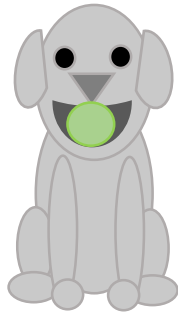


# Dogs and tennis balls

## An ROC and AUC example

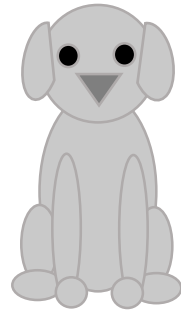
Emily Miller



### Reference

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861-874.

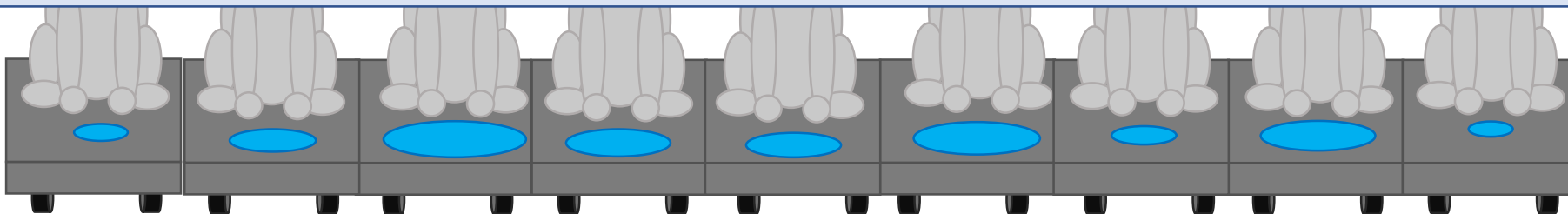
9 dogs, each on  
a rolling platform  
behind a screen



4 tennis balls



Each tennis ball is  
held by a dog; each  
dog can hold at most  
one tennis ball.



9 dogs, each on  
a rolling platform  
behind a screen

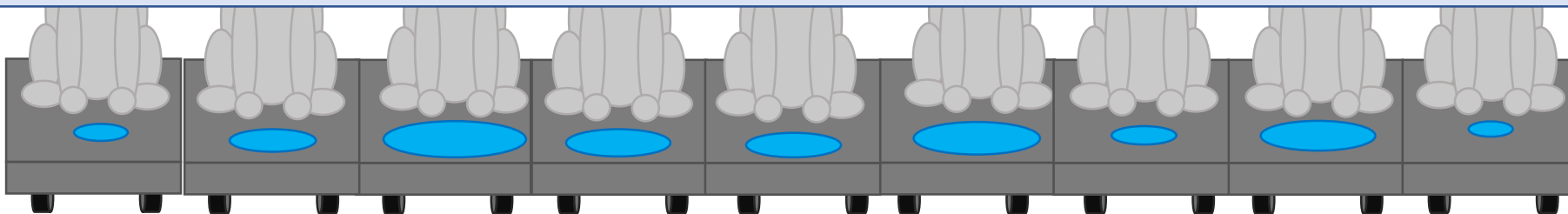


4 tennis balls



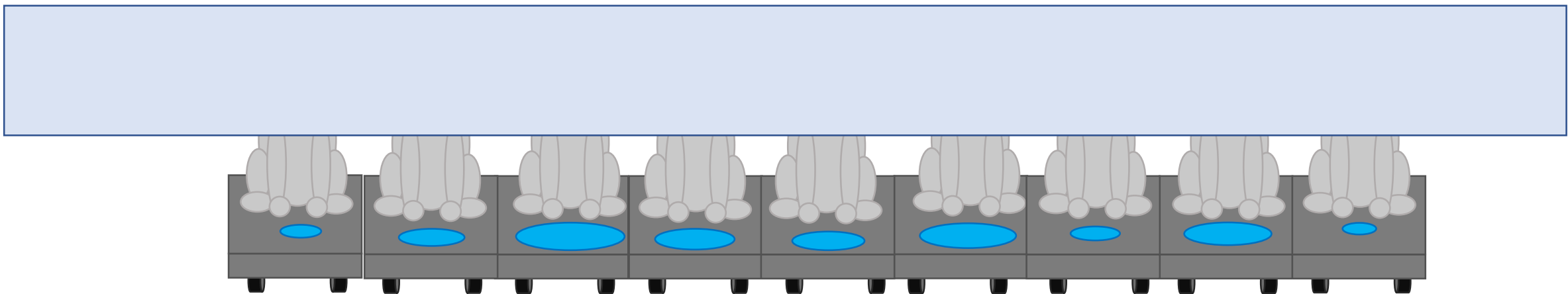
Each tennis ball is  
held by a dog; each  
dog can hold at most  
one tennis ball.

