GENERAL ASSEMBLY

DATA SCIENCE

MODEL EVALUATION METRICS

Regression:

- *Root Mean Squared Error*

Classification:

- *Confusion Matrix*
- *ROC Curve (and AUC)*

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

- *Used for regression problems*
- *Square root of the mean of the squared errors*
- *Easily interpretable (in the "y" units)*
- *"Punishes" larger errors*

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

Example:

y_true = [100, 50, 30]

y_preds = [90, 50, 50]

*RMSE = np.sqrt((10**2 + 0**2 + 20**2) / 3) = 12.88*

Confusion Matrix: table to describe the performance of a classifier

n=165	Actual: YES	Actual: NO
Predicted: YES	100	10
Predicted: NO	5	50

Example: Test for presence of disease

YES = positive test = True = 1

NO = negative test = False = 0

- *How many classes are there?*
- *How many patients?*
- *How many predictions of disease?*
- *How many patients actually have the disease?*

n=165	Actual: YES	Actual: NO	
Predicted: YES	TP = 100	FP = 10	110
Predicted: NO	FN = 5	TN = 50	55
	105	60	

Basic Terminology:

- *True Positives (TP)*
- *True Negatives (TN)*
- *False Positives (FP)*
- *False Negatives (FN)*

Accuracy:

- *Overall, how often is it correct?*
- *(TP + TN) / total = 150/165 = 0.91*

Misclassification Rate (Error Rate):

- *Overall, how often is it wrong?*
- *1 - accuracy = 1 - 0.91 = 0.09*

n=165	Actual: YES	Actual: NO	
Predicted: YES	TP = 100	FP = 10	110
Predicted: NO	FN = 5	TN = 50	55
	105	60	

Precision:

- *TP / predicted yes = 100/110 = 0.91*

True Positive Rate:

- *TP / actual yes = 100/105 = 0.95*
- *"sensitivity" or "recall"*

False Positive Rate:

- *FP / actual no = 10/60 = 0.17*

Specificity:

- *1 - FPR = 1 - 0.17 = 0.83*

Email Number	Score	True Label
5	0.93	Spam
8	0.91	Spam
2	0.84	Spam
1	0.6	Ham
7	0.54	Spam
3	0.22	Ham
4	0.10	Ham
6	0.02	Ham

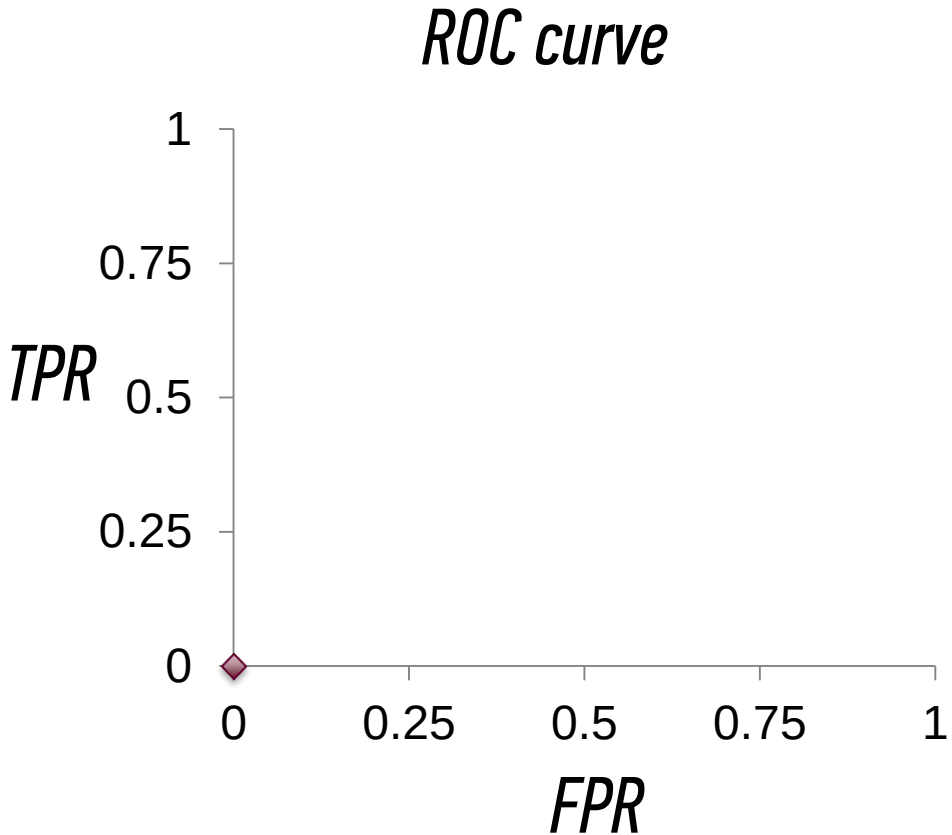
Every email gets a spamminess score.

Choosing a cut-off, this becomes a classification.

How do we choose a cut-off?
How do we evaluate the ranking without choosing a cut-off?

ROC CURVE / AUC

Email Number	Score	True Label
5	0.93	Spam
8	0.91	Spam
2	0.84	Spam
1	0.6	Ham
7	0.54	Spam
3	0.22	Ham
4	0.10	Ham
6	0.02	Ham



DATA SCIENCE