

Evaluation Measures for Classifiers

Cynthia Rudin

Machine Learning Course, Duke

Classification

Example 1:	[5	3	120	12	1	0]
Example 2:	[0	0	89	5	1	1]
Example 3:	[1	0	20	0	0	1]
:								:	



Features, called x

Classification

Example 1:	[5	3	120	12	1	0]	-1
Example 2:	[0	0	89	5	1	1]	1
Example 3:	[1	0	20	0	0	1]	-1
:							:			:



Features, called x



Labels, y

Classification

Example 1:	[5	3	120	12	1	0]	-1	-22.1
Example 2:	[0	0	89	5	1	1]	1	17.2
Example 3:	[1	0	20	0	0	1]	-1	5.2
:							:			:	:

↑
Features, called x

↑
Labels, y

↑
 $f(x)$

Classification

Example 1:	[5	3	120	12	1	0]	-1	-22.1	-1
Example 2:	[0	0	89	5	1	1]	1	17.2	1
Example 3:	[1	0	20	0	0	1]	-1	5.2	1
:							:			:		:

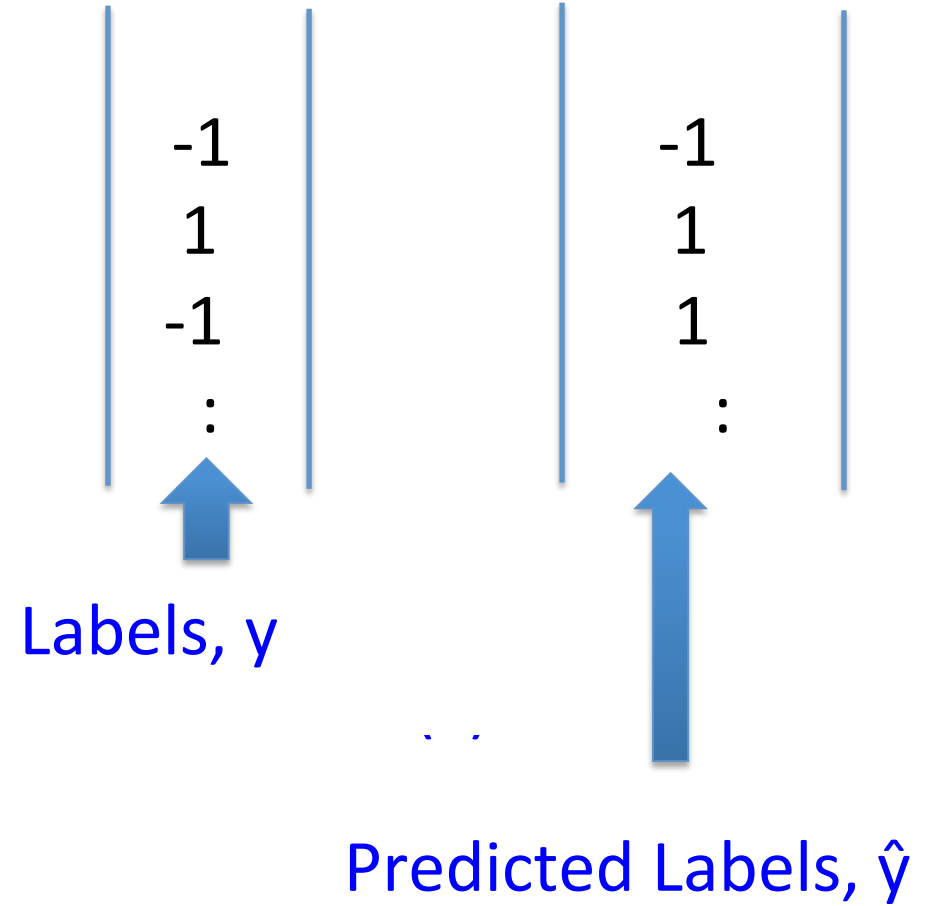
↑
Features, called x

↑
Labels, y

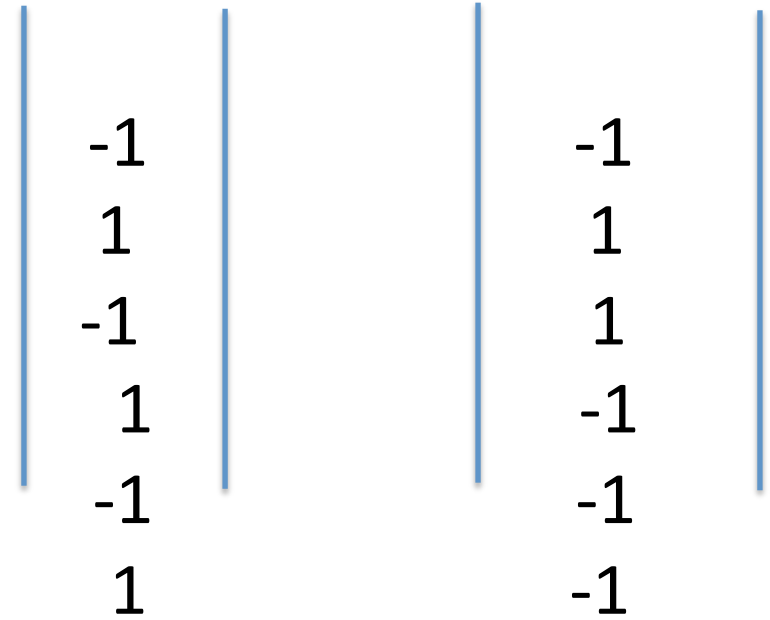
↑
 $f(x)$

↑
Predicted Labels, \hat{y}