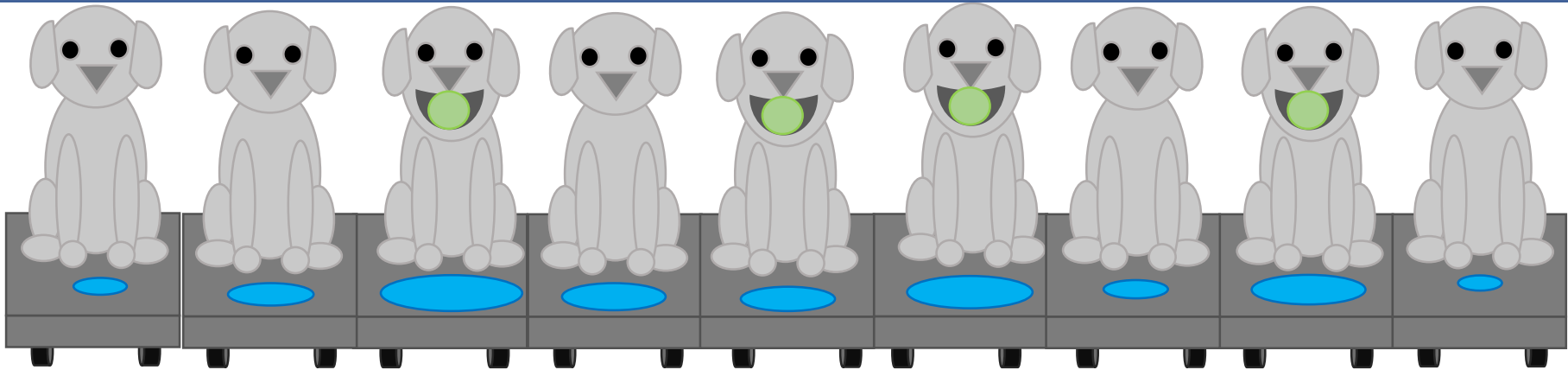
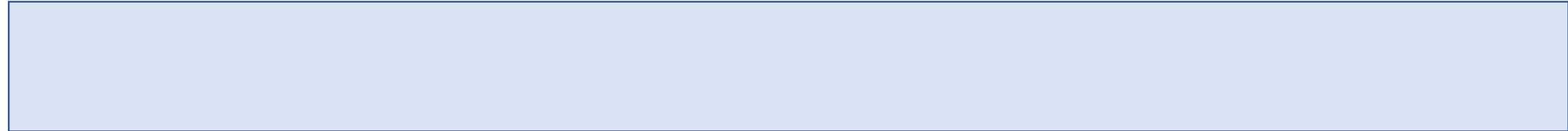
Let's consider two models that predict which dogs are holding tennis balls.



### Model 1

We suppose that the farther a dog is now sitting to the left, the more likely that dog is to have a tennis ball.

