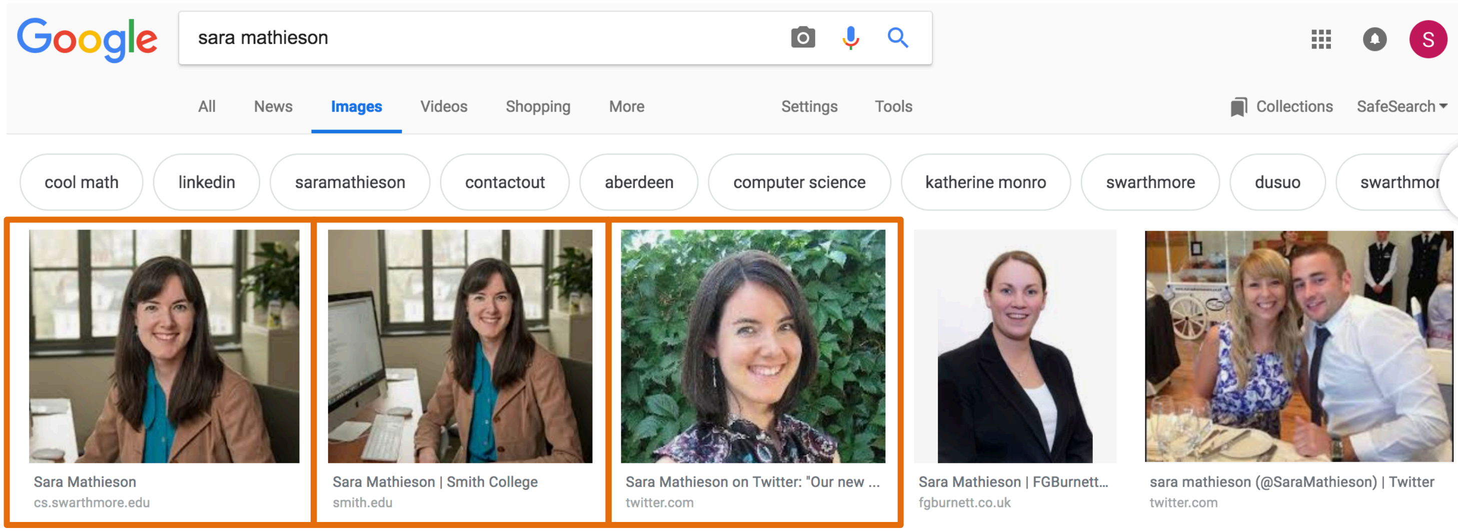
Precision and Recall



P=6 (number of images that are actually me)

- Precision?
- Recall?