Classification



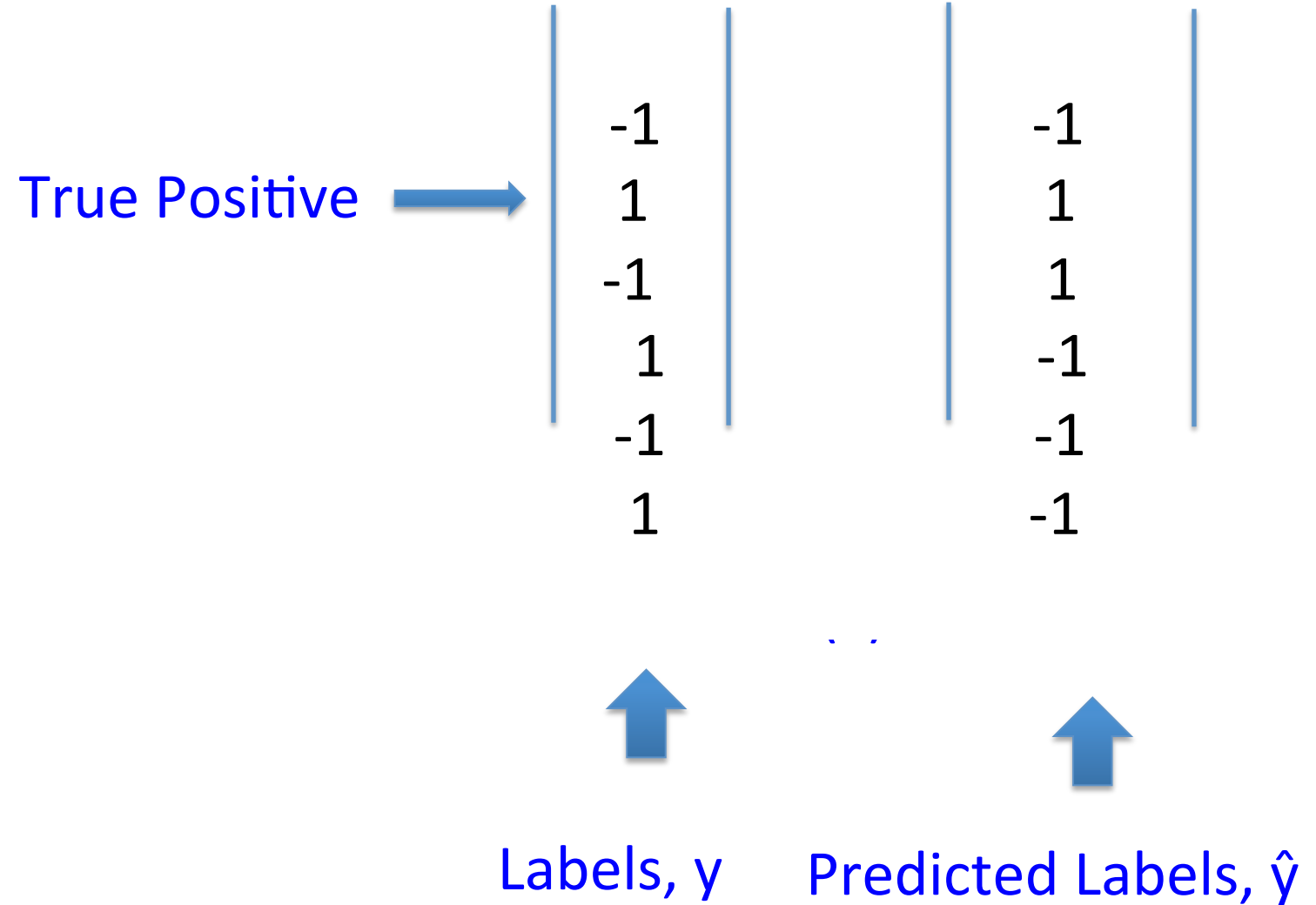
Classification



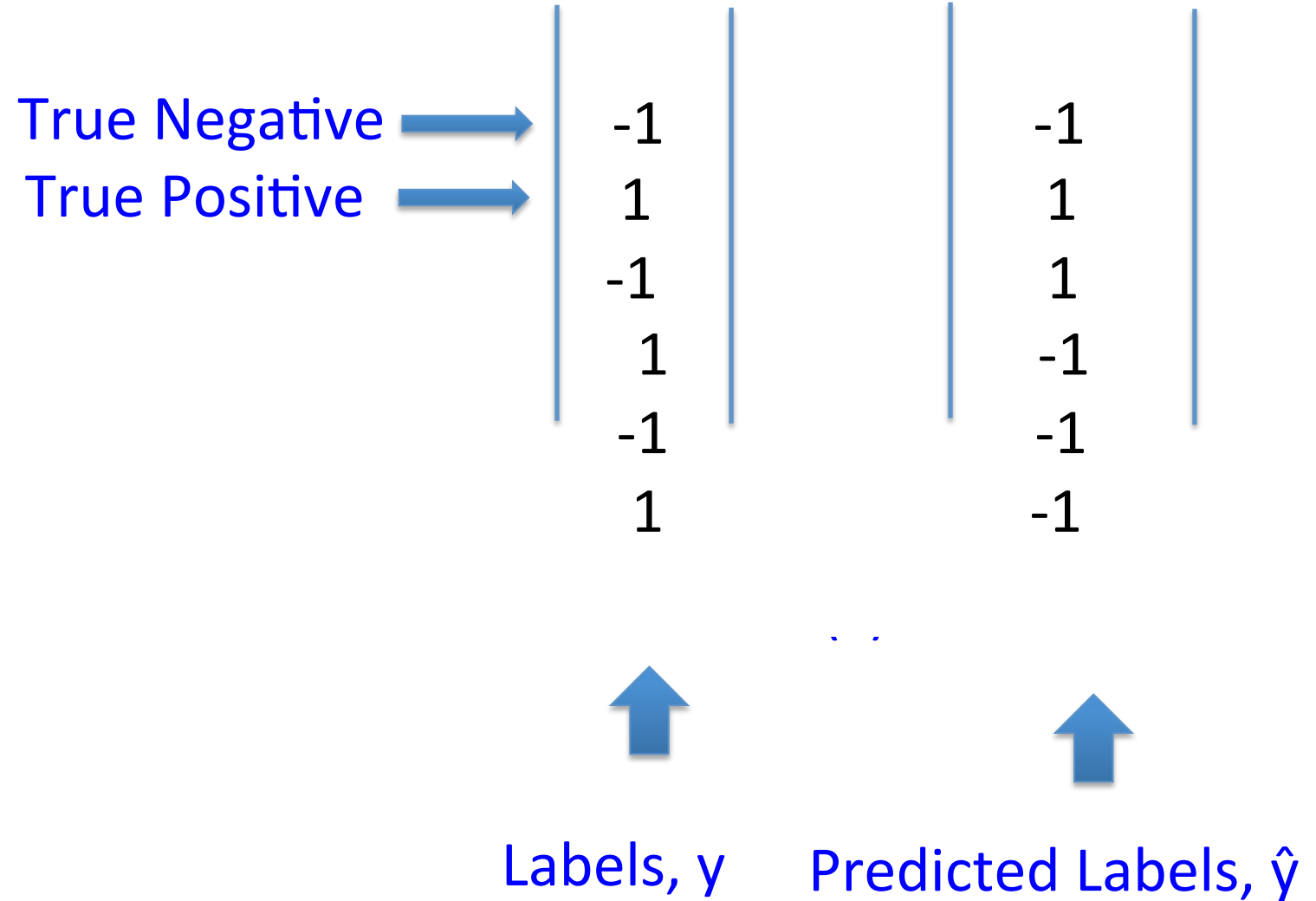
Labels, y

Predicted Labels, \hat{y}

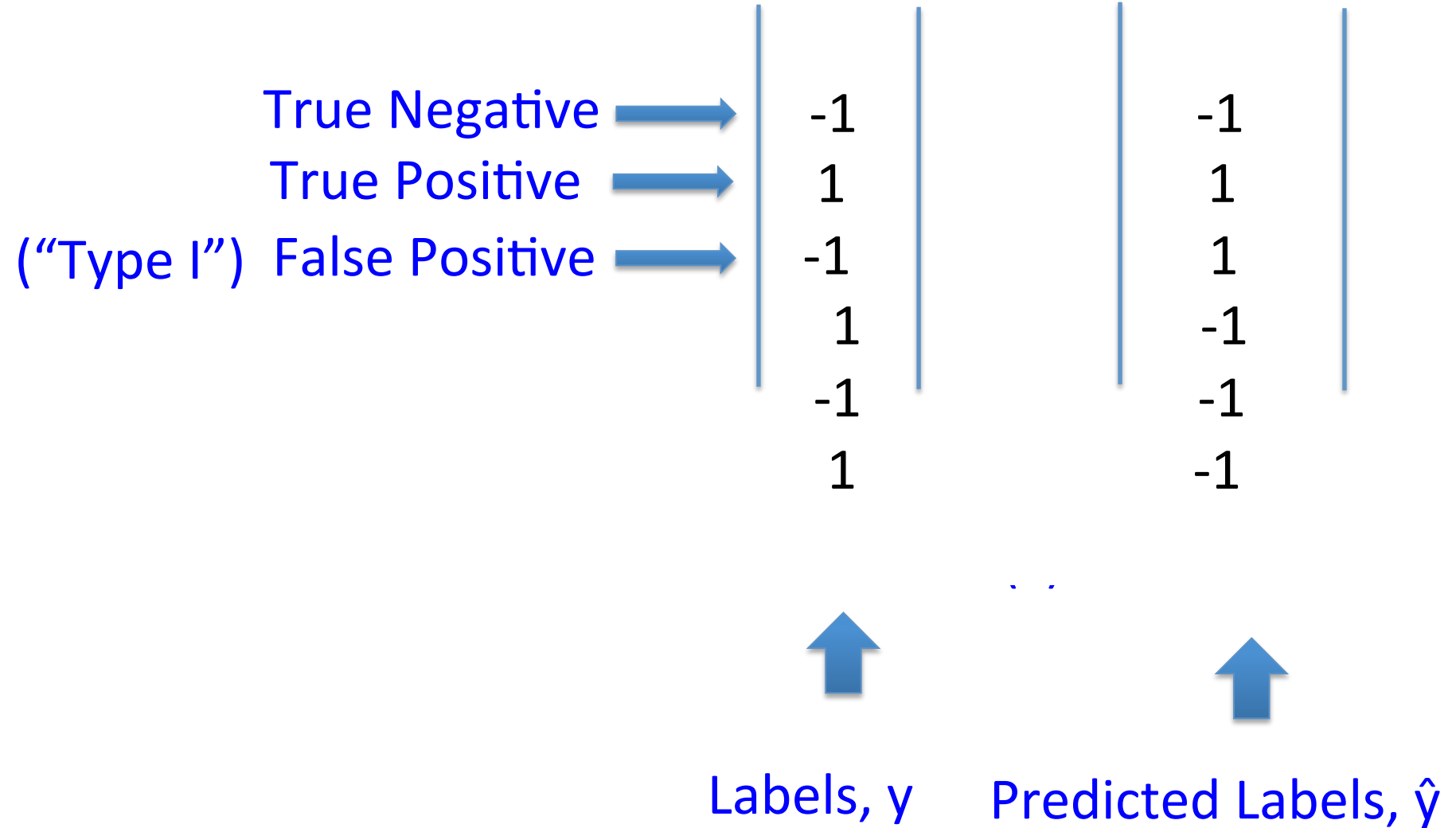
Classification



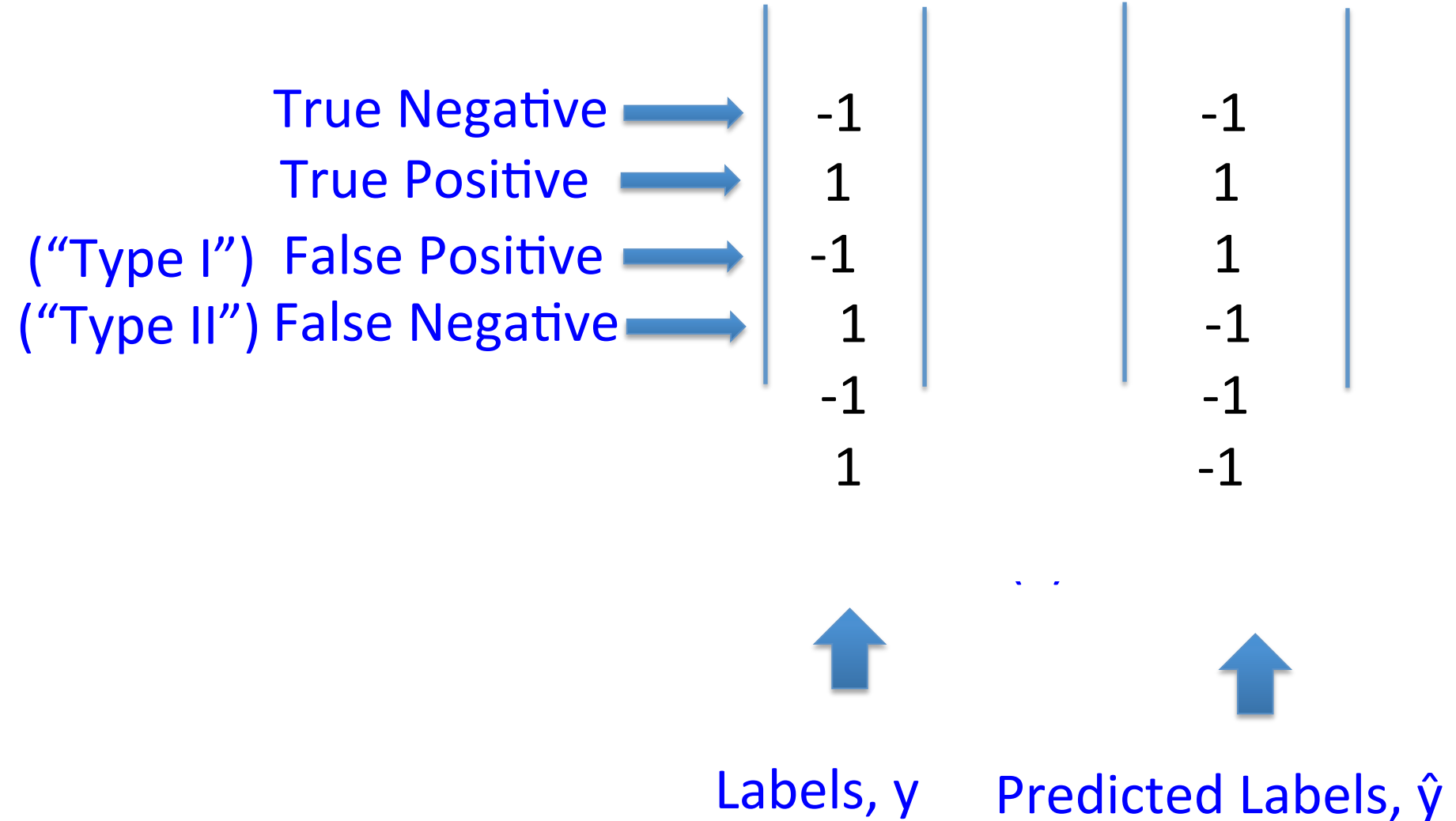
Classification



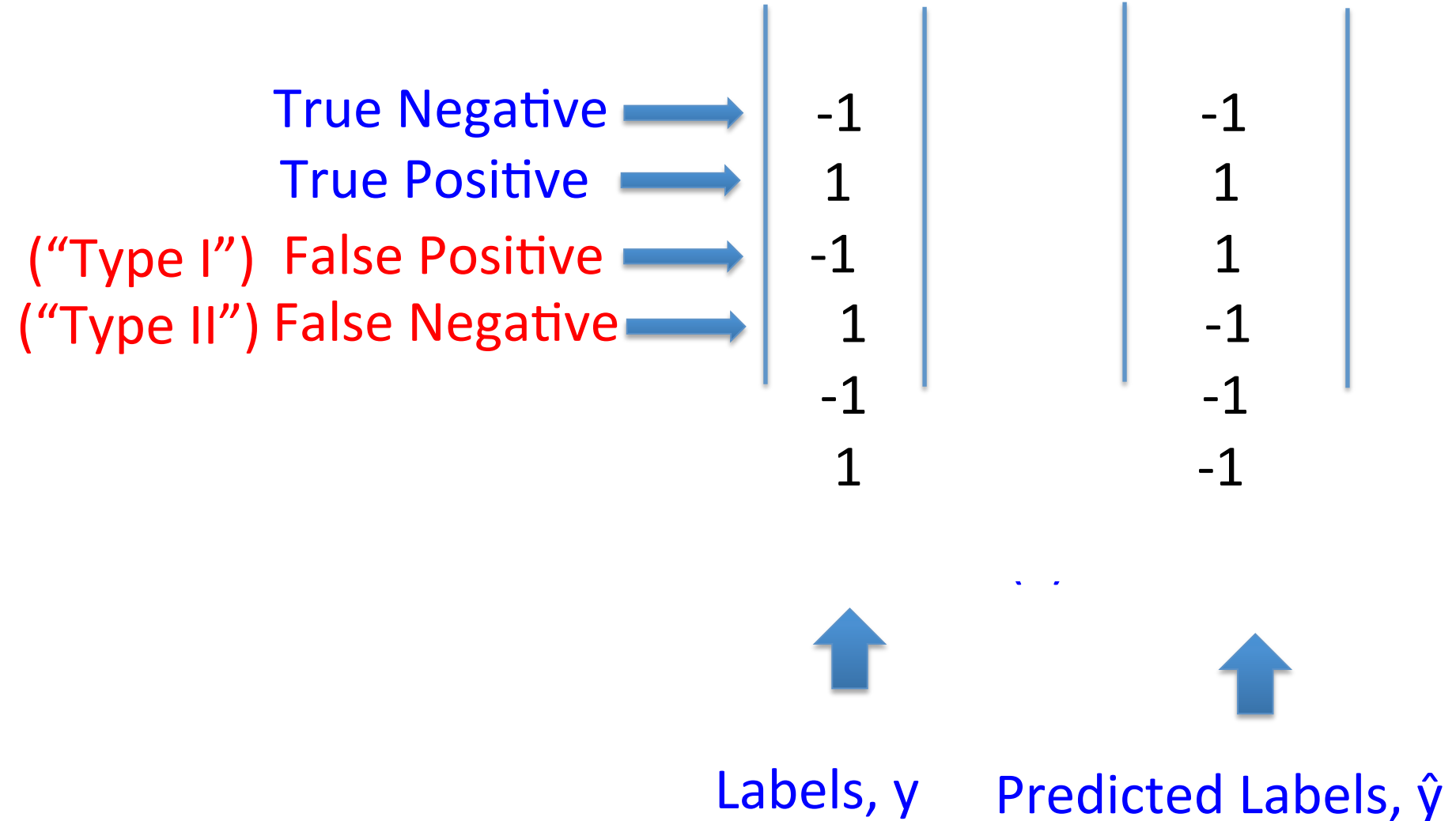
Classification



Classification



Classification



Confusion Matrix

	$y=+1$	$y=-1$
$\hat{y}=1$	723	15
$\hat{y}=-1$	72	409

Confusion Matrix

	$y=+1$	$y=-1$
$\hat{y}=1$	723	15
$\hat{y}=-1$	72	409

	$y=+1$	$y=-1$
$\hat{y}=1$	TP	FP (Type I)
$\hat{y}=-1$	FN (Type II)	TN

Evaluation Measures for Classifiers

- Misclassification error

$$\frac{\text{FP} + \text{FN}}{n}$$

Evaluation Measures for Classifiers

- Misclassification error

$$\frac{\text{FP}+\text{FN}}{n} = \frac{1}{n} \sum_{i=1}^n 1_{[y_i \neq \hat{y}_i]}$$

