GENERAL ASSEMBLY

# CLASSIFICATION METRICS

INTRODUCTION

# ADVANCED CLASSIFICATION METRICS

# ADVANCED CLASSIFICATION METRICS

- While we've already discussed three different metrics to measure the effectiveness of a classification model, they've only given us an overall picture of how a model is performing.

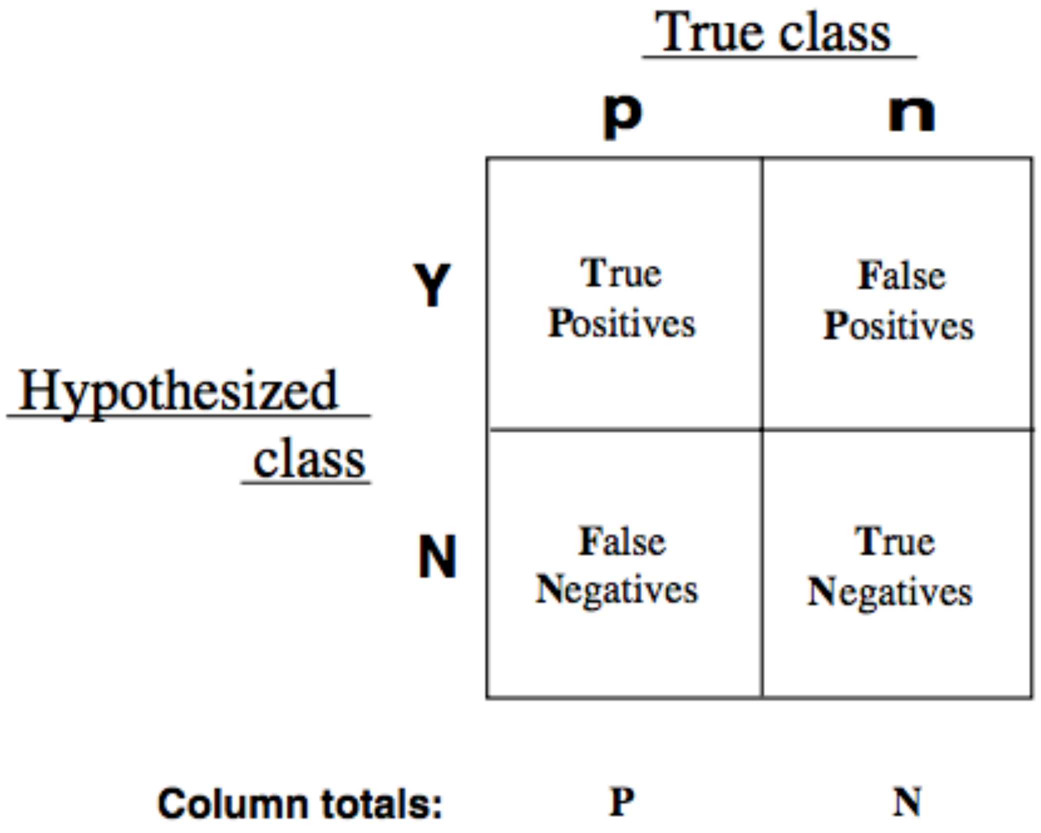
- What if we wanted to know exactly how a classifier was performing (e.g. what is predicting correctly vs incorrectly)?

INTRODUCTION

# Confusion Matrix

# ADVANCED CLASSIFICATION METRICS

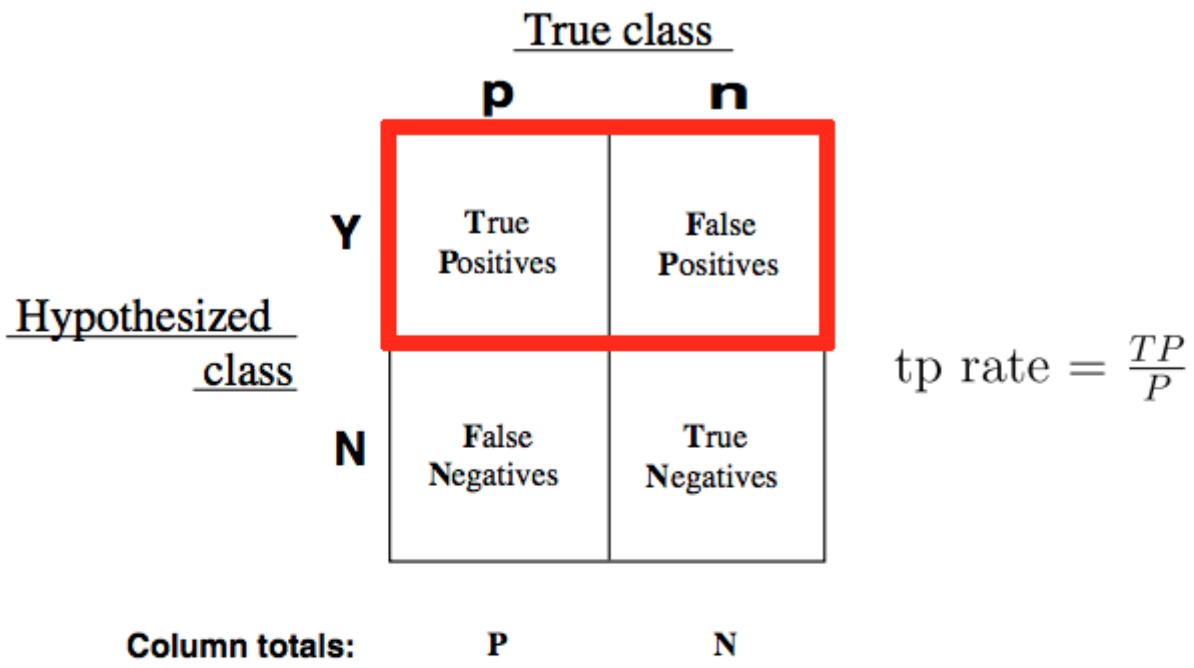
- We can use a confusion matrix to obtain more granular accuracy ratings for of each class by using the *true positive rate* and the *false positive rate*.



# ADVANCED CLASSIFICATION METRICS

- The **true positive rate (TPR)** asks, “Out of all of the target classes, how many were accurately predicted to belong to that class?”

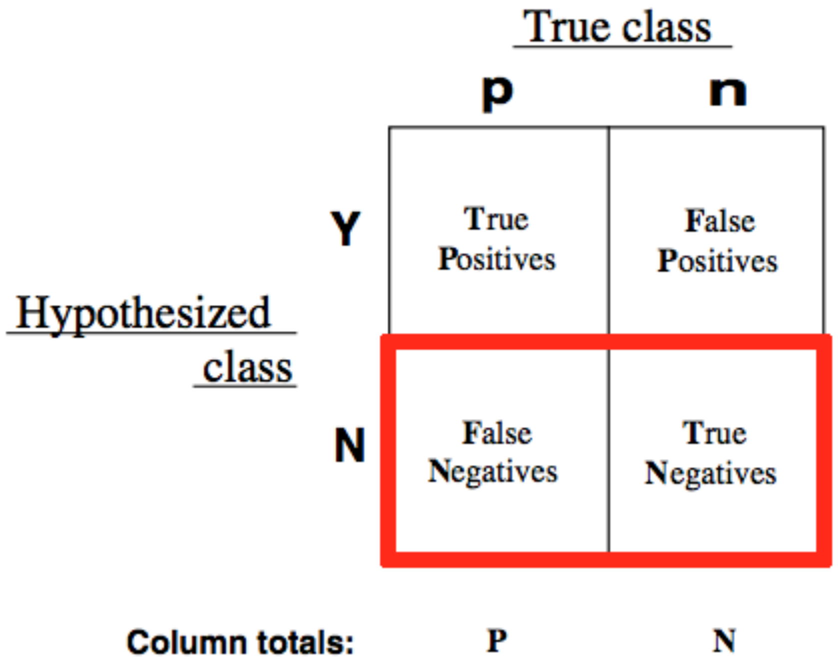
- Using our example, the TPR would be how often does our model <u>correctly</u> identify customer who will default on their credit card debt.



# ADVANCED CLASSIFICATION METRICS

- The **false positive rate (FPR)** asks, “Out of all items not belonging to a class, how many were predicted as belonging to that target class label?”

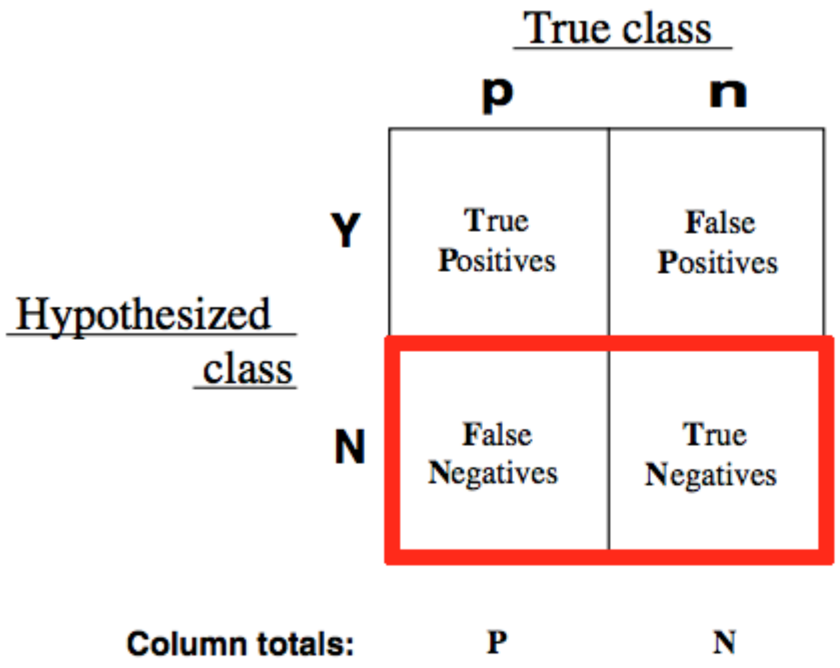
- Using our example, the FPR would be how often the model predict that a customer will default when they end up not doing so.



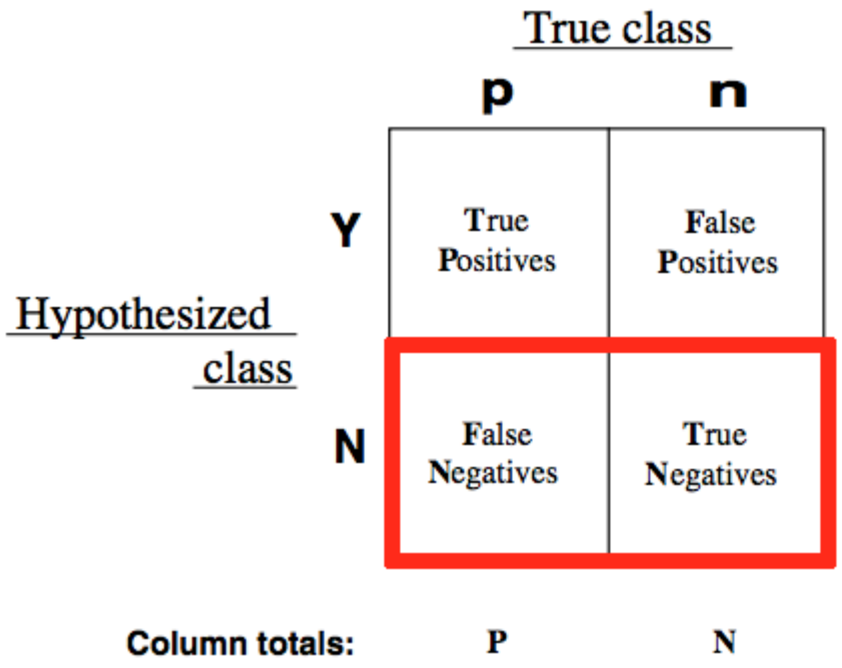

$$\text{fp rate} = \frac{FP}{N}$$

# ADVANCED CLASSIFICATION METRICS

- We can also measure the inverse of TPR/FPR or the false negative rate and the false negative rate (TNR).




$$\frac{TN}{TN+FN}$$




$$\frac{FN}{TN+FN}$$

# ADVANCED CLASSIFICATION METRICS

- These rates gives us a much clearer pictures of where model predictions begin to fall apart and exactly what business cases are being mishandled.

- This allows us to adjust our models accordingly and use metrics that best align to our business needs.

$$TPR = \frac{TP}{TP + FN}$$

$$FPR = \frac{FP}{TN + FP} = 1 - TPR$$

$$TNR = \frac{TN}{TN + FP}$$

$$FNR = \frac{FN}{TP + FN} = 1 - TNR$$

# ADVANCED CLASSIFICATION METRICS

- Ideally, our classifier would have a TPR approaching 1 and a FPR approaching 0. This would mean that our model is correctly predicting all customers who defaulted and not mistakenly predict that they wouldn't default.

- We can vary the classification threshold for our model to get different predictions, but how do we know if a model is better overall than other model?

- We can compare the FPR and TPR of the models, but it can often be difficult to optimize two numbers at once. *Can you think of any ways to combine our two metrics?*

# AREA UNDER THE CURVE

- If we have a TPR of 1 (all positives are marked positive) and FPR of 0 (all negatives are not marked positive), we'd have an AUC of 1. This means everything was accurately predicted.

- If we have a TPR of 0 (all positives are not marked positive) and an FPR of 1 (all negatives are marked positive), we'd have an AUC of 0. This means nothing was predicted accurately.

- An AUC of 0.5 would suggest a model no better than random is an excellent benchmark to use for comparing predictions (e.g. is my AUC above 0.5?).

# MORE CLASSIFICATION METRICS!

- There are several other common metrics that are similar to TPR and FPR that can also be useful.