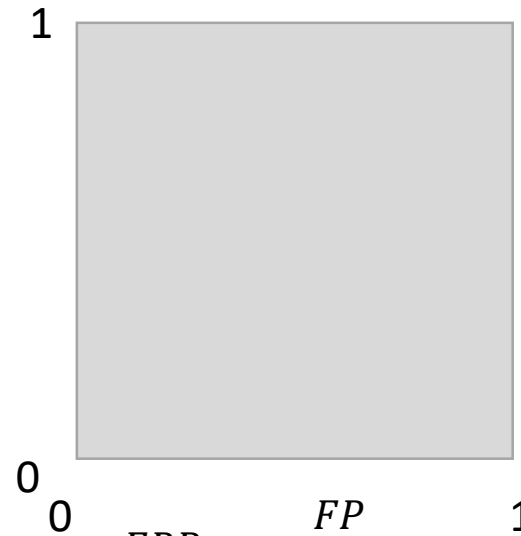




What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

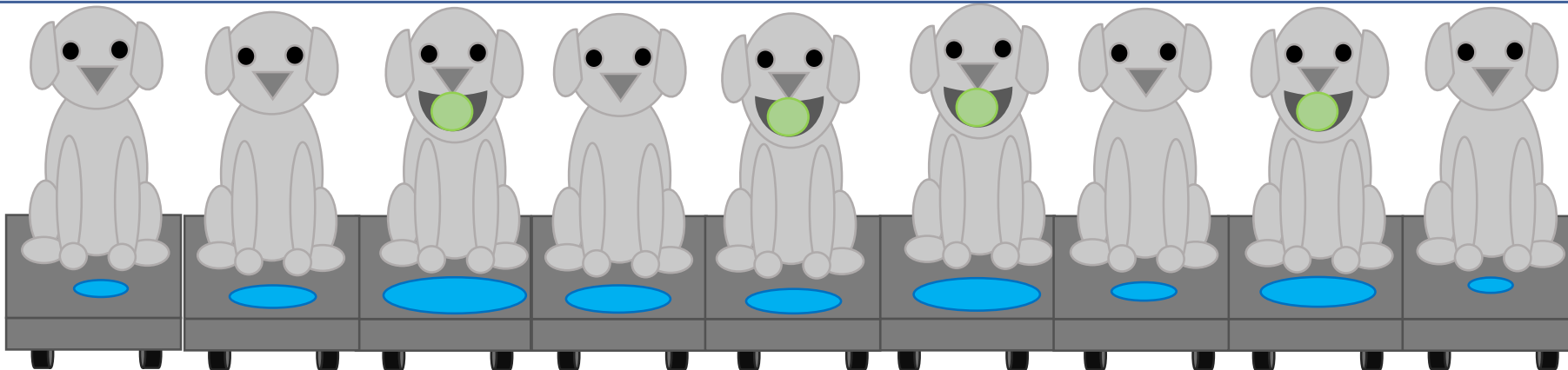
all actual positives



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

all actual negatives

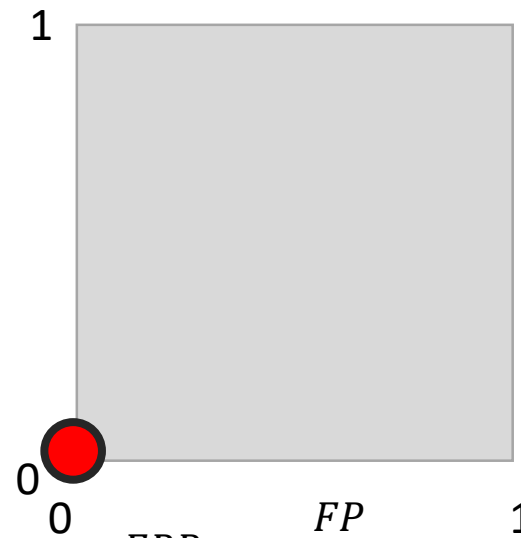
What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