Precision and Recall

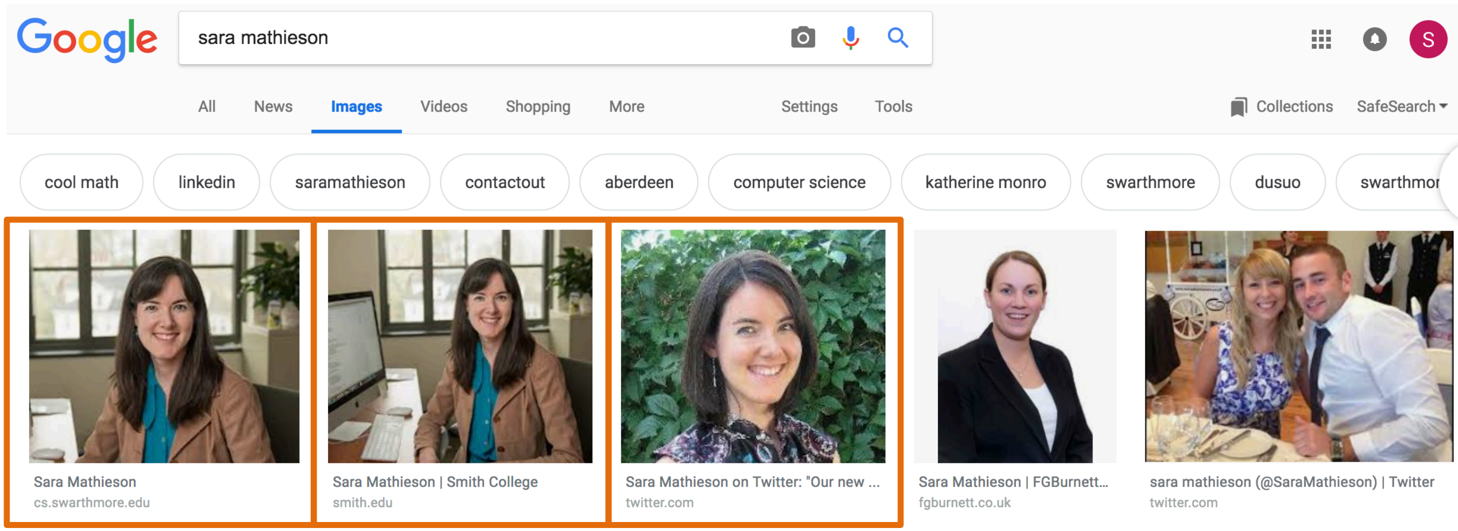


P=6 (number of images that are actually me)

- Precision = TP/(FP+TP) = 3/5
- Recall?