True class

| Hypothesized class | p | n |
|---|---|---|
| Y | True Positives | False Positives |
| N | False Negatives | True Negatives |
| Column totals: | P | N |

$$\text{fp rate} = \frac{FP}{N} \qquad \text{tp rate} = \frac{TP}{P}$$

$$\text{precision} = \frac{TP}{TP+FP} \qquad \text{recall} = \frac{TP}{P}$$

$$\text{accuracy} = \frac{TP+TN}{P+N}$$

$$\text{F-measure} = \frac{2}{1/\text{precision}+1/\text{recall}}$$

- Sklearn has

INTRODUCTION

# Building a Confusion Matrix

# HOW… MEASURE PERFORMANCE?

*Confusion Matrix: table to describe the performance of a classifier*

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

*Example: Test for presence of disease*
*NO = negative test = False = 0*
*YES = positive test = True = 1*

- *How many classes are there?*
- *How many patients?*
- *How many times is disease predicted?*
- *How many patients actually have the disease?*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*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**?*
- *(FP + FN) / total = 15/165 = 0.09*

# CONFUSION MATRIX

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

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

*Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# CONFUSION MATRIX

| n=165 | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | TN = 50 | FP = 10 | 60 |
| Actual: YES | FN = 5 | $\boxed{\text{TP = 100}}$ | 105 |
| | 55 | 110 | |

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

→ *Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# CONFUSION MATRIX

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

*Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# CONFUSION MATRIX

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

*Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# CONFUSION MATRIX

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

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

*Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# CONFUSION MATRIX

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

*Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# CONFUSION MATRIX

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

*Sensitivity:*

- *When actual value is **positive**, how often is prediction **correct**?*
- *TP / actual yes = 100/105 = 0.95*
- *"True Positive Rate" or "Recall"*

*False Positive Rate:*

- *When actual value is **negative**, how often is prediction **wrong**?*
- *FP / actual no = 10/60 = 0.17*

*Specificity:*

- *When actual value is **negative**, how often is prediction **correct**?*
- *TN / actual no = 50/60 = 0.83*

# ROC AUC

## THE ROC CURVE

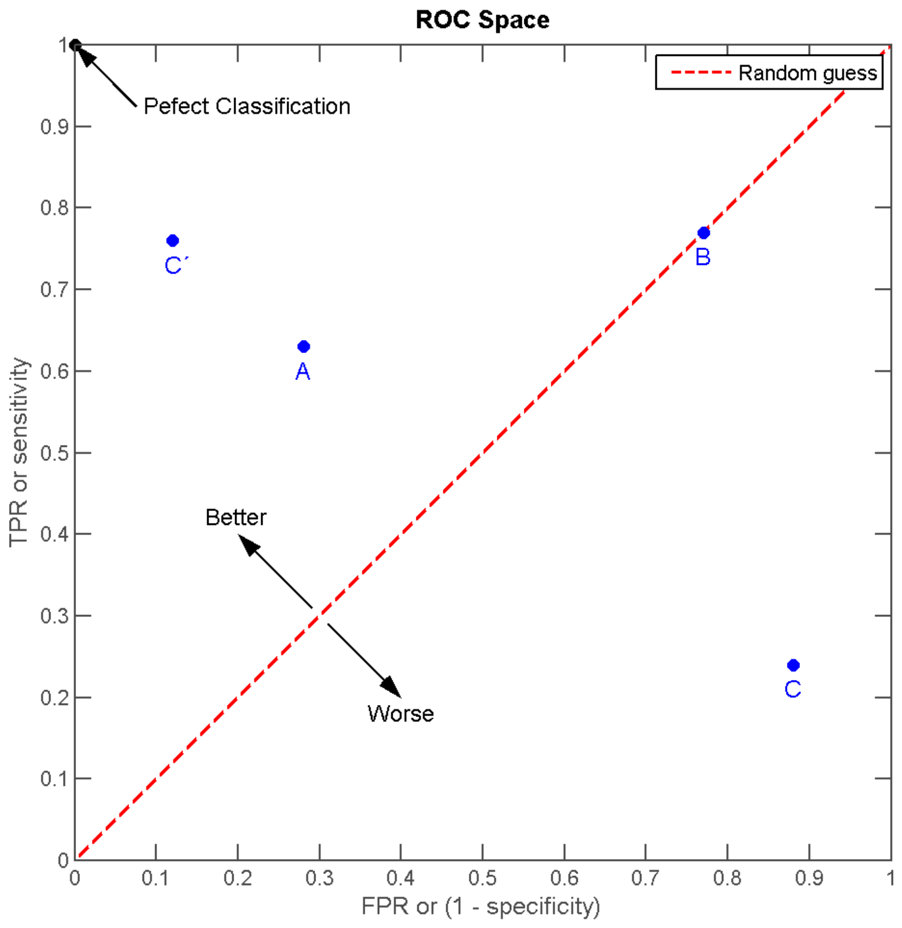
- This is where the Receiver Operation Characteristic (ROC) curve comes in handy.

- The curve is created by plotting the TPR against the FPR at various model classification settings.

- Area Under the Curve (AUC) summarizes the impact of TPR and FPR in a single value.

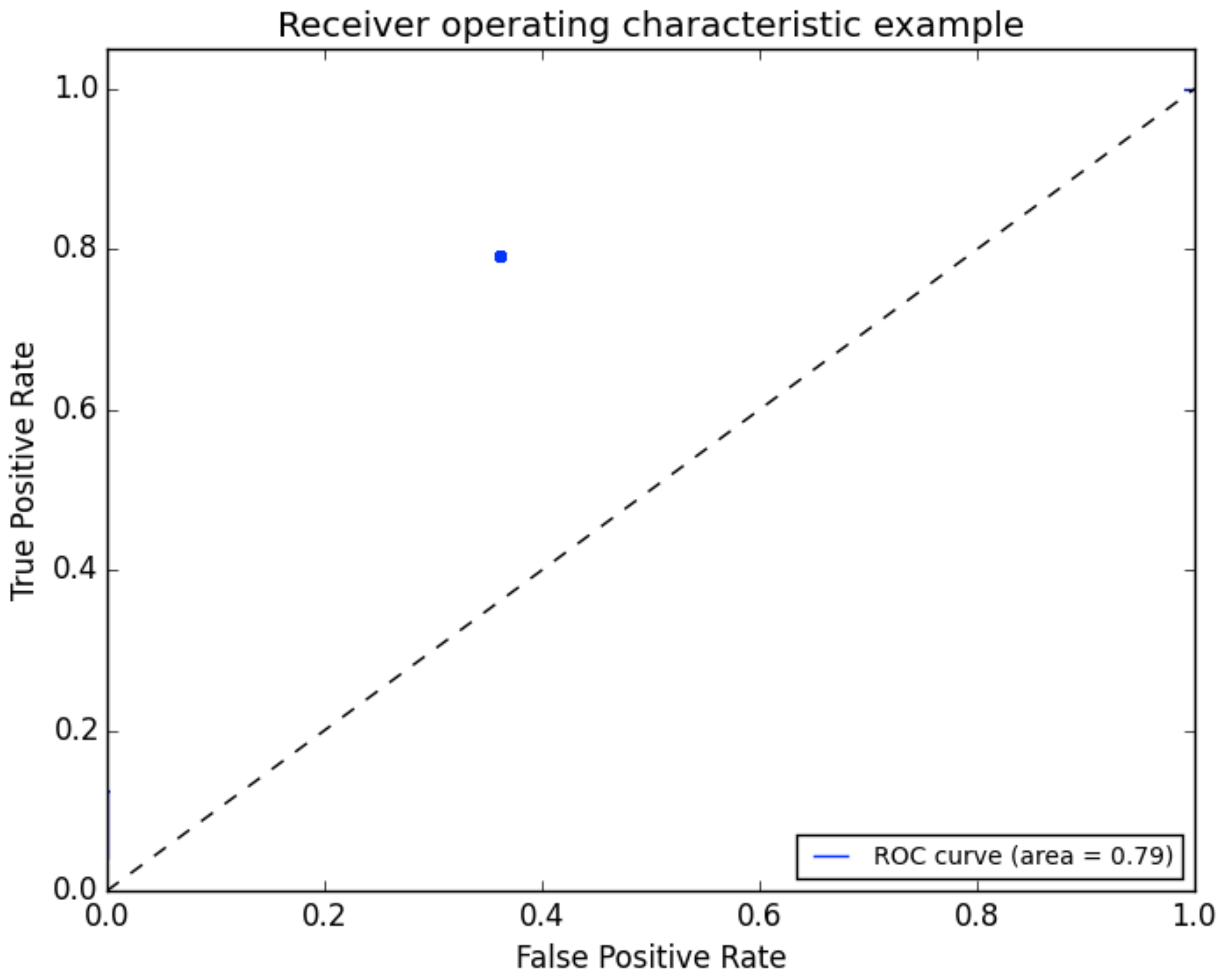
# THE ROC CURVE

- There can be a variety of points on an ROC curve.



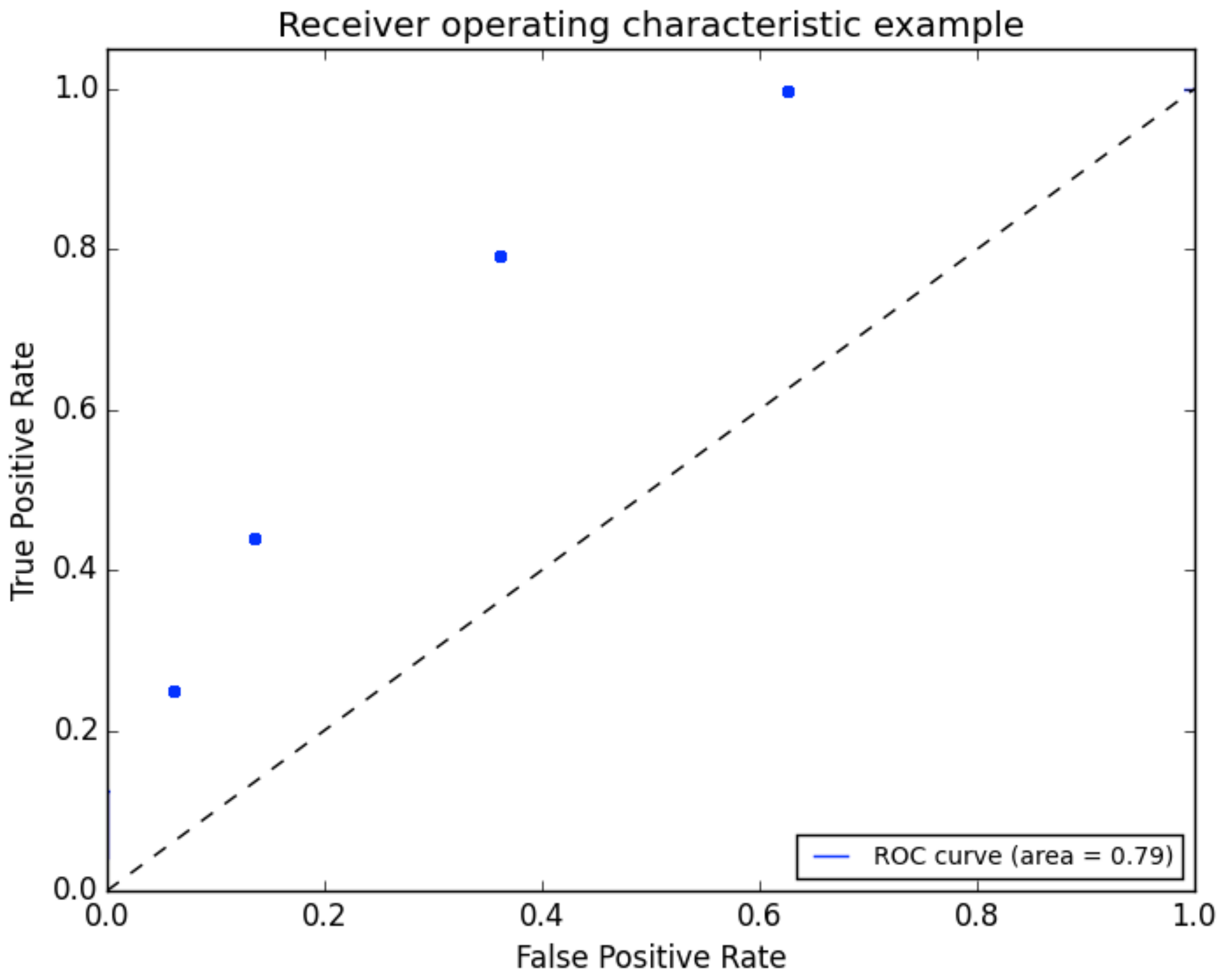
# THE ROC CURVE

- We can begin by plotting an individual TPR/FPR pair for one threshold.



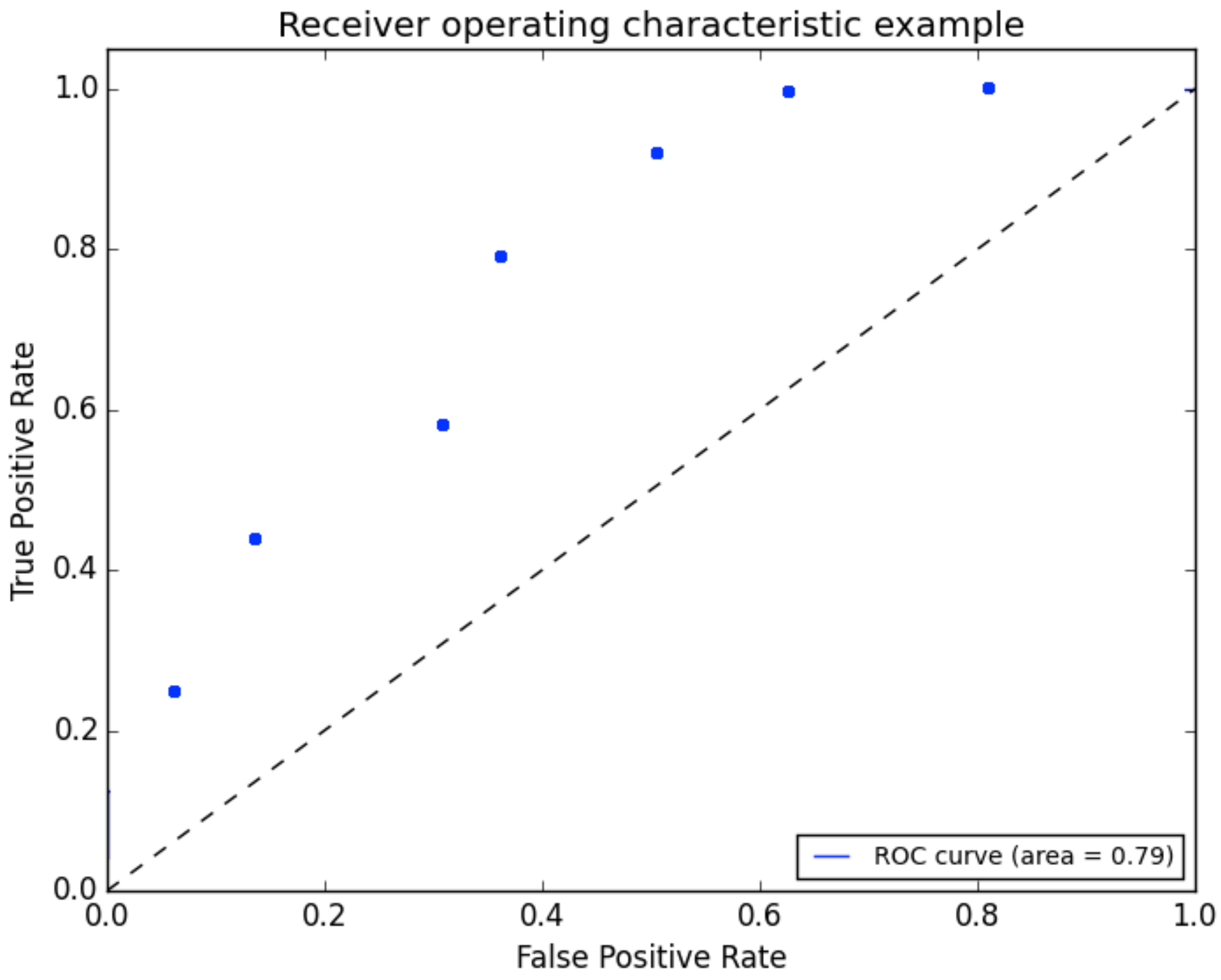
# THE ROC CURVE

- We can continue adding pairs for different thresholds.