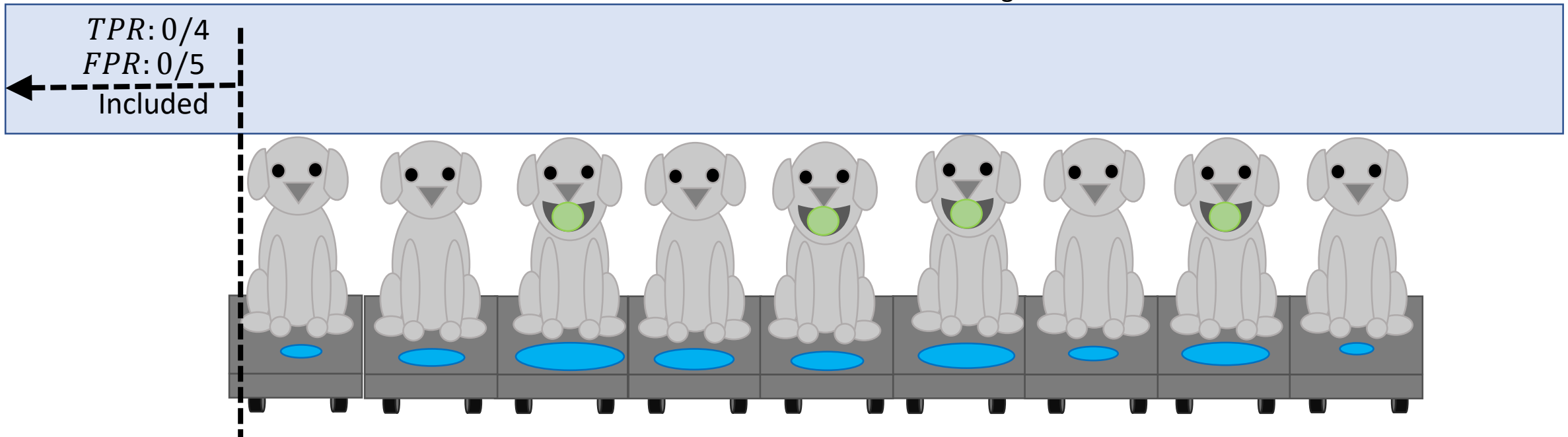
all actual positives



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

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

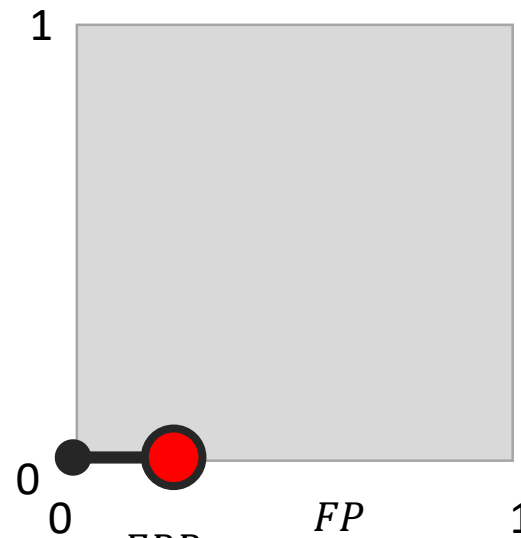




What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives



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

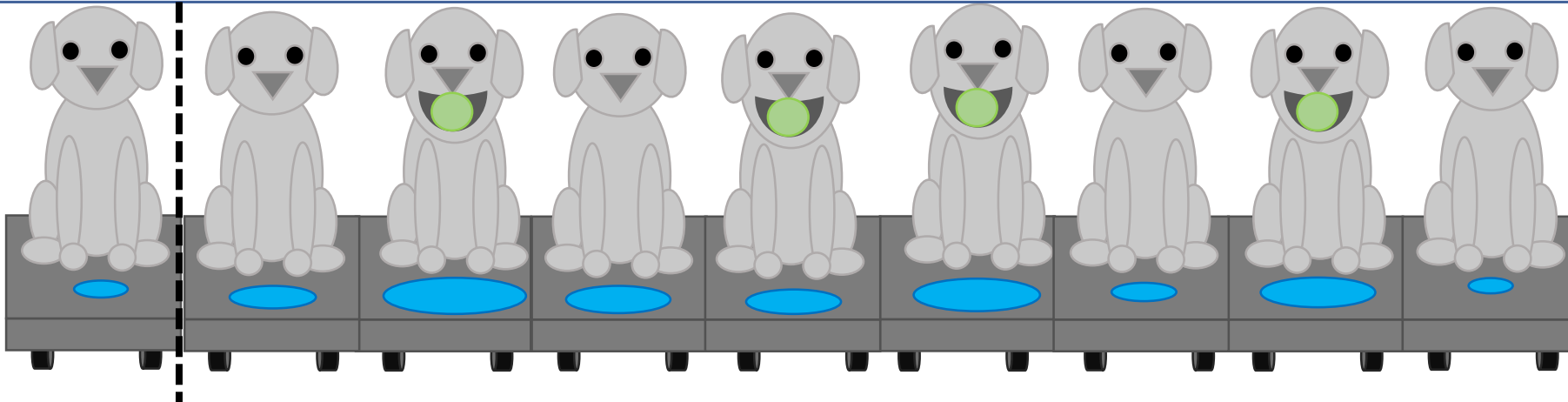
all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