Modified from Ameet Soni

Precision and Recall

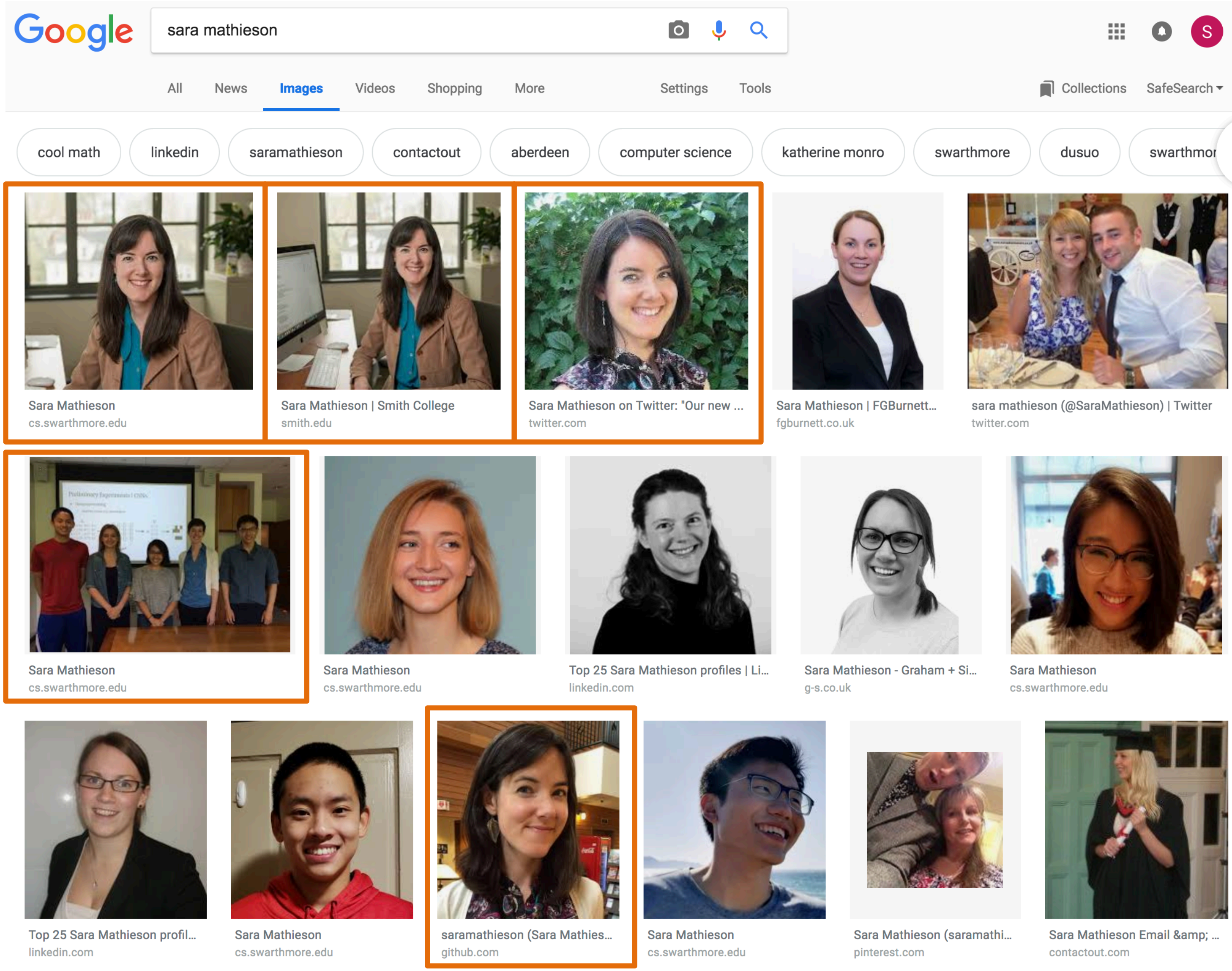


P=6 (number of images that are actually me)

- Precision = TP/(FP+TP) = 3/5
- Recall = TP/(FN+TP) = 3/6

Modified from Ameet Soni

Precision and Recall



P=6 (number of images that are actually me)

- Precision = 5/16
- Recall = 5/6

Modified from Ameet Soni

Outline for March 1

- Evaluation metrics
  - Precision and Recall
  - Confusion matrices revisited
  - ROC curves
  - Relationship to probabilistic methods
  - Cross-validation

Recap Confusion Matrices

		Predicted class	
		Negative	Positive
True class	Negative	True negative (TN)	False positive (FP)
	Positive	False negative (FN)	True positive (TP)

Recap Confusion Matrices

		Predicted class: Negative	Predicted class: Positive	
True class	Negative	True negative (TN)	False positive (FP) "false alarm"	N (total number of true negatives)
	Positive	False negative (FN) "miss"	True positive (TP)	P (total number of true positives)
		N* (what we said was negative)	P* (what we said was positive "flagged")	

Modified from Jessica Wu

Recap Confusion Matrices

Predicted class

True class	Negative	Positive	
Negative	True negative (TN)	False positive (FP) “false alarm”	N
Positive	False negative (FN) “miss”	True positive (TP)	P
	N*	P*	

Modified from Jessica Wu

Recap Confusion Matrices

		Predicted class		
		Negative	Positive	
True class	Negative	True negative (TN)	False positive (FP) "false alarm"	N
	Positive	False negative (FN) "miss"	True positive (TP)	P
		N*	P*	

Error:

(FN+FP)/(TN+FP+FN+TP)

= (FN+FP)/(N+P)

Modified from Jessica Wu

Recap Confusion Matrices

True class		Predicted class: Negative	Predicted class: Positive	
	Negative	True negative (TN)	False positive (FP) "false alarm"	N
	Positive	False negative (FN) "miss"	True positive (TP)	P
		N*	P*	

Accuracy = 1-Error:

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

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

Recap Confusion Matrices

		Predicted class		
		Negative	Positive	
True class	Negative	True negative (TN)	False positive (FP) “false alarm”	N
	Positive	False negative (FN) “miss”	True positive (TP)	P
		N*	P*	

Precision:

TP/(FP+TP) = TP/P*

Modified from Jessica Wu

Recap Confusion Matrices

Predicted class

True class

	Negative	Positive	
Negative	True negative (TN)	False positive (FP) “false alarm”	N
Positive	False negative (FN) “miss”	True positive (TP)	P
	N*	P*	

Recall
(True Positive Rate):

TP/(FN+TP) = TP/P

Modified from Jessica Wu

Recap Confusion Matrices

True class	Predicted class: Negative	Predicted class: Positive	
Negative	True negative (TN)	False positive (FP) "false alarm"	N
Positive	False negative (FN) "miss"	True positive (TP)	P
	N*	P*	