	$y=+1$	$y=-1$
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

Evaluation Measures for Classifiers

- True Positive Rate (TPR), Sensitivity, Recall

$$\frac{\text{TP}}{\# \text{Pos}} = \frac{\sum_i^n 1_{[y_i = \hat{y}_i \text{ and } y_i = 1]}}{\sum_i^n 1_{[y_i = 1]}}$$

	y=+1	y=-1
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

Evaluation Measures for Classifiers

- True Negative Rate (TNR), Specificity

$$\frac{\text{TN}}{\# \text{ Neg}} = \frac{\sum_i^n 1_{[y_i = \hat{y}_i \text{ and } y_i = -1]}}{\sum_i^n 1_{[y_i = -1]}}$$

	y=+1	y=-1
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

Evaluation Measures for Classifiers

- False Positive Rate (FPR)

$$\frac{\text{FP}}{\# \text{ Neg}} = \frac{\sum_i^n 1_{[y_i \neq \hat{y}_i \text{ and } y_i = -1]}}{\sum_i^n 1_{[y_i = -1]}}$$

	y=+1	y=-1
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

Evaluation Measures for Classifiers

- Precision

$$\frac{\text{TP}}{\# \text{ predicted positive}} = \frac{\sum_i^n 1_{[y_i = \hat{y}_i \text{ and } y_i = 1]}}{\sum_i^n 1_{[\hat{y}_i = 1]}}$$

	y=+1	y=-1
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

Evaluation Measures for Classifiers

- F1-score

$$F1 = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Precision

	y=+1	y=-1
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

Recall

	y=+1	y=-1
$\hat{y}=1$	TP	FP
$\hat{y}=-1$	FN	TN

ROC Curves, Part I

Cynthia Rudin

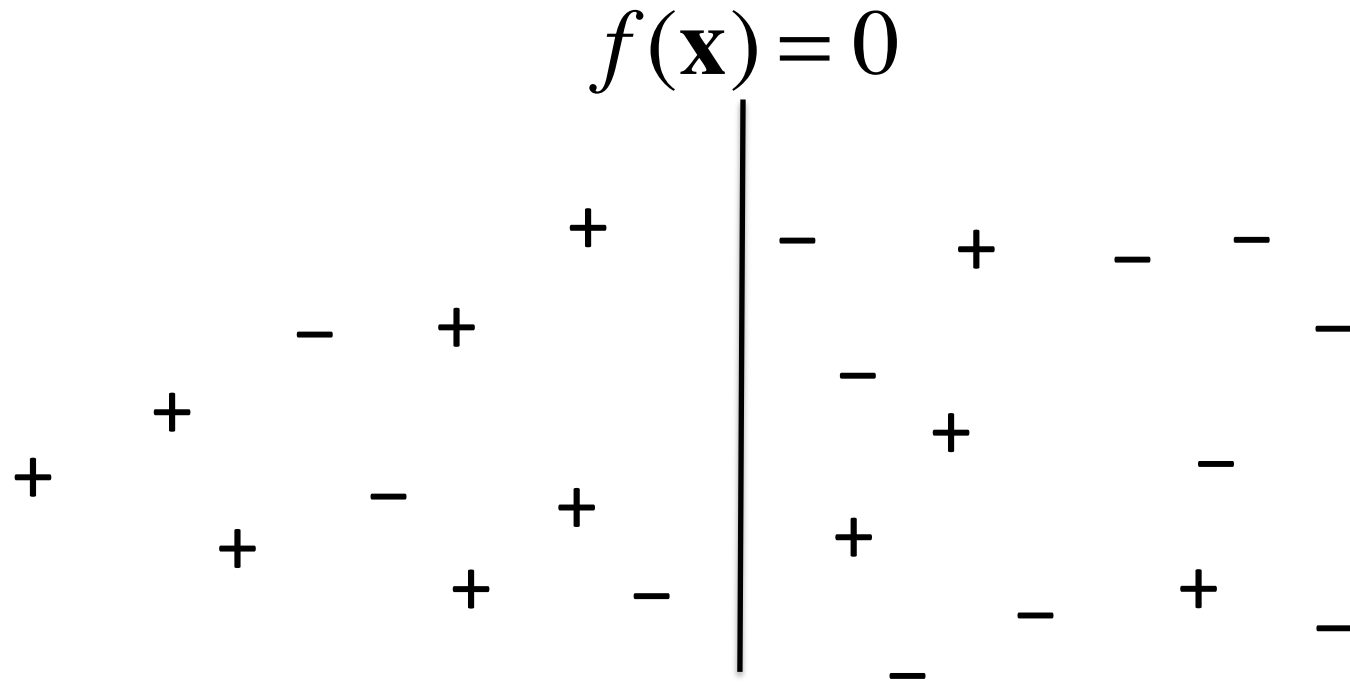
Machine Learning Course, Duke

ROC Curves

- Started during WWII for analyzing radar signals.
- For a particular False Positive Rate (FPR), what is the True Positive Rate (TPR)?
- $\text{FPR} = \text{number of negatives that were classified by the ML algorithm as positives} / \text{total number of negatives}$
- $\text{TPR} = \text{number of positives that were classified by the ML algorithm as positives} / \text{total number of positives}$