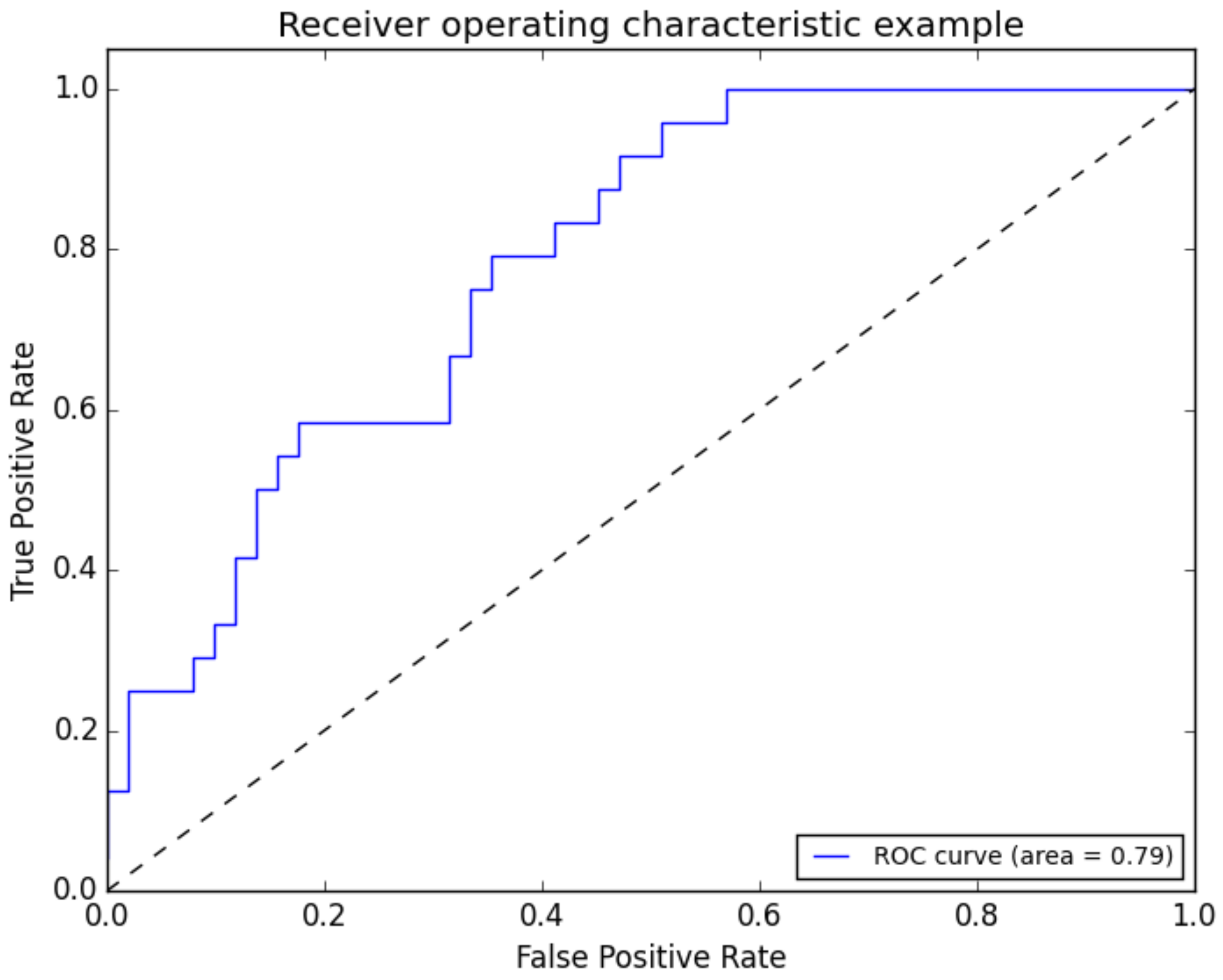
# THE ROC CURVE

- We can continue adding pairs for even more thresholds.



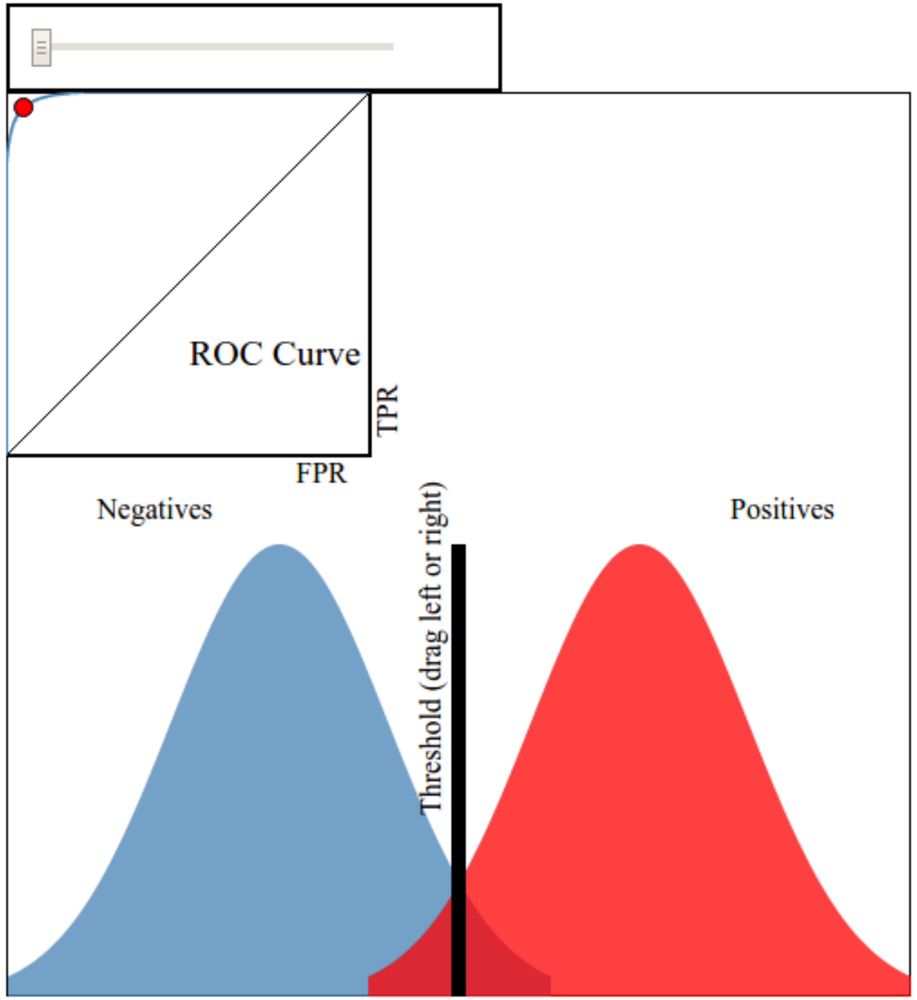
# THE ROC CURVE

- Finally, we create a full "curve" that is described by both TPR and FPR.



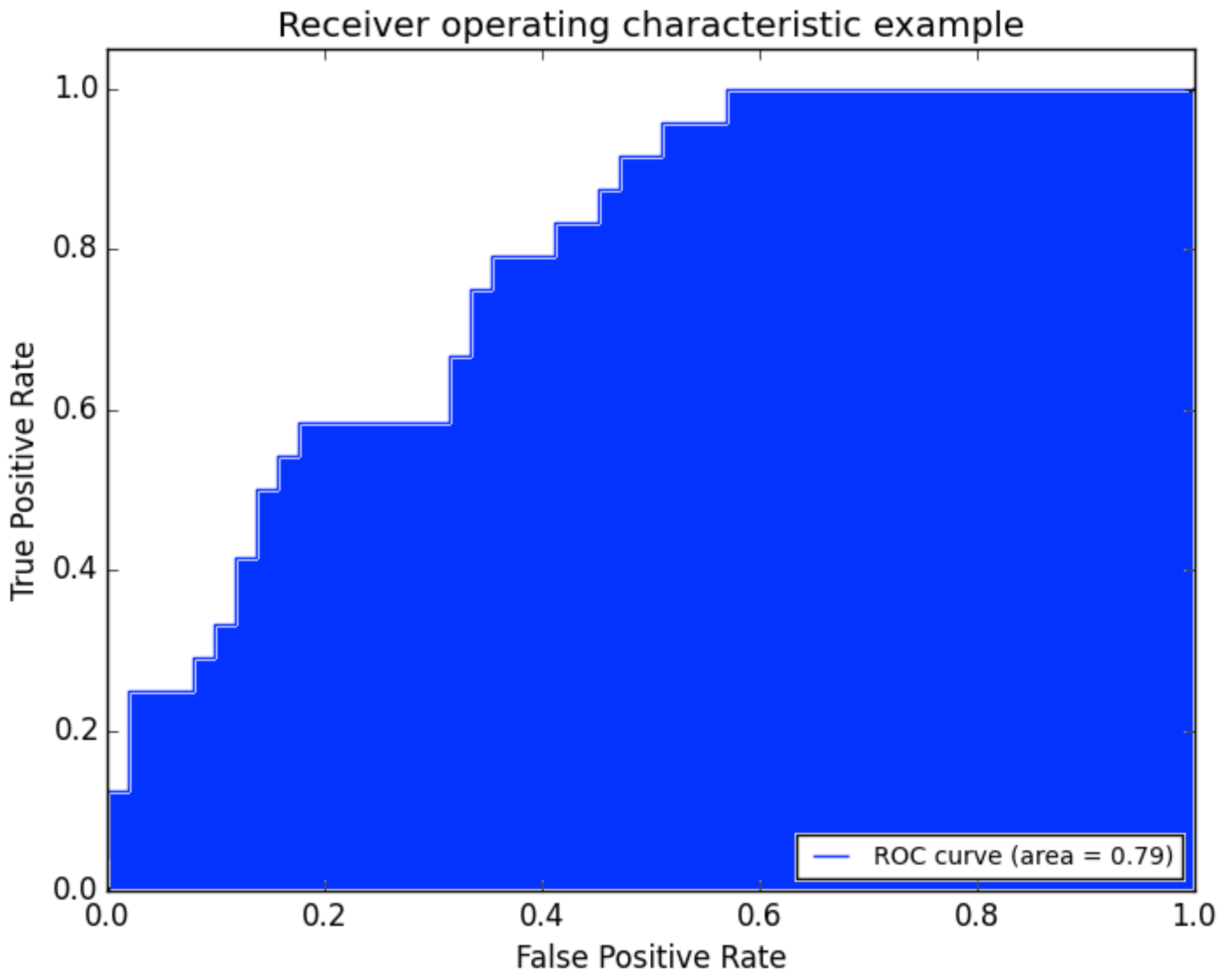
# THE ROC CURVE

- This interactive visualization can help practice visualizing ROC curves.



# AREA UNDER THE CURVE

- With this curve, we can find the Area Under the Curve (AUC).



# RECEIVING OPERATOR CHARACTERISTIC (ROC) CURVE

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

*Every email is assigned a “spamminess” score by our classification algorithm. To actually make our predictions, we choose a numeric cutoff for classifying as spam.*

*An ROC Curve will help us to visualize how well our classifier is doing without having to choose a cutoff!*

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

*The ROC plots the True Positive Rate (TRP) on the y-axis against the False Positive Rate (FPR) on the x-axis.*

*<u>TPR</u>: When actual value is **spam**, how often is prediction **correct**?*

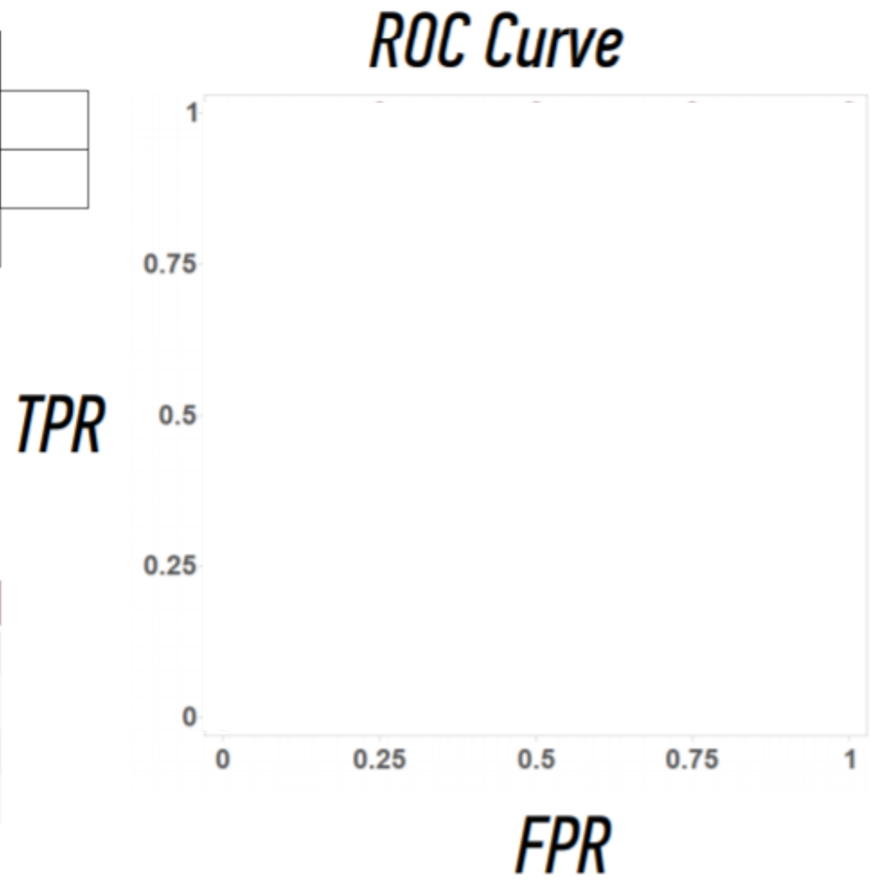
*<u>FPR</u>: When actual value is **ham**, how often is prediction **wrong**?*

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | | |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| 0 | ? | ? | 0.50 | | |
| 0.05 | | | 0.65 | | |
| 0.15 | | | 0.85 | | |
| 0.25 | | | 1 | | |