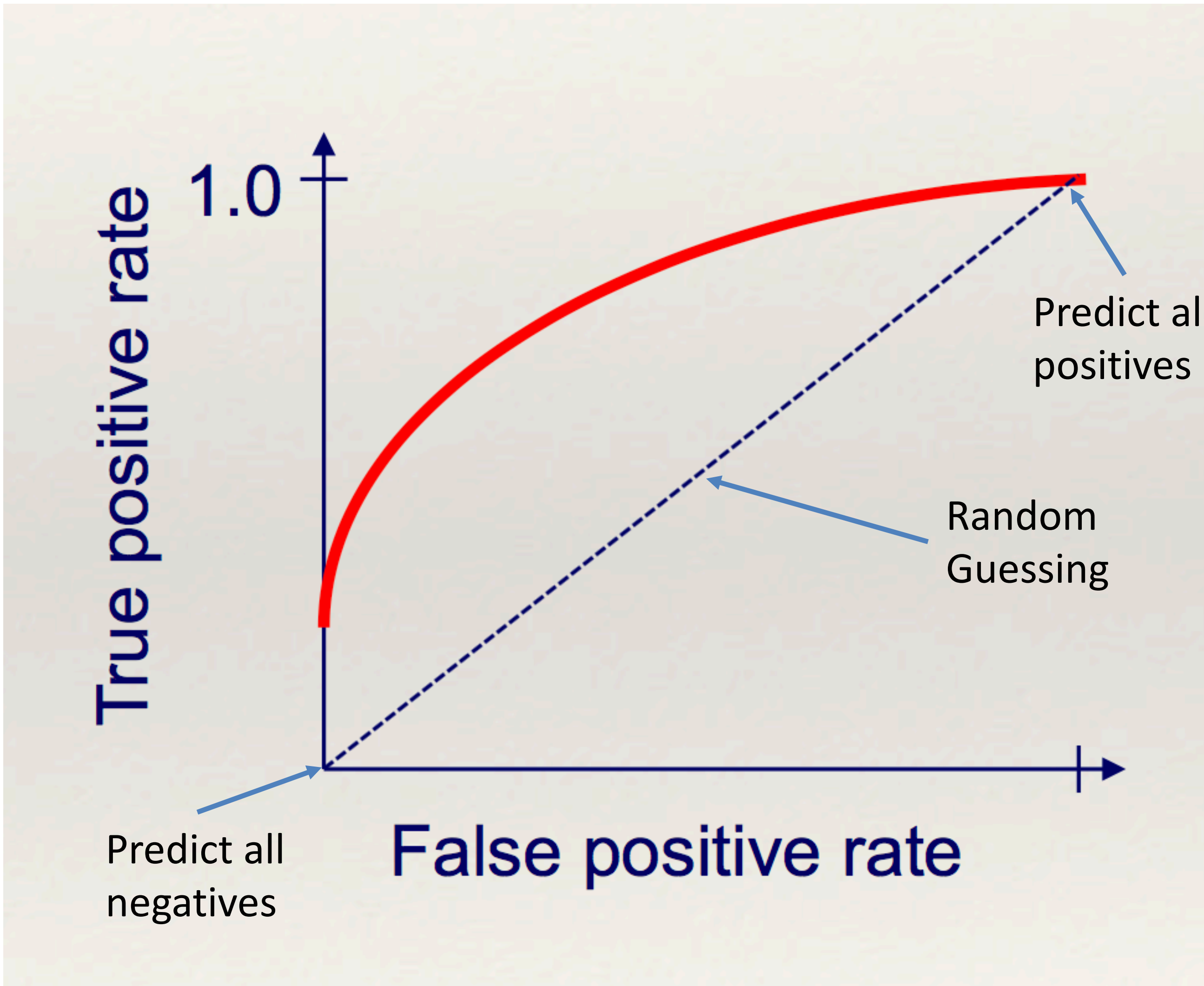
False Positive Rate:

FP/(TN+FP) = FP/N

Outline for March 1

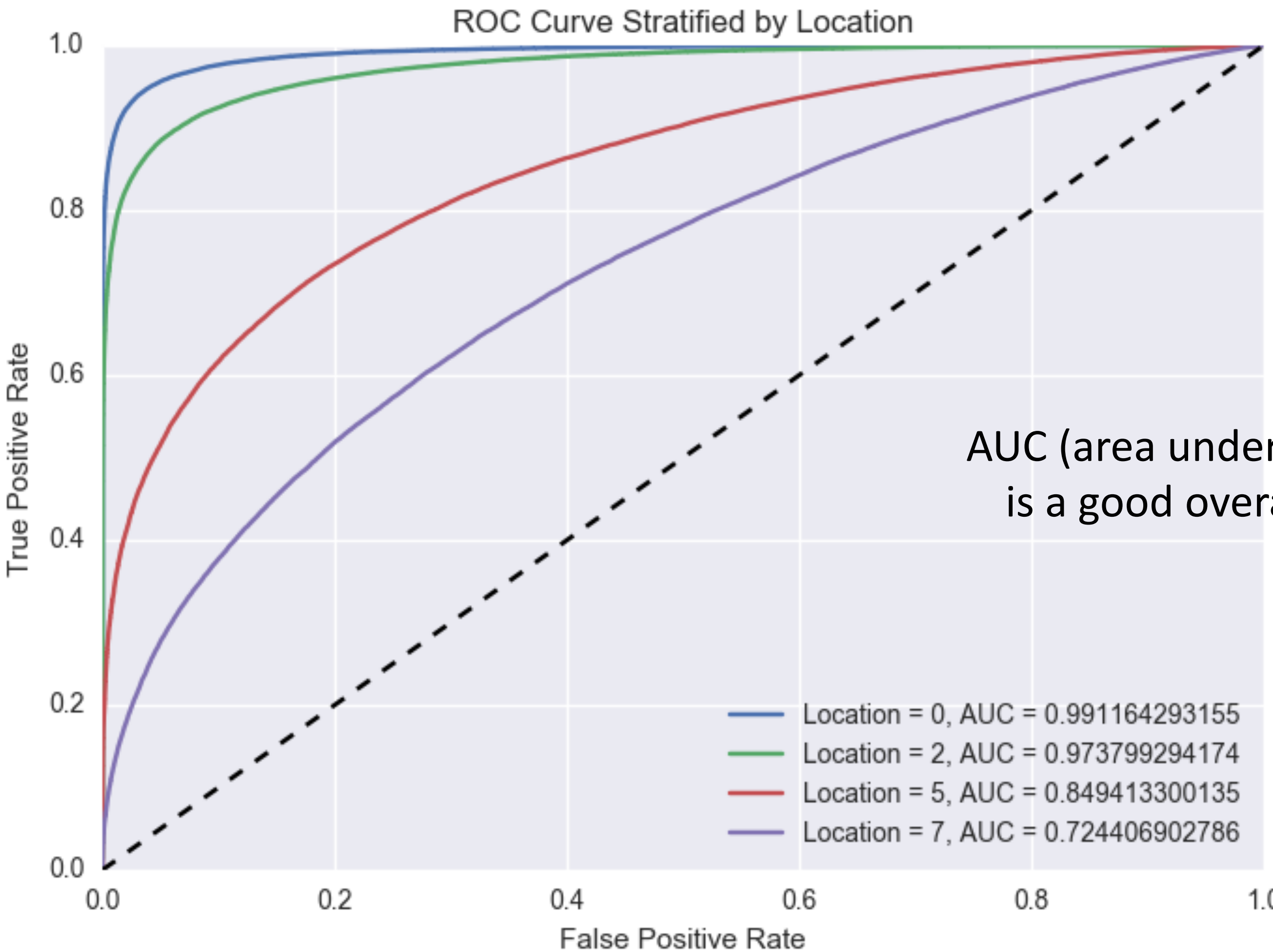
- Evaluation metrics
  - Precision and Recall
  - Confusion matrices revisited
  - ROC curves
  - Relationship to probabilistic methods
  - Cross-validation

ROC curve (Receiver Operating Characteristic)