ROC Curves

- Adjust the decision boundary

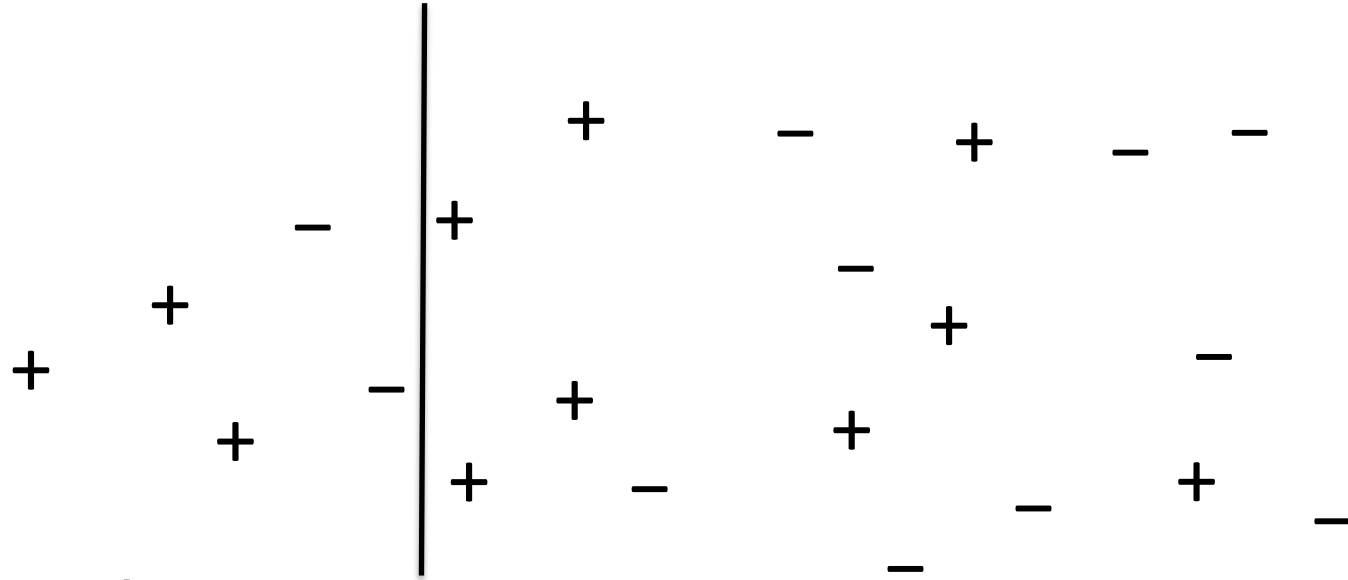


- TPR = 7/11
- FPR = 3/12

ROC Curves

- Adjust the decision boundary

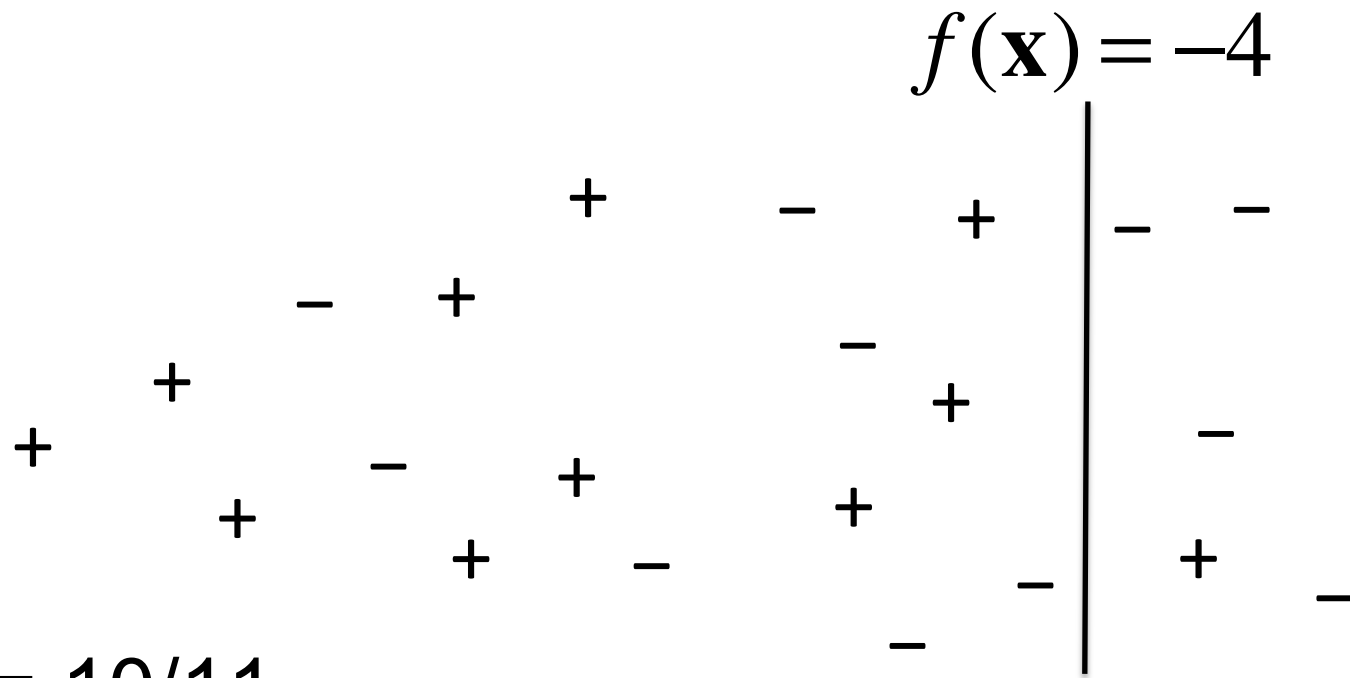
$$f(\mathbf{x}) = 3$$



- TPR = 3/11
- FPR = 2/12

ROC Curves

- Adjust the decision boundary

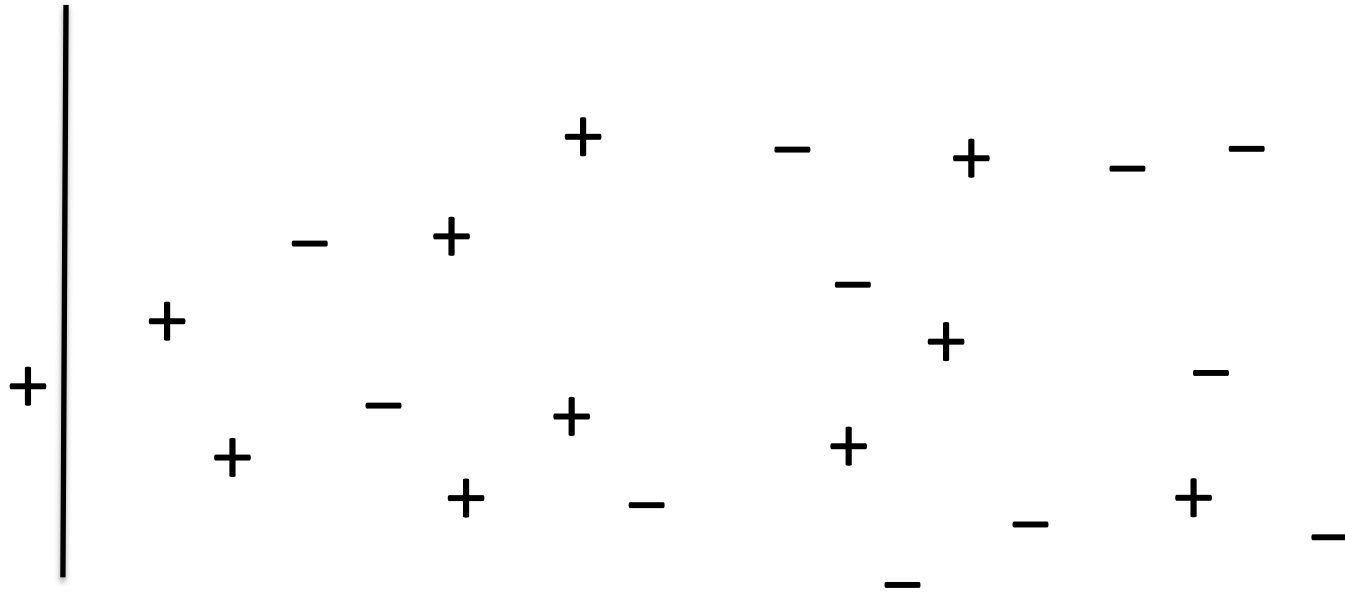


- TPR = 10/11
- FPR = 7/12

ROC Curves