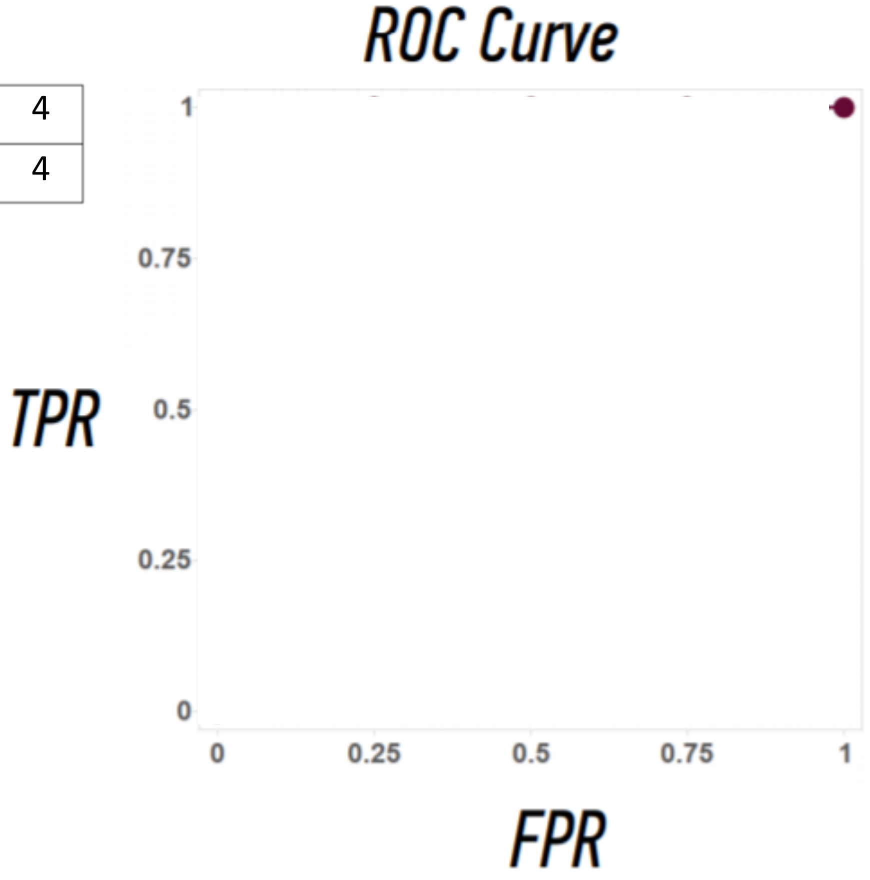


# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 0 | 4 | 4 |
| Actual: YES | 0 | 4 | 4 |
| | 0 | 8 | |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| 0 | 1 | 1 | 0.50 | | |
| 0.05 | | | 0.65 | | |
| 0.15 | | | 0.85 | | |
| 0.25 | | | 1 | | |

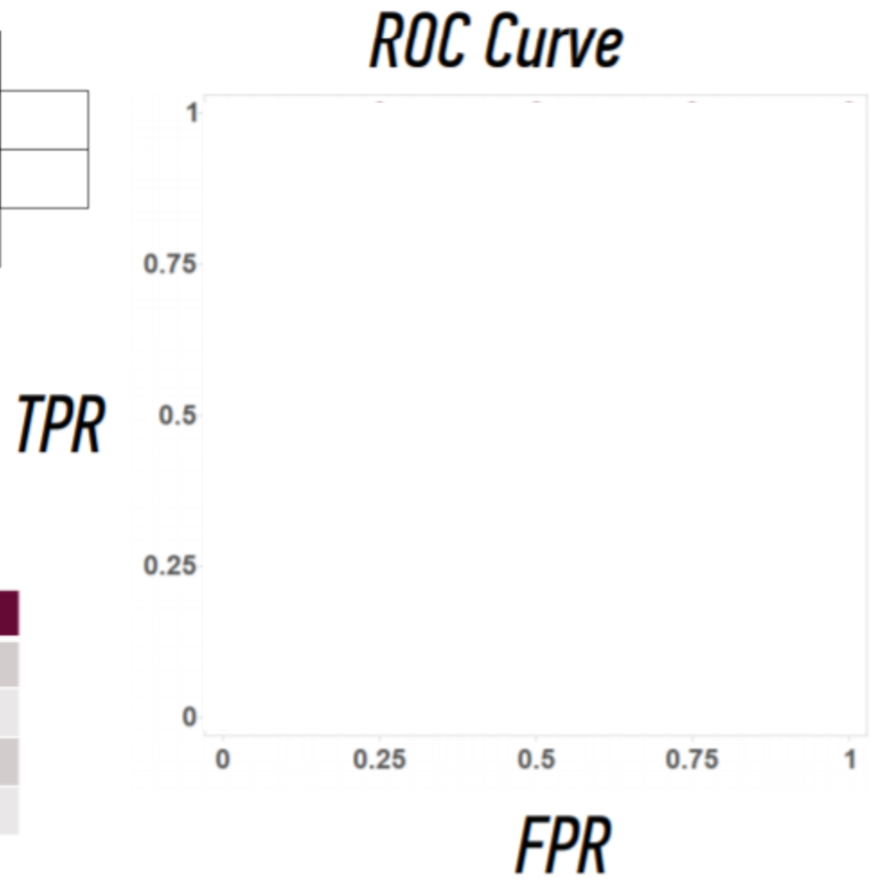


# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | | |

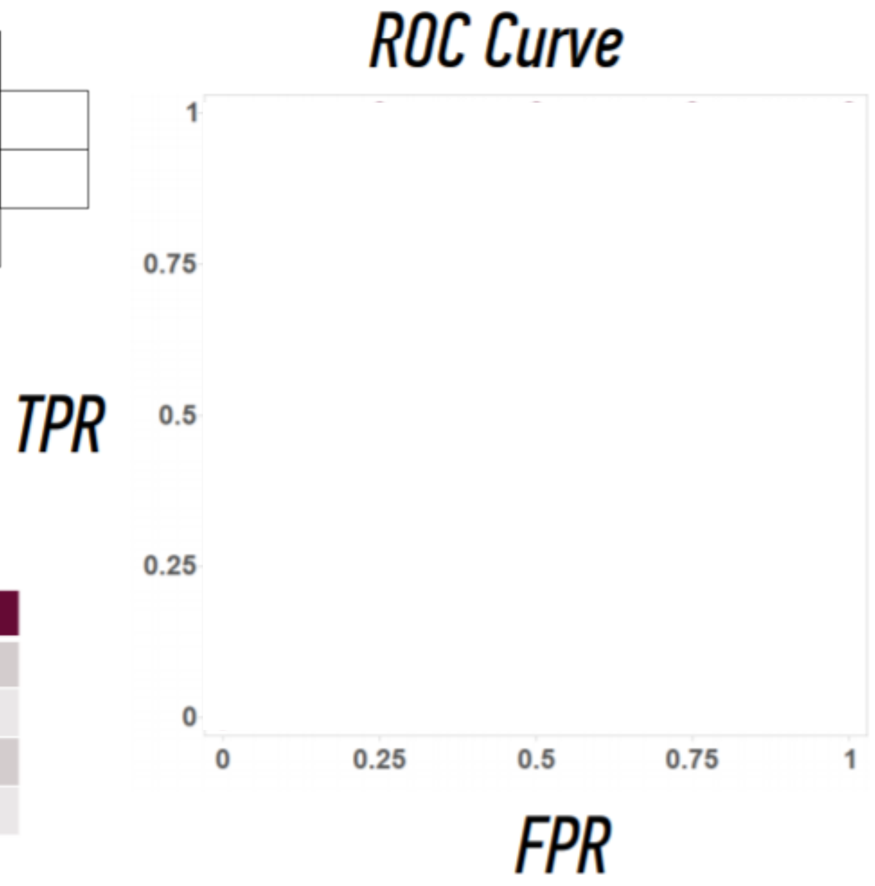
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | | |



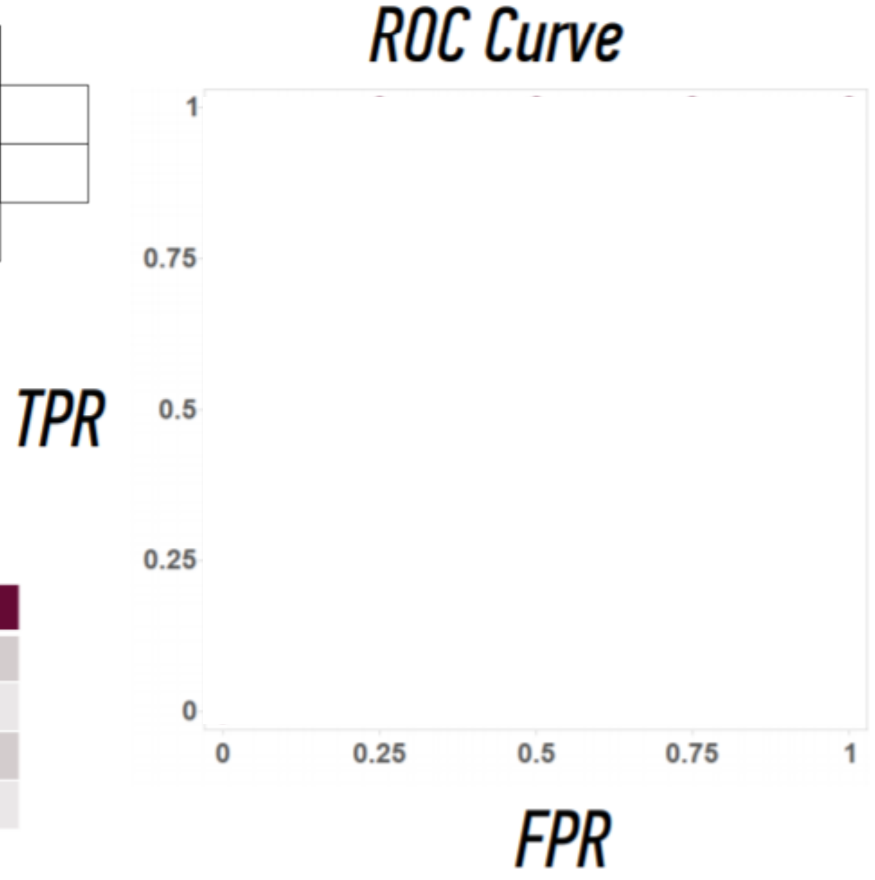

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | 1 | |

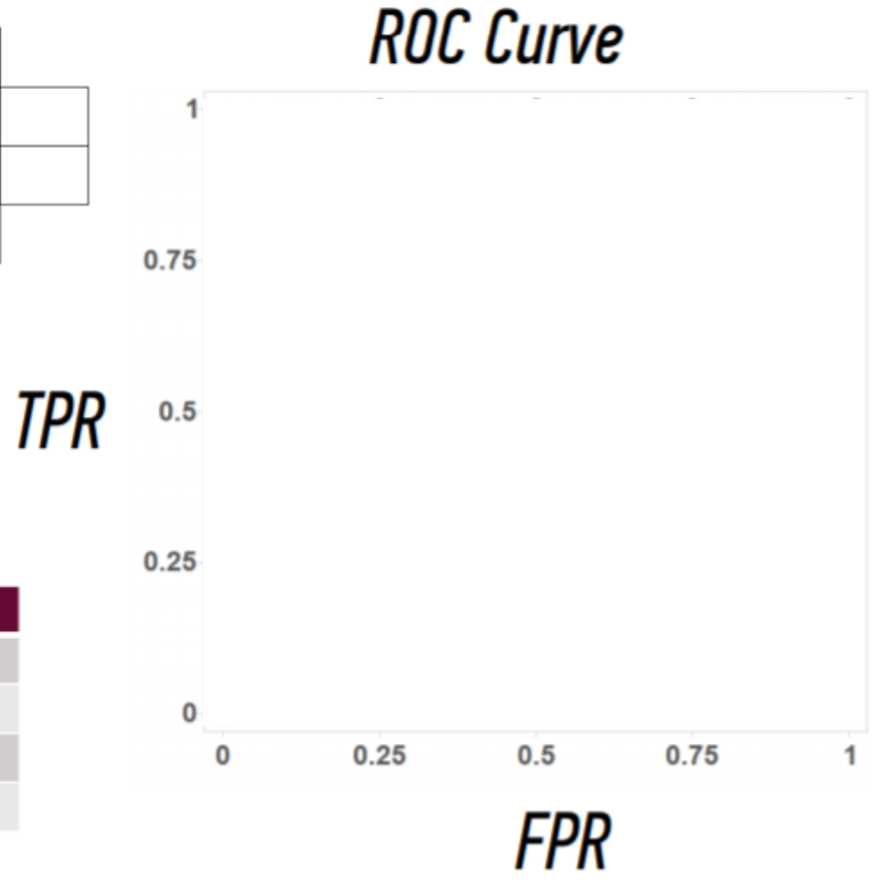
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | 2 | |




| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | 3 | |

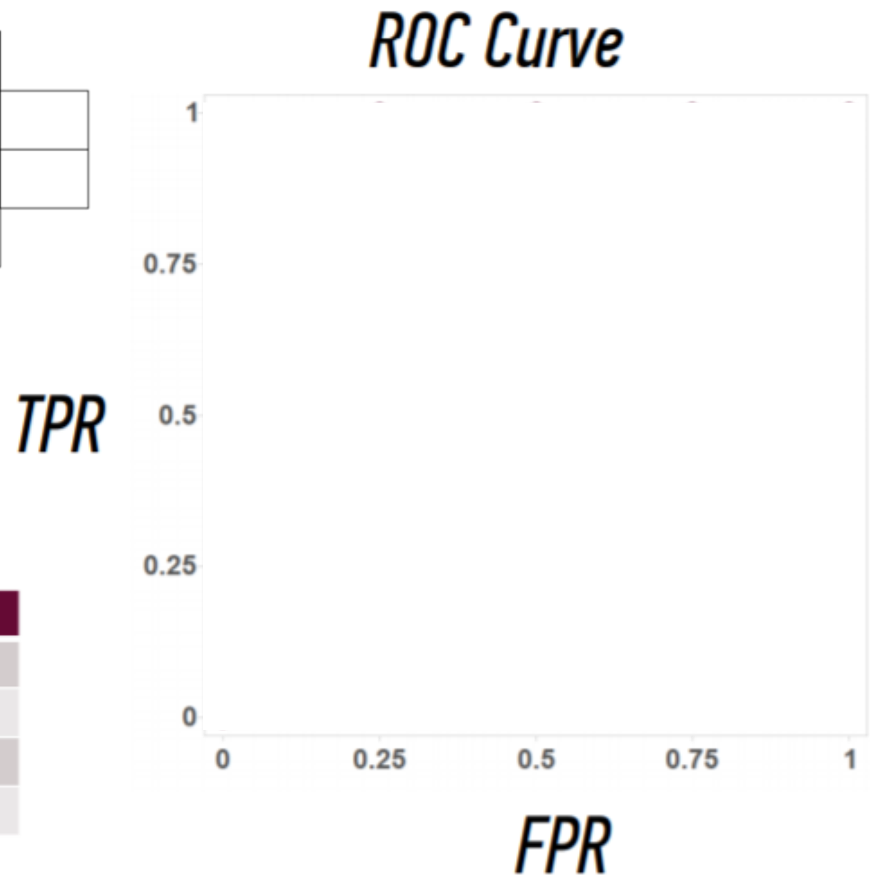
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | 4 | |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