Modified from Ameet Soni

ROC curve example: comparing methods



AUC (area under the curve) is a good overall metric

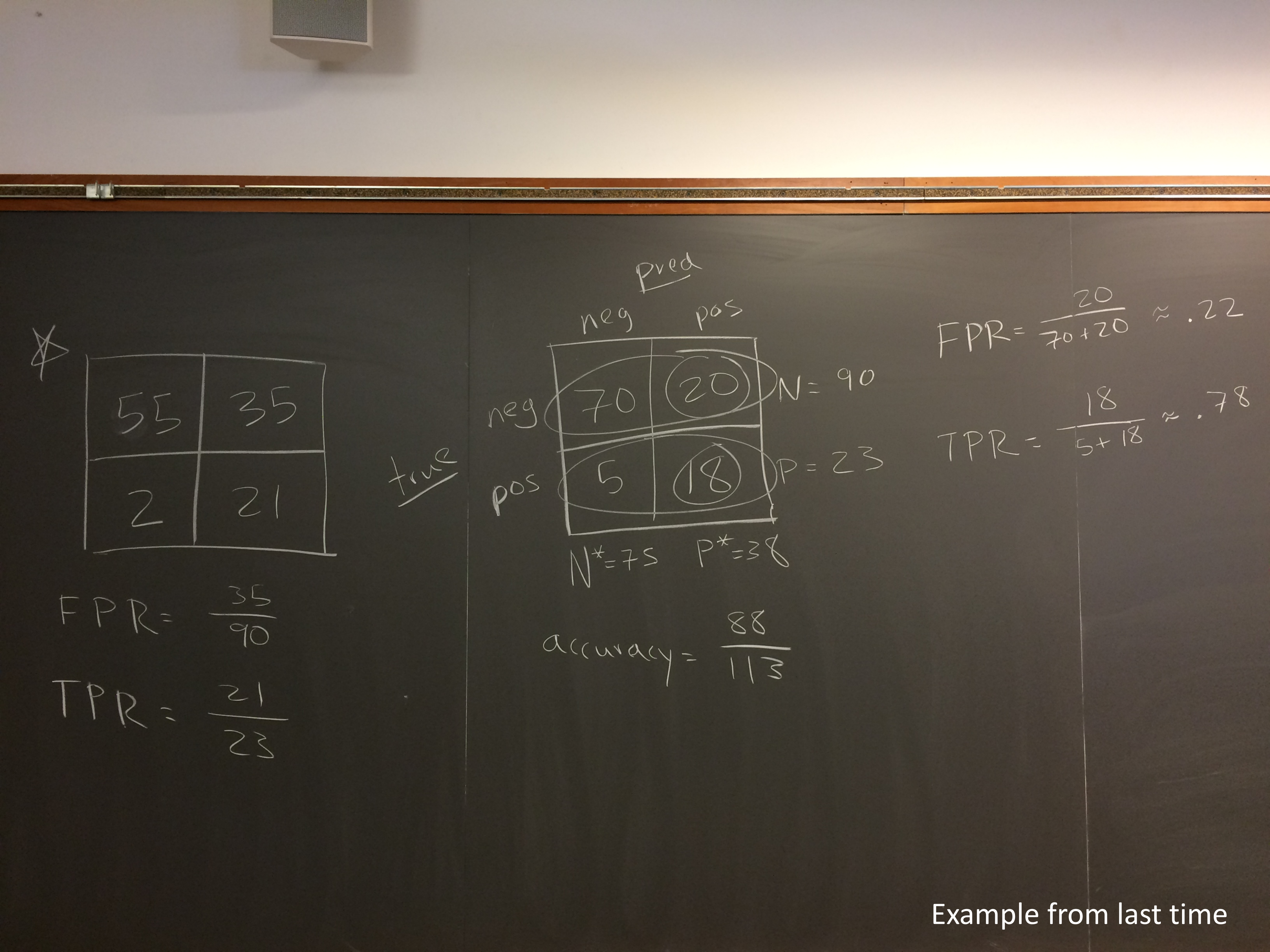
Example of a ROC curve from my research
Chan, Perrone, Spence, Jenkins, Mathieson, Song

How to get a ROC curve for probabilistic methods?