- Adjust the decision boundary

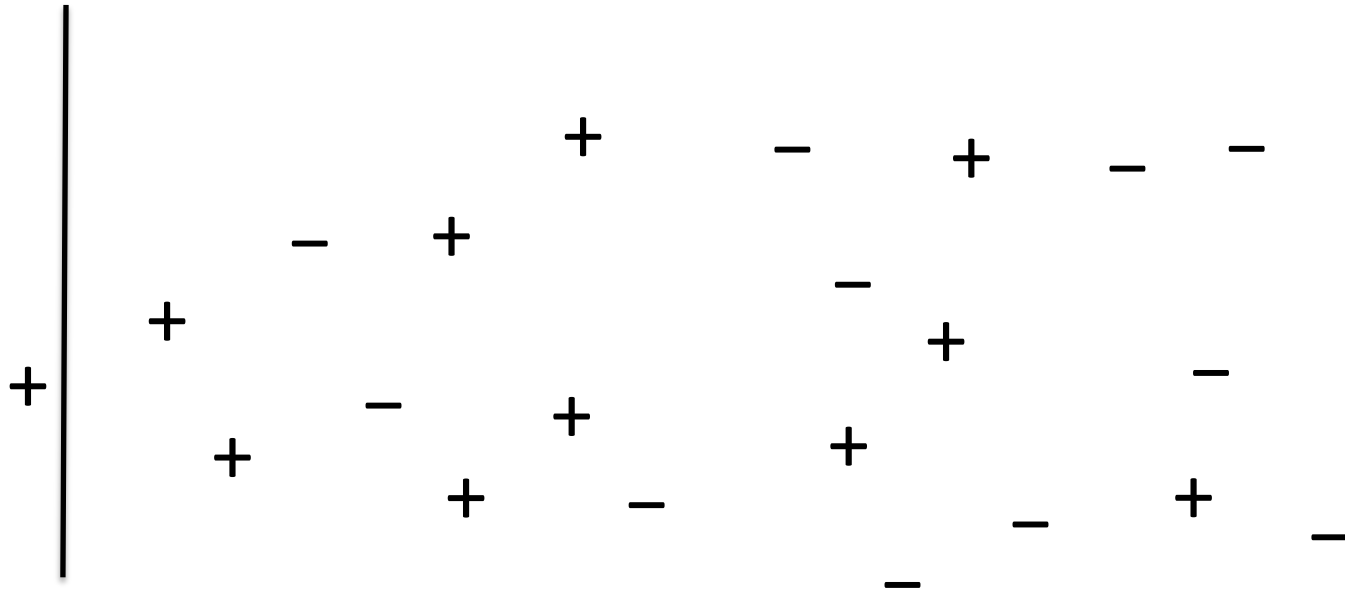
$$f(\mathbf{x}) = 7$$



- TPR = 1/11
- FPR = 0/12

ROC Curves

- Adjust the decision boundary



ROC Curves

- For a particular False Positive Rate (FPR), what is the True Positive Rate (TPR)?



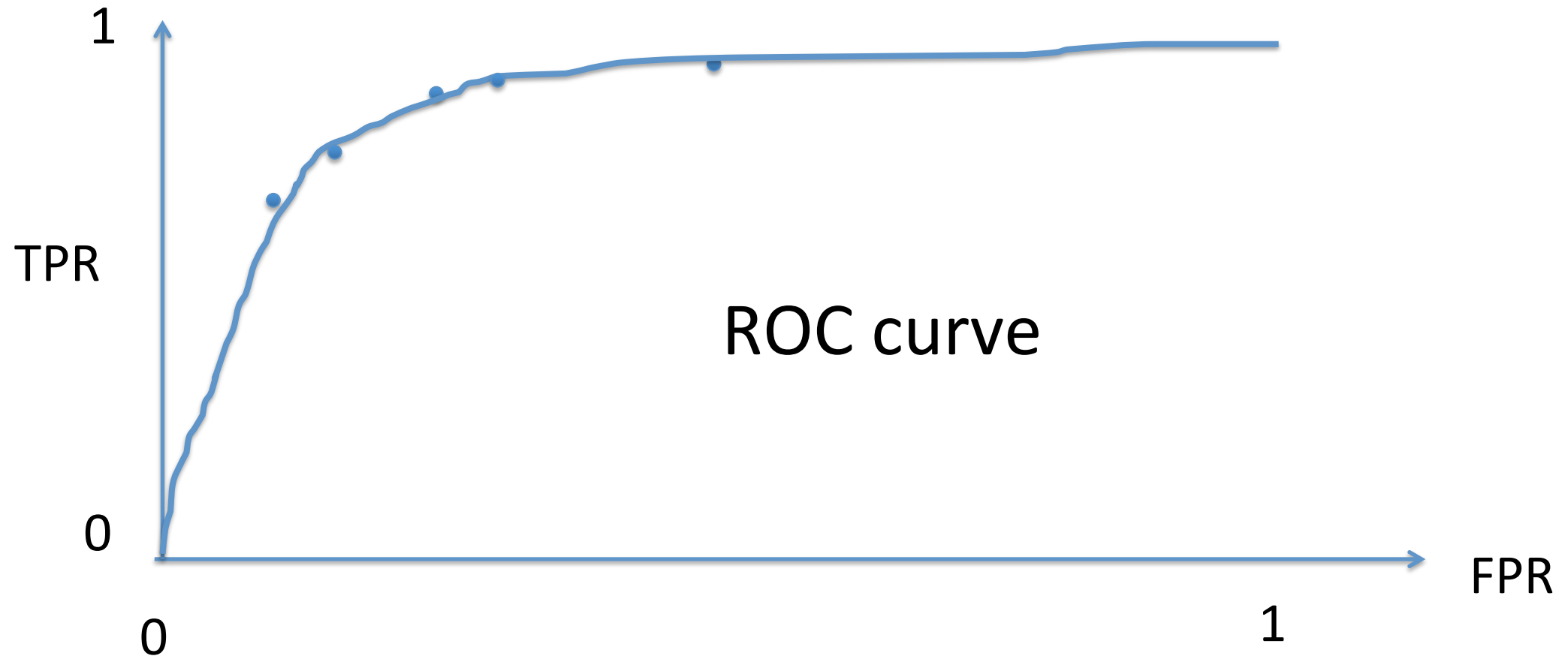
ROC Curves

- For a particular False Positive Rate (FPR), what is the True Positive Rate (TPR)?