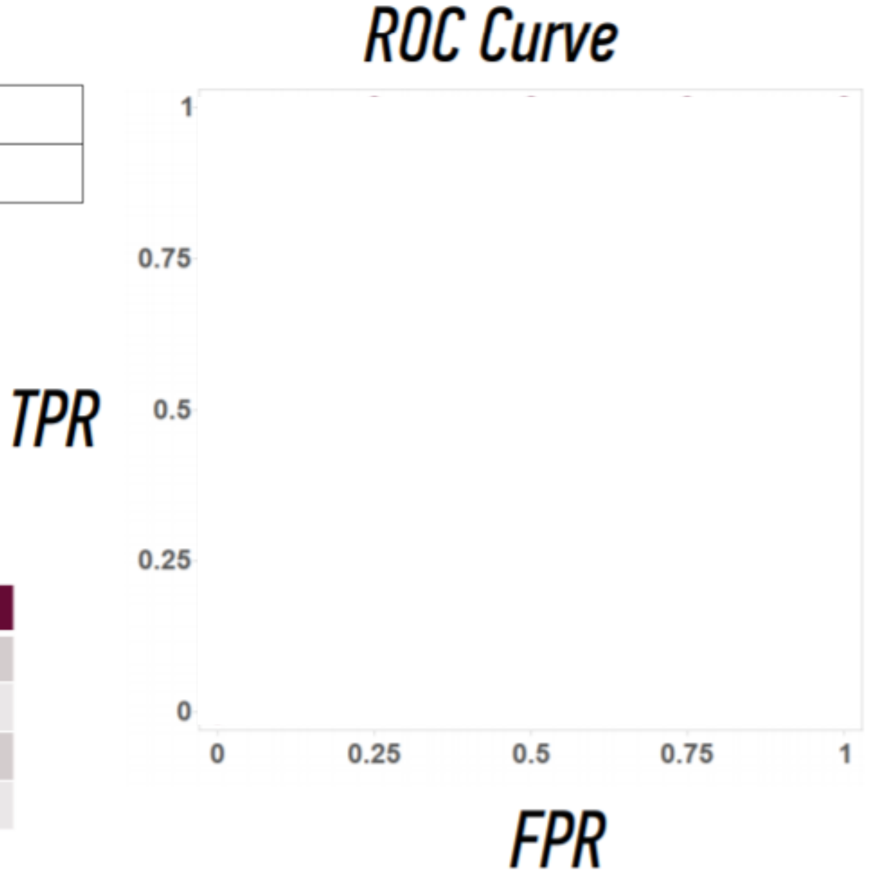


# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | 5 | |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

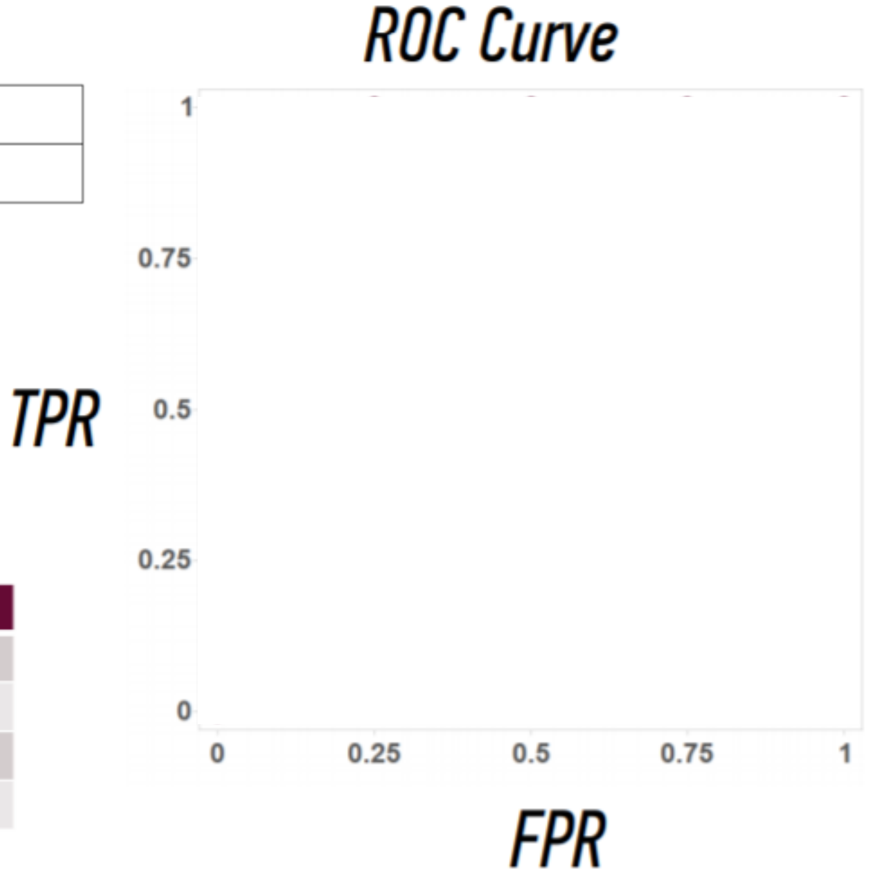


# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | | 7 | |

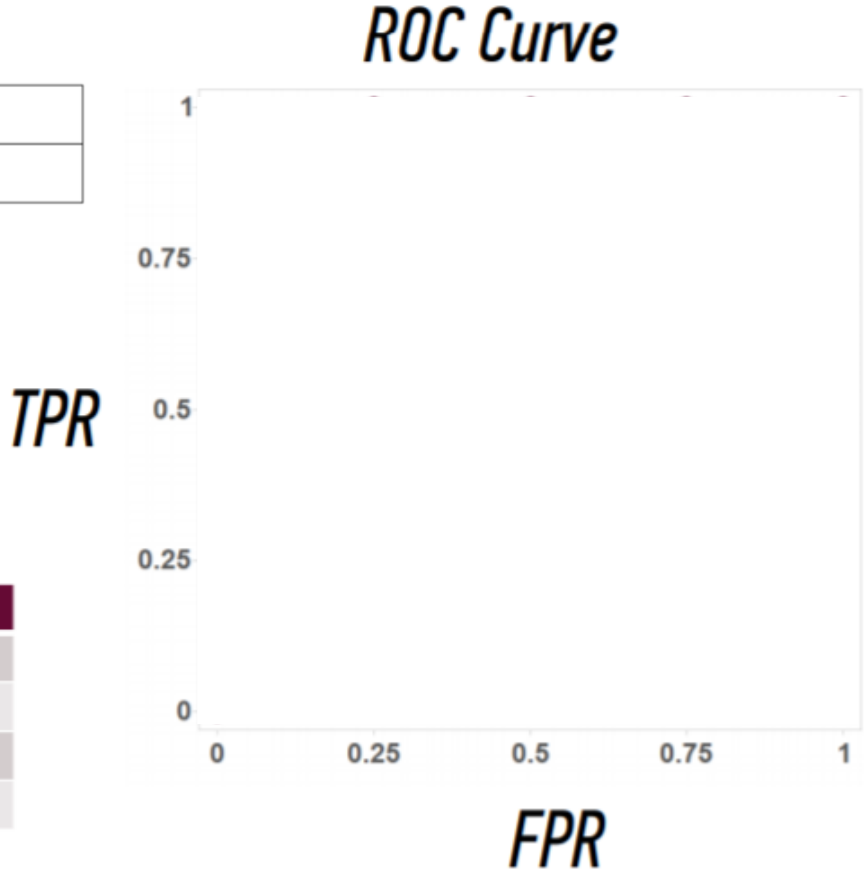
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | | |
| Actual: YES | | | |
| | 1 | 7 | |

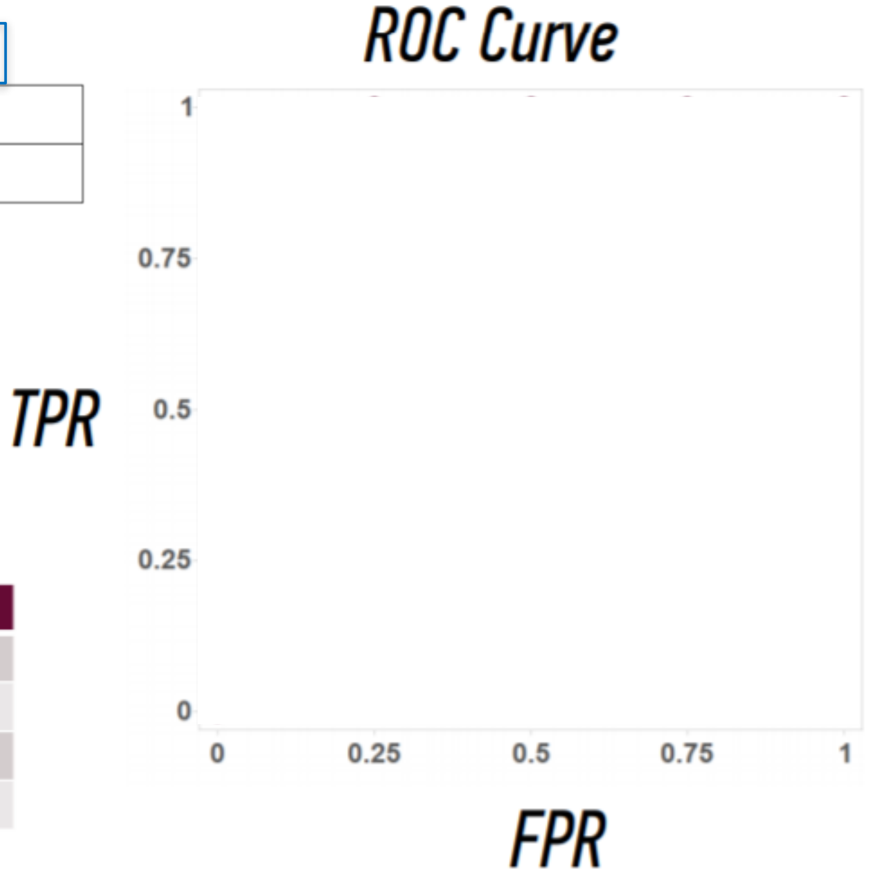



| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | | 3 | |
| Actual: YES | | | |
| | 1 | 7 | |

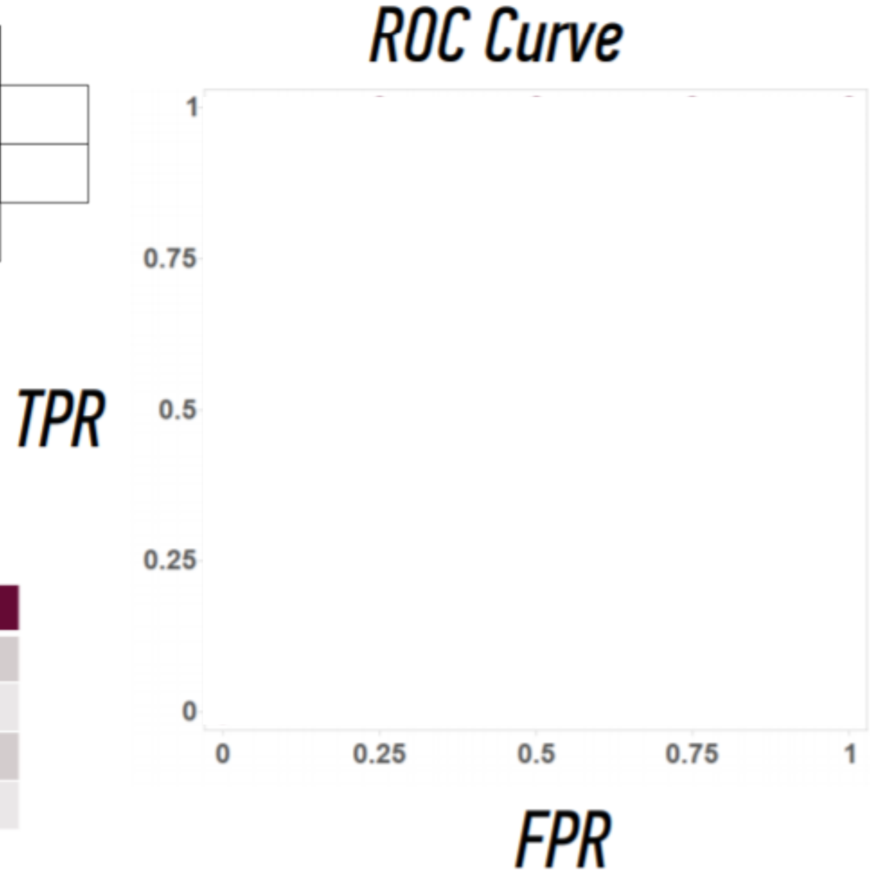



| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | |
| Actual: YES | | | |
| | 1 | 7 | |