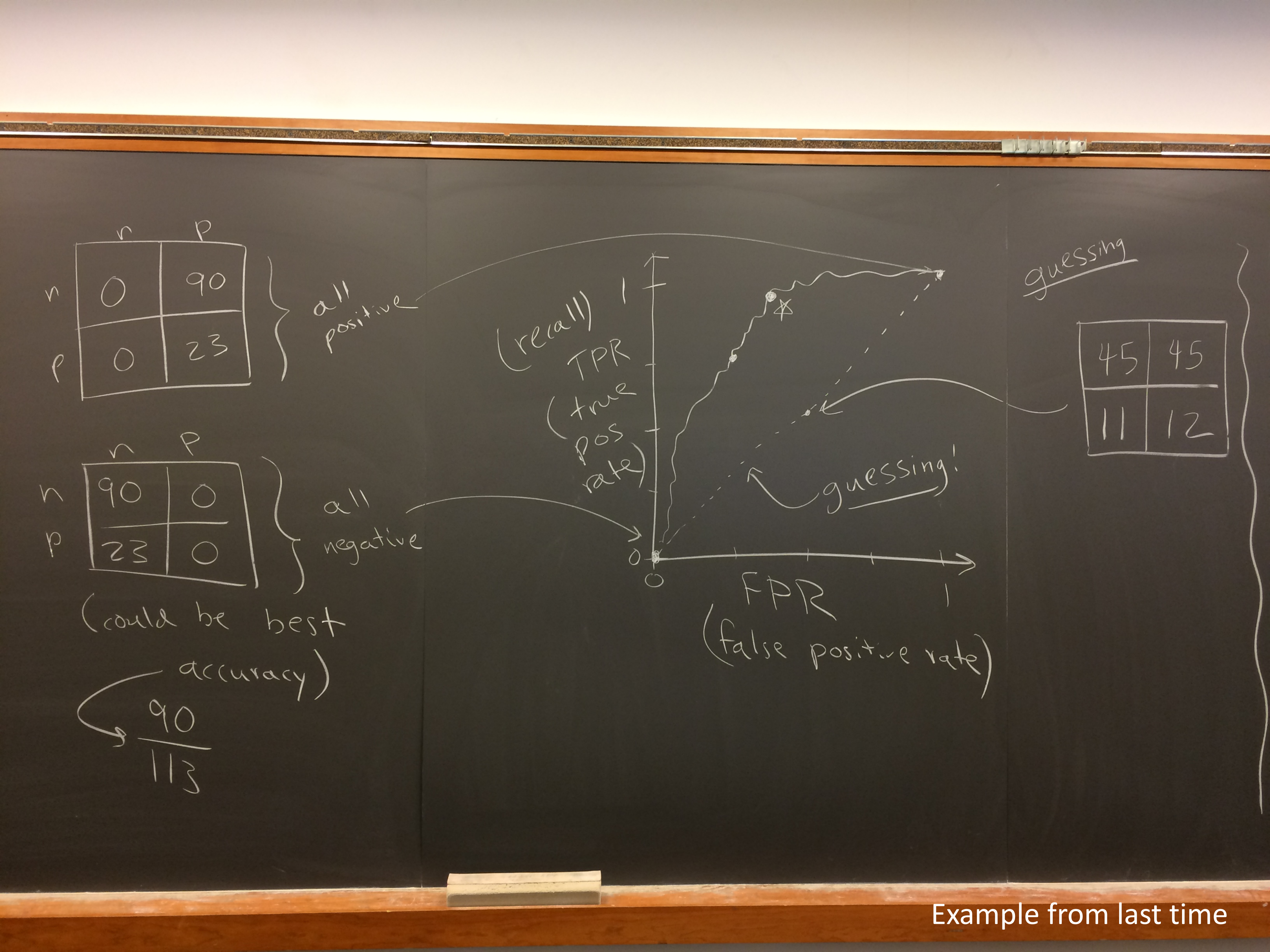
- Usually we use 0.5 as a threshold for binary classification

- Vary the threshold! (i.e. choose 0.25)

  - $P(y=1 \mid x) > 0.25$ => classify as 1 (positive)
  - $P(y=1 \mid x) <= 0.25$ => classify as 0 (negative)



Example from last time



Example from last time

Handout 7

(a)

true	n	p	
n	77	3	80
p	13	7	20
	90	10	

precision: $\frac{7}{10} = .7$

recall: $\frac{7}{20} = .35$

FPR: $\frac{3}{77+3} = .04$

(d)

68	12	80
2	18	20
70	30	

precision: $\frac{18}{30}$

FPR $= \frac{12}{80} = .15$

TPR $= \frac{18}{20} = .9$