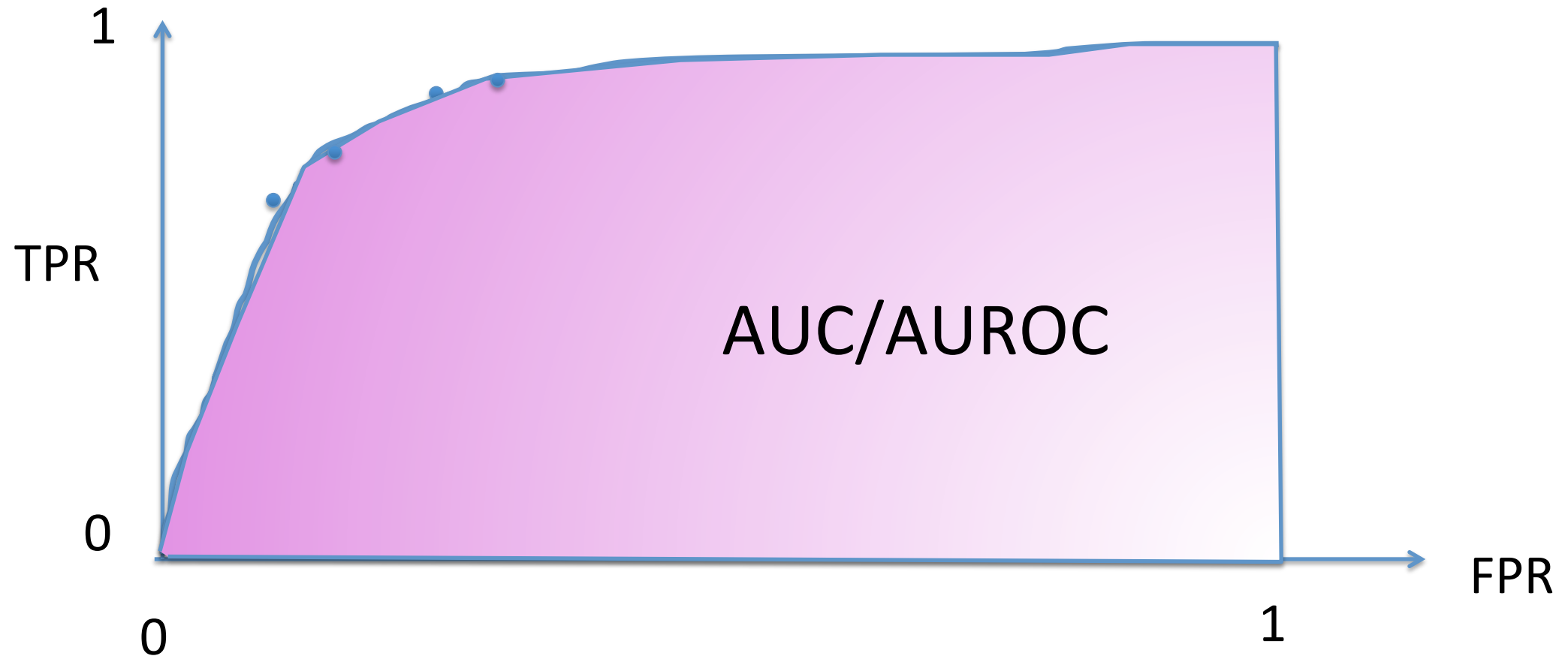
ROC Curves

- For a particular False Positive Rate (FPR), what is the True Positive Rate (TPR)?



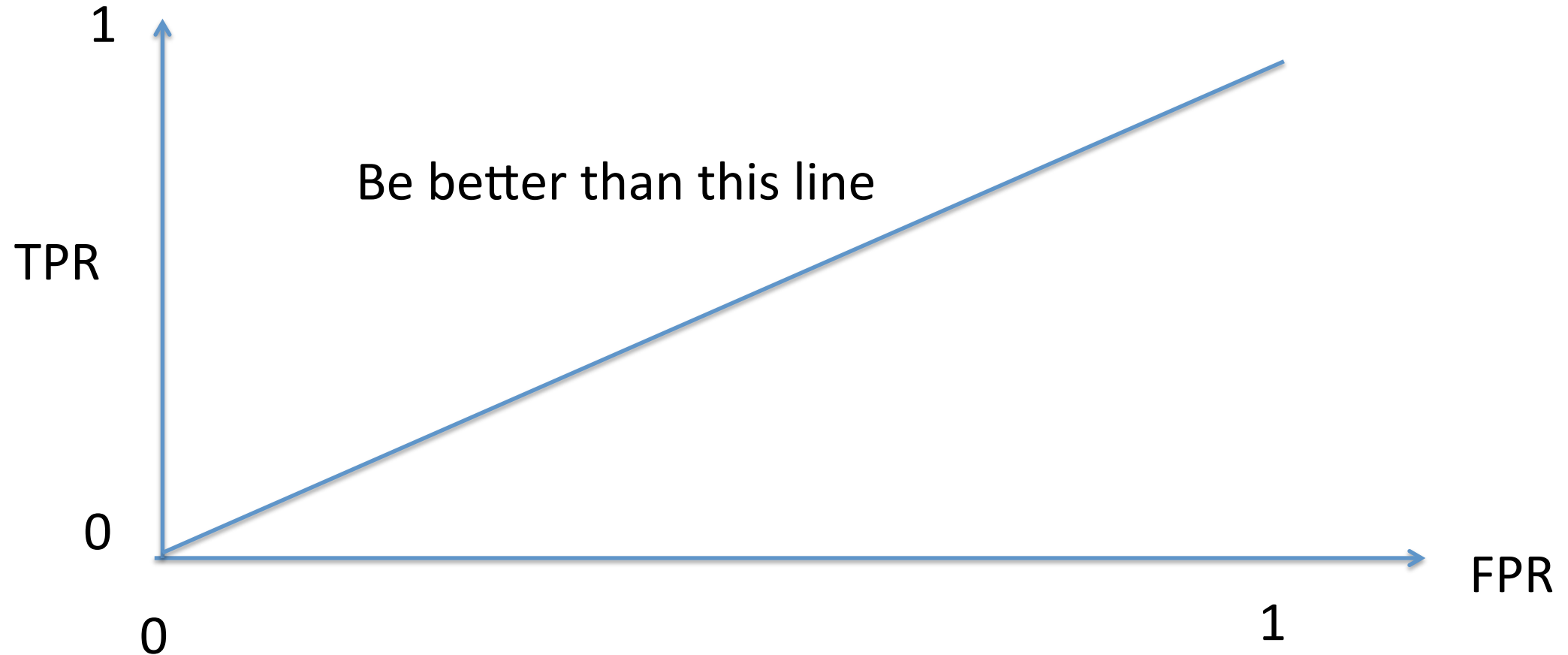
ROC Curves

- For a particular False Positive Rate (FPR), what is the True Positive Rate (TPR)?



ROC Curves

- For a particular False Positive Rate (FPR), what is the True Positive Rate (TPR)?



Evaluation

Many ways to evaluate a model:

- Confusion matrix (TP, TN, FP, FN)
- Accuracy / misclassification error
- Precision, Recall, F1-score
- ROC curves, AUC/AUROC