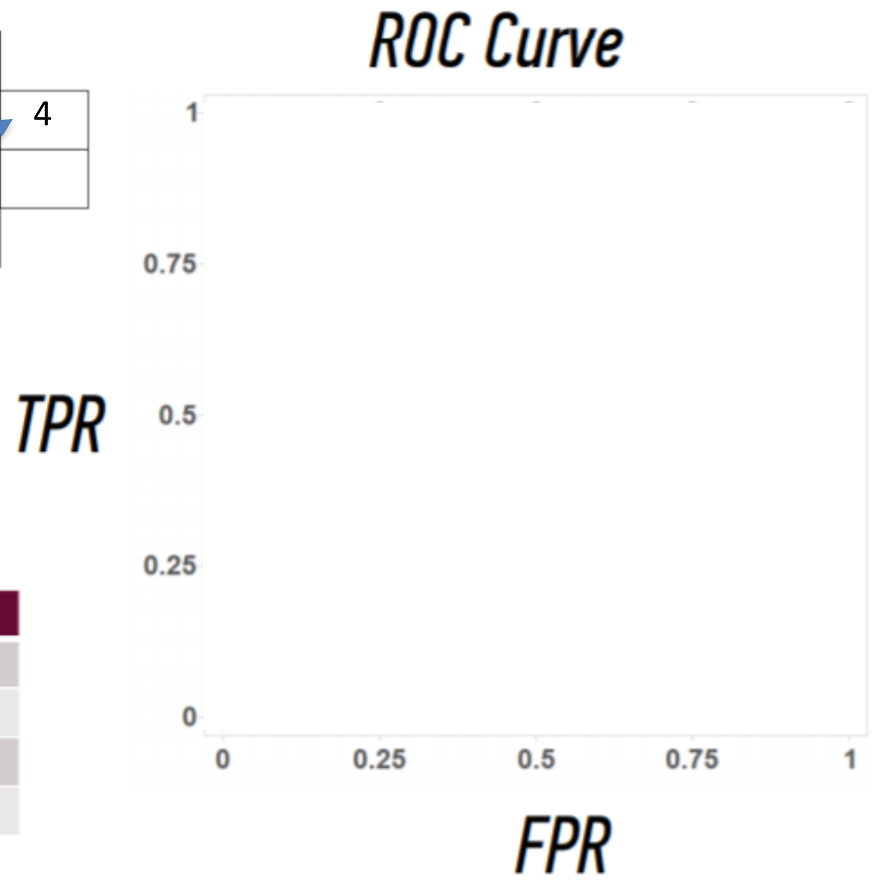

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | | | |
| | 1 | 7 | |

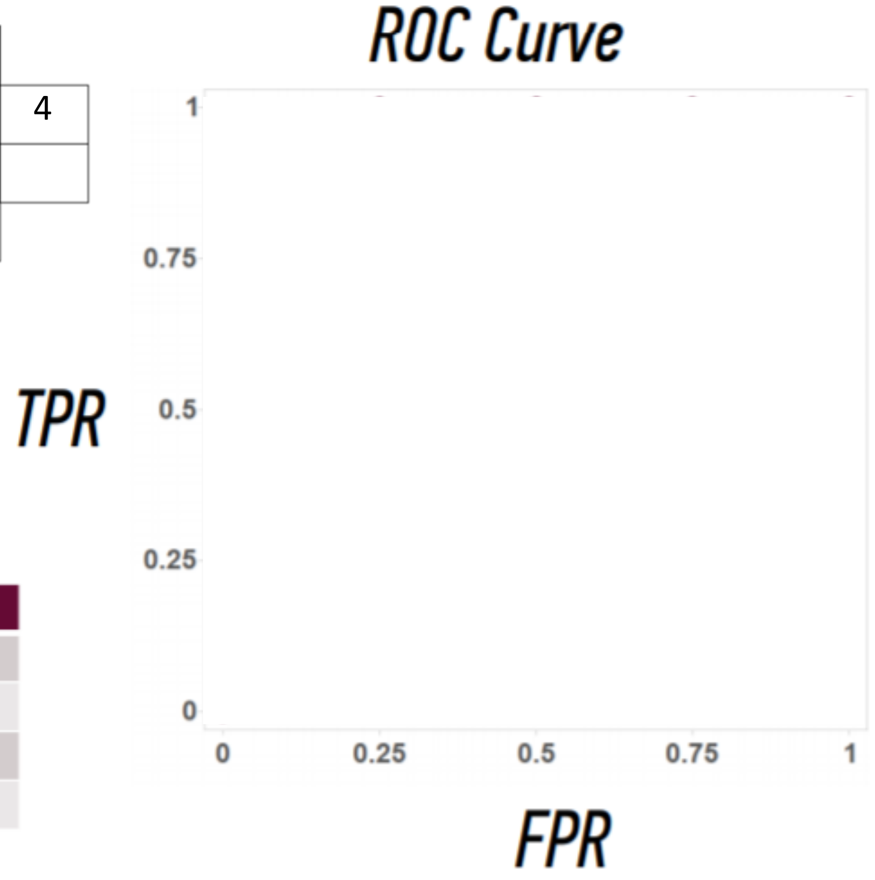
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | ? | | |
| | 1 | 7 | |



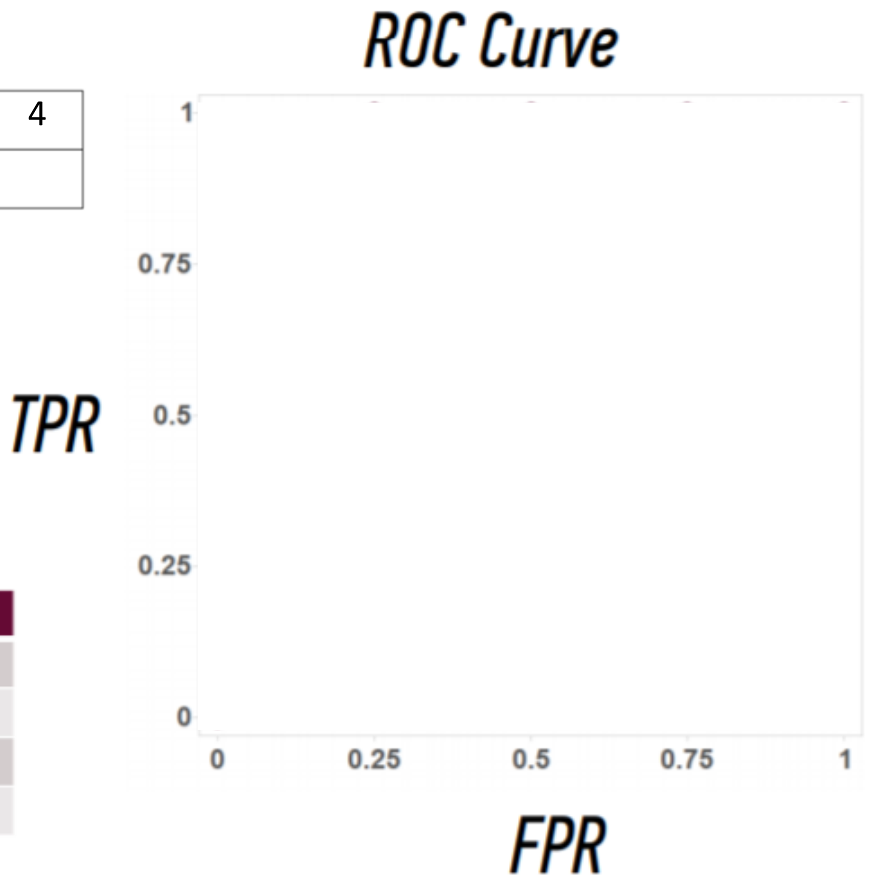

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | | |
| | 1 | 7 | |

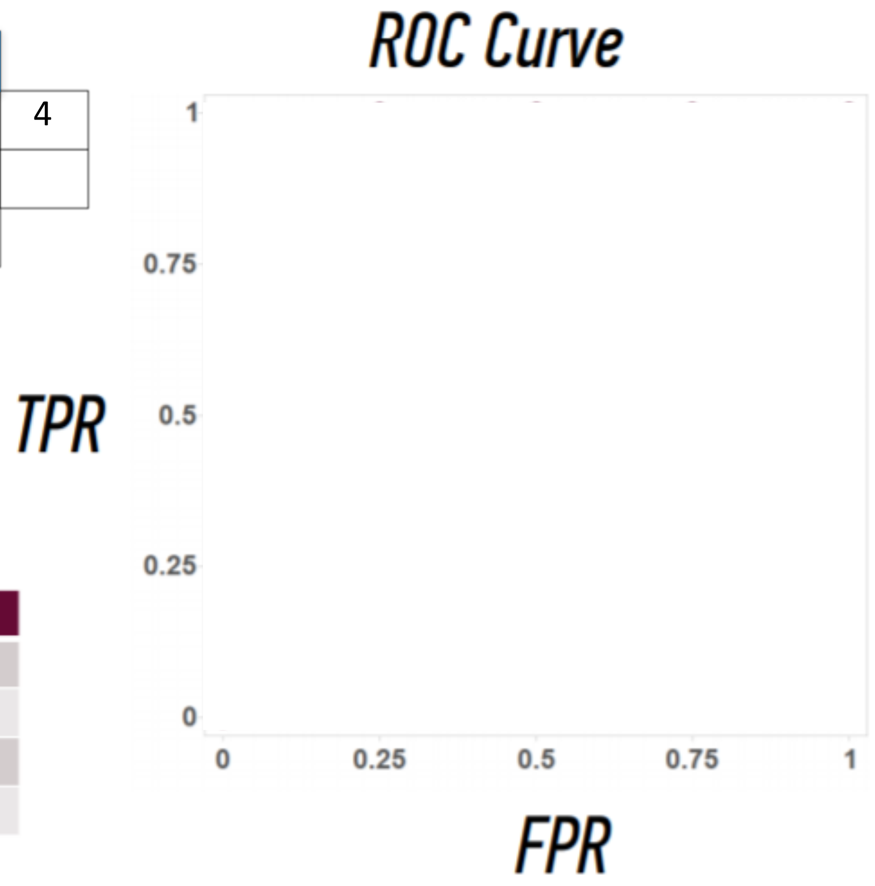
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | ? | |
| | 1 | 7 | |



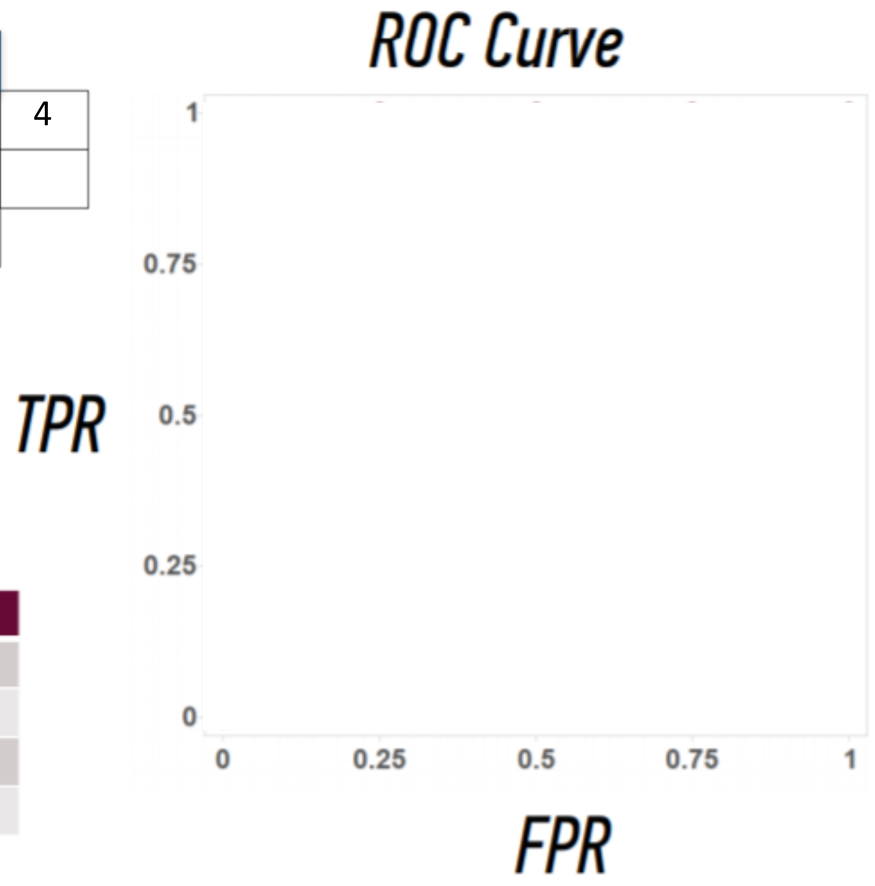

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | ? | |
| | 1 | 7 | |

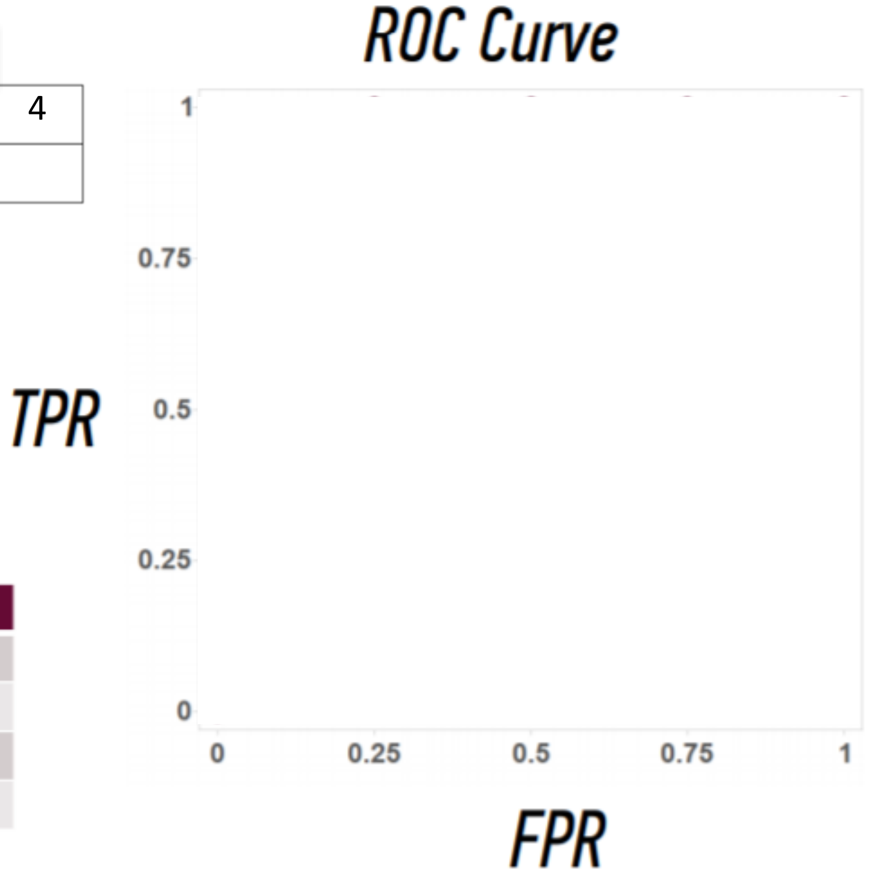
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | ? | |
| | 1 | 7 | |



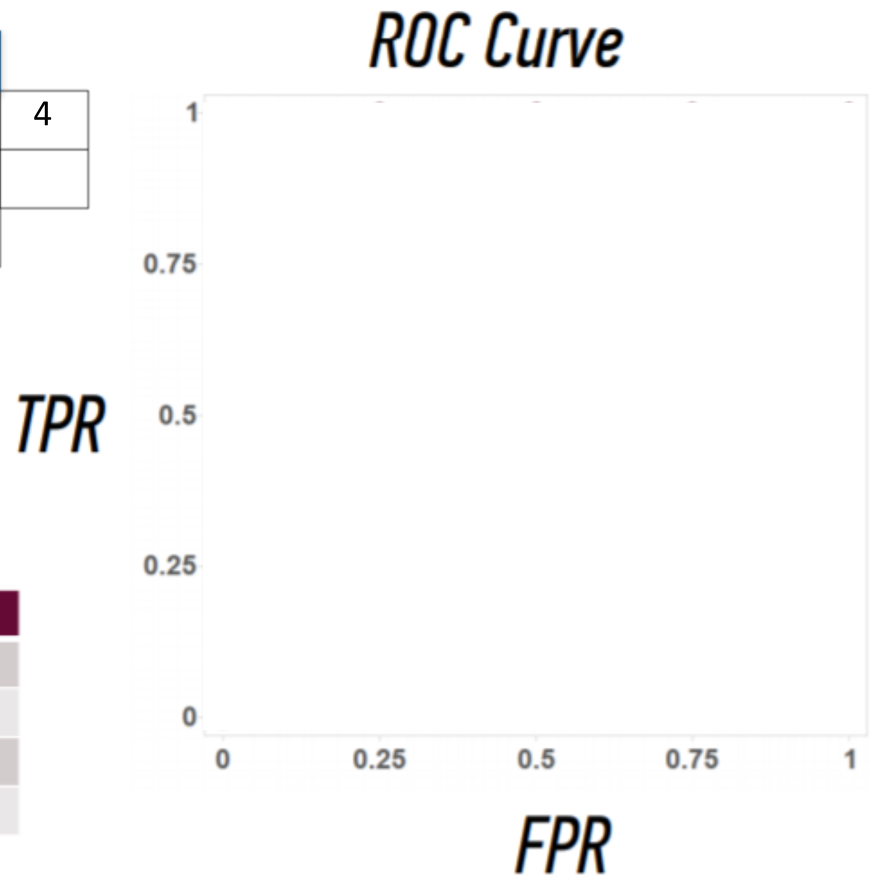

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | 4 | |
| | 1 | 7 | |

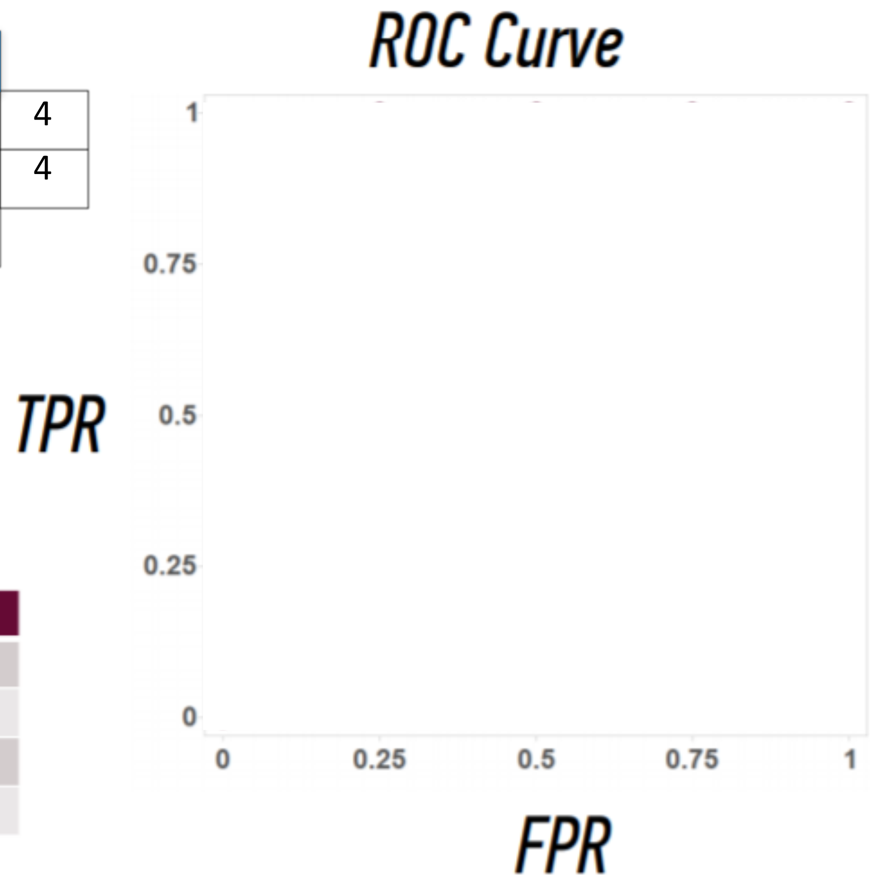
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | 4 | 4 |
| | 1 | 7 | |




| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

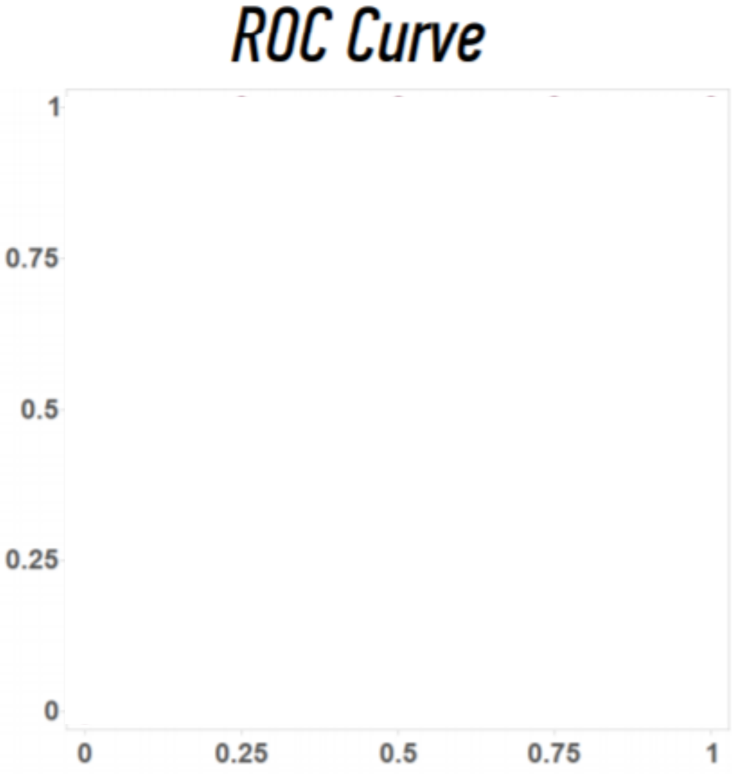
# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | 4 | 4 |
| | 1 | 7 | |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

TPR



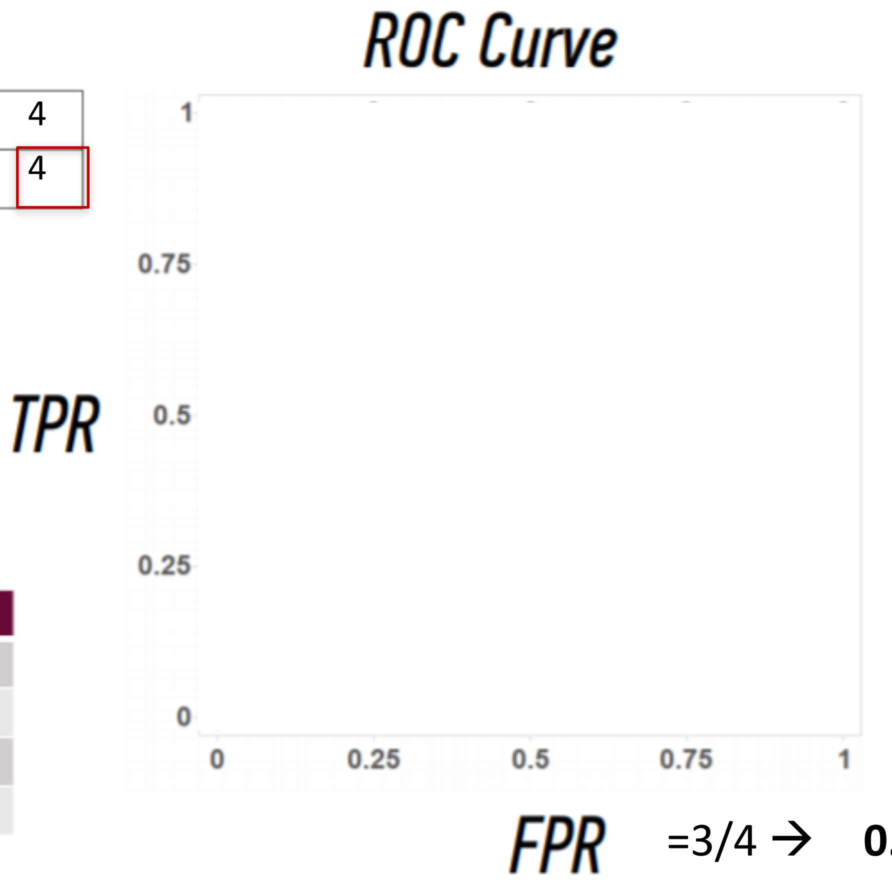

FPR =3/4 → **0.75**

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | 4 | 4 |
| | 1 | 7 | |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |




=3/4 → **0.75**

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

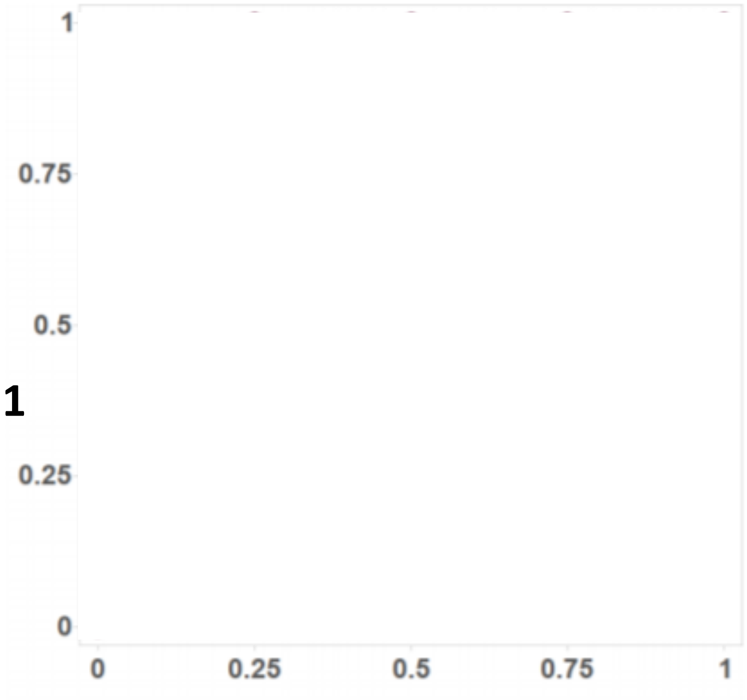
| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | 4 | 4 |
| | 1 | 7 | |

TPR

=4/4 → **1**

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

ROC Curve




FPR

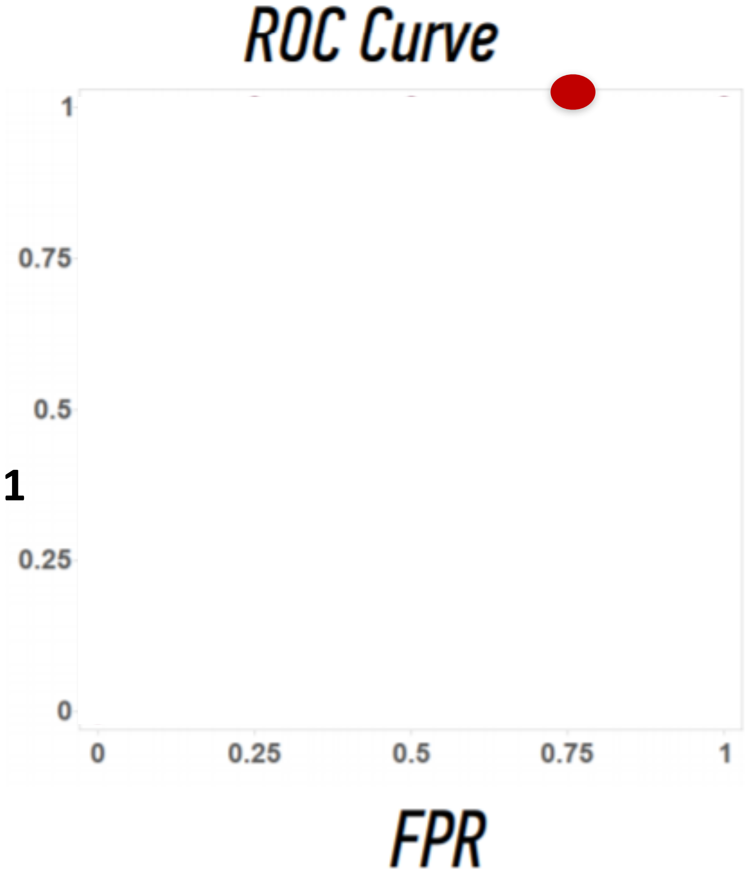
# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| | Predicted: NO | Predicted: YES | |
|---|---|---|---|
| Actual: NO | 1 | 3 | 4 |
| Actual: YES | 0 | 4 | 4 |
| | 1 | 7 | |

TPR

=4/4 → **1**

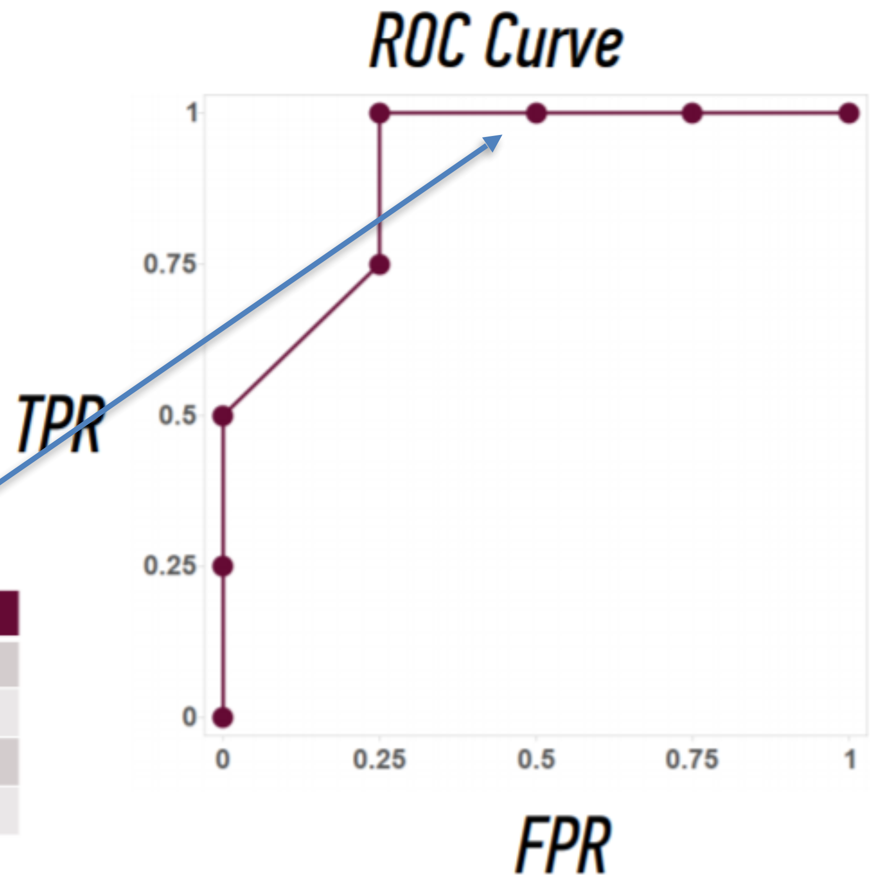



| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |

# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

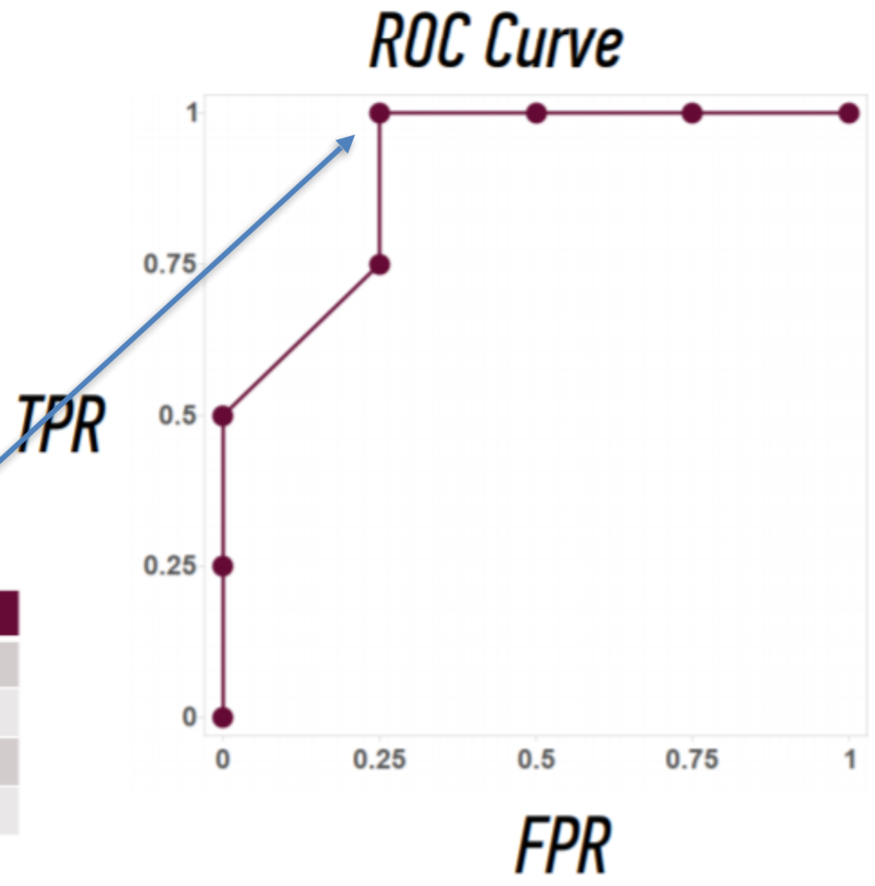
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

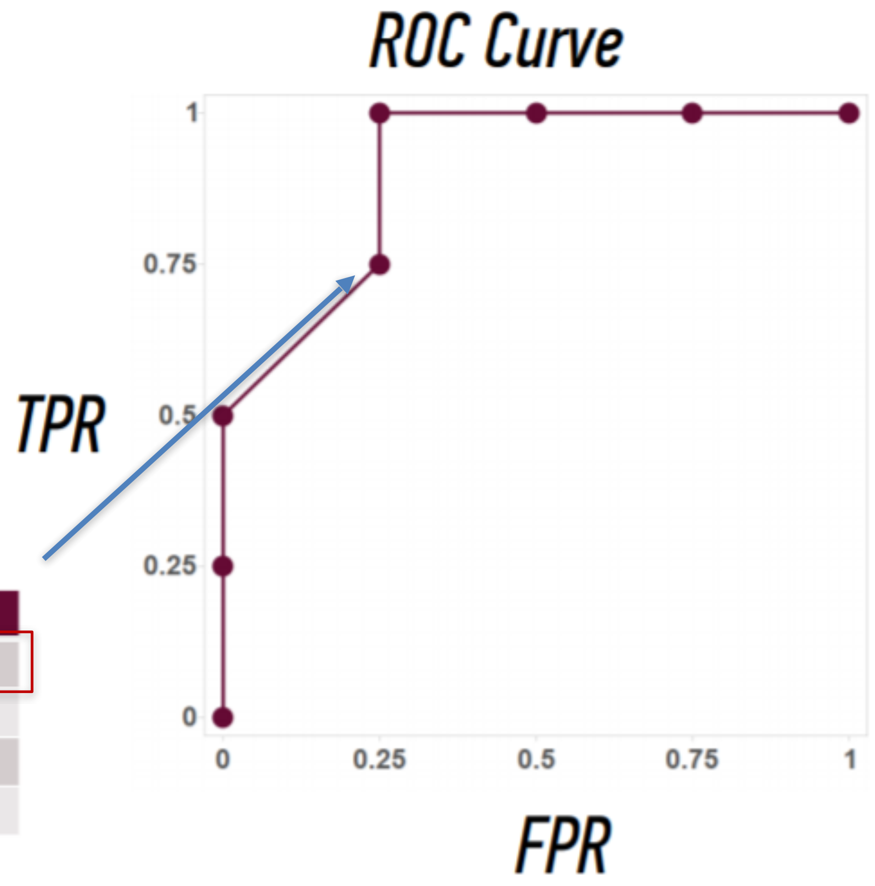
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

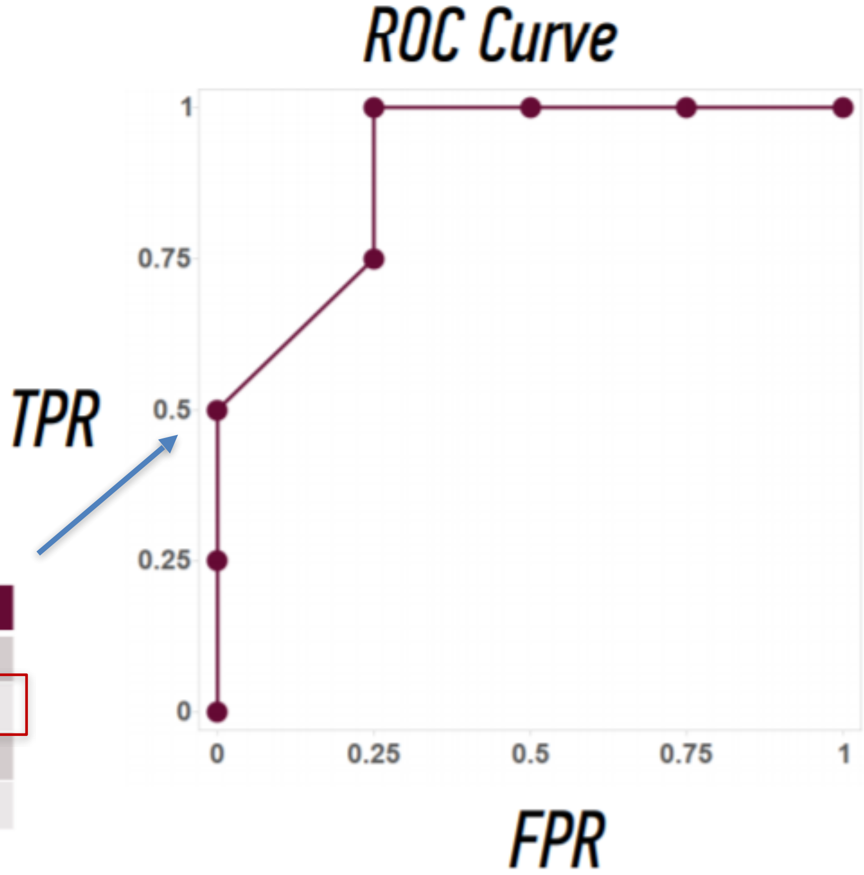
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

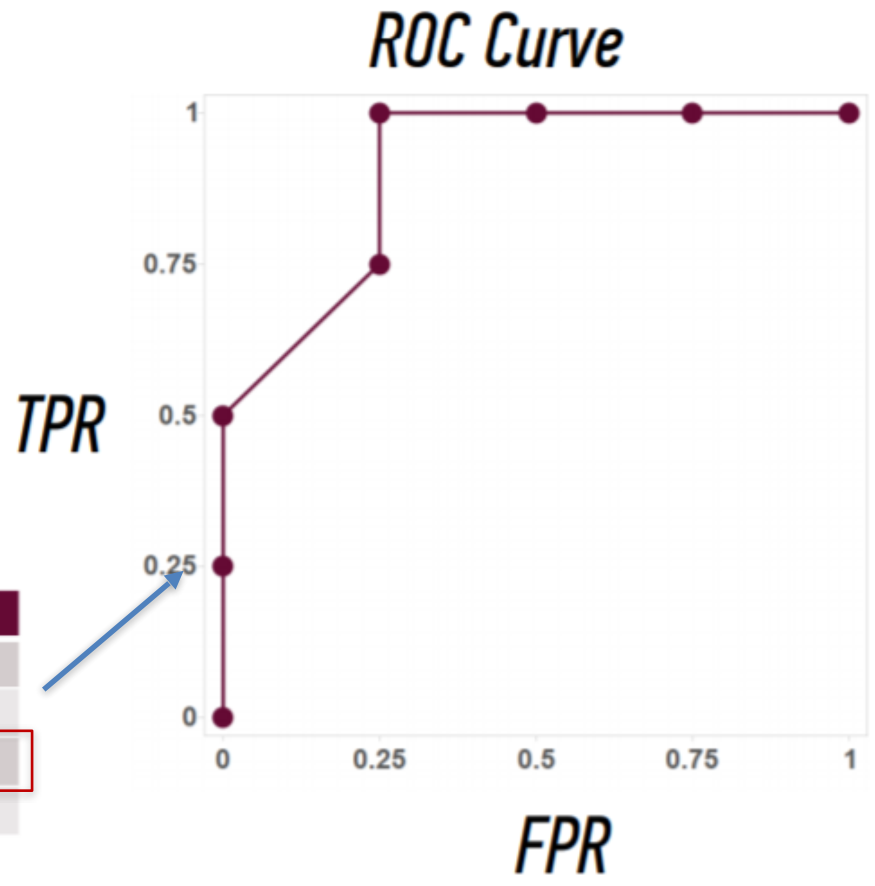
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

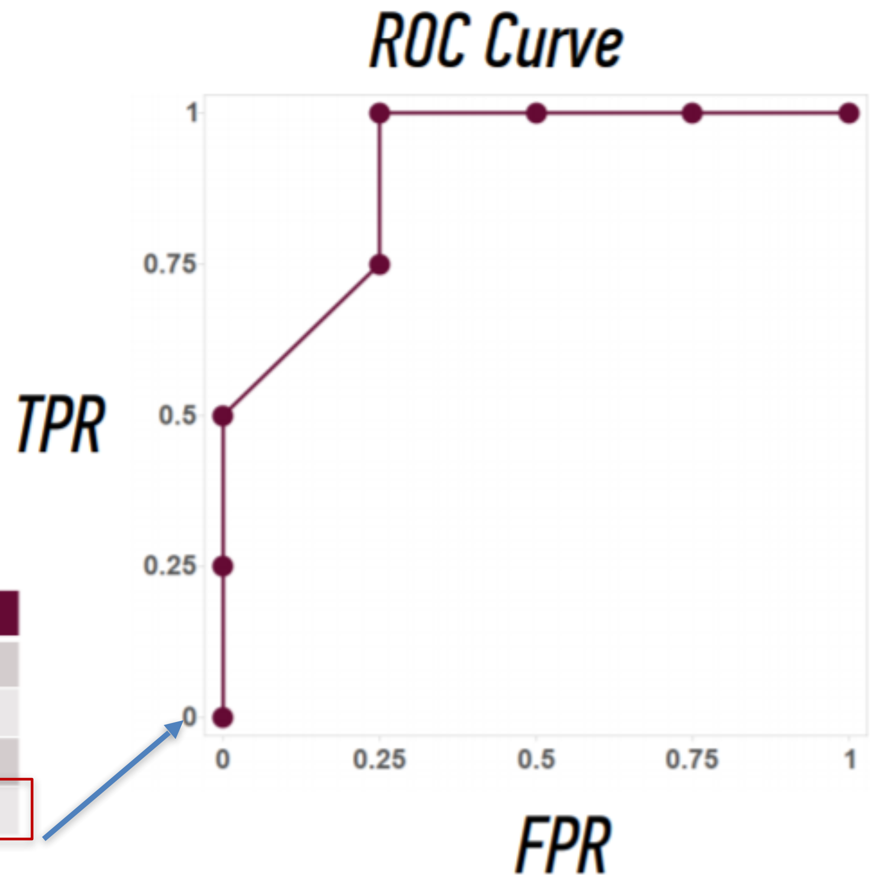
| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



# ROC CURVE / AUC

| Email Number | Score | True Label |
|---|---|---|
| 5 | 0.99 | Spam |
| 8 | 0.82 | Spam |
| 2 | 0.60 | Spam |
| 1 | 0.60 | Ham |
| 7 | 0.48 | Spam |
| 3 | 0.22 | Ham |
| 4 | 0.10 | Ham |
| 6 | 0.02 | Ham |

| Cutoff | TPR (y) | FPR (x) | Cutoff | TPR (y) | FPR (x) |
|---|---|---|---|---|---|
| **0** | 1 | 1 | **0.50** | 0.75 | 0.25 |
| **0.05** | 1 | 0.75 | **0.65** | 0.5 | 0 |
| **0.15** | 1 | 0.5 | **0.85** | 0.25 | 0 |
| **0.25** | 1 | 0.25 | **1** | 0 | 0 |



## GUIDED PRACTICE

# WHICH METRIC MATTERS?

# ACTIVITY: WHICH METRIC MATTERS?

EXERCISE

## DIRECTIONS (15 minutes)

While AUC seems like a “golden standard”, it could be *further* improved depending upon your problem. There will be instances where error in positive or negative matches will be very important. For each of the example on the next slide:

1. Write a confusion matrix: true positive, false positive, true negative, false negative. Then decide what each square represents for that specific example.

2. Define the *benefit* of a true positive and true negative and the *cost* of a false positive and false negative.

## DELIVERABLE

Answers for each example

# ACTIVITY: WHICH METRIC MATTERS?

EXERCISE

## DIRECTIONS (15 minutes)

**Examples**:

1. A test is developed for determining if a patient has cancer or not.
2. A newspaper company is targeting a marketing campaign for "at risk" users that may stop paying for the product soon.
3. You build a spam classifier for your email system.

## DELIVERABLE

Answers for each example

# CONFUSION VS ROC?

Discuss in your groups:

- What information do you take away from each of these evaluation techniques?
- What decisions can be made from each tool?