TPR: 0/4

FPR: 1/5

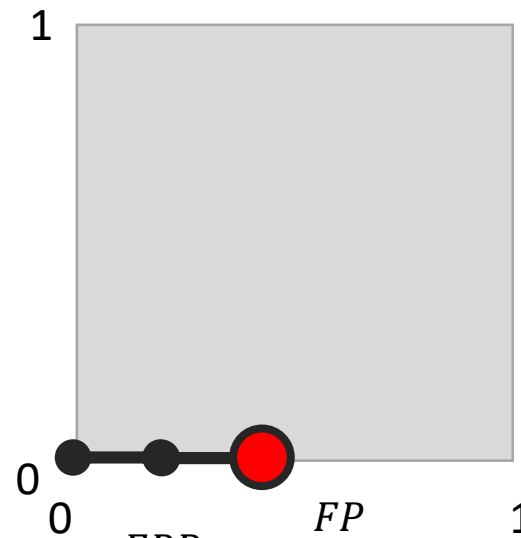
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

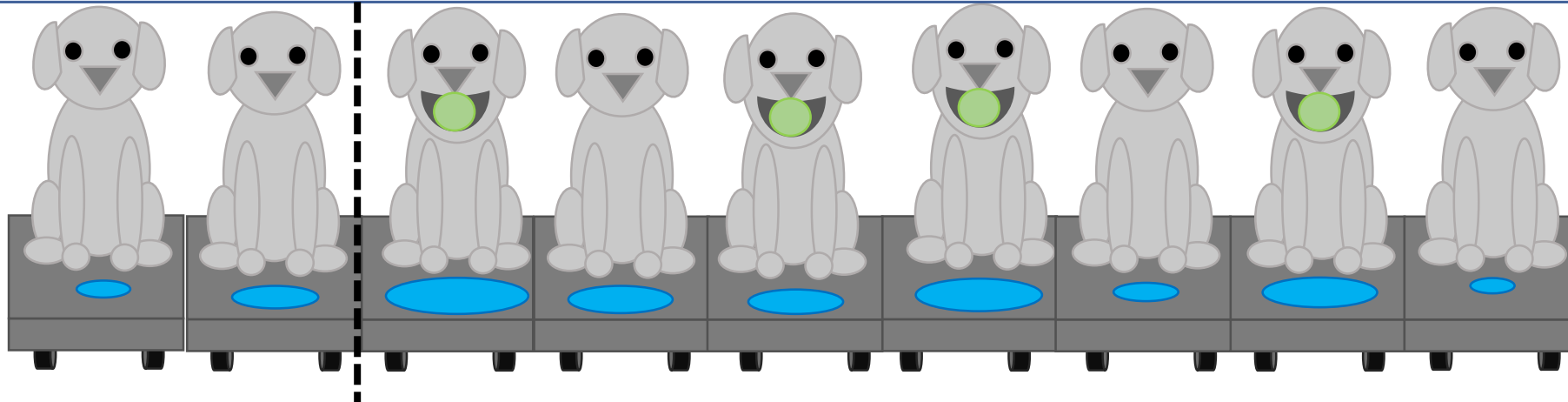
$$TPR = \frac{TP}{TP + FN}$$

all actual positives



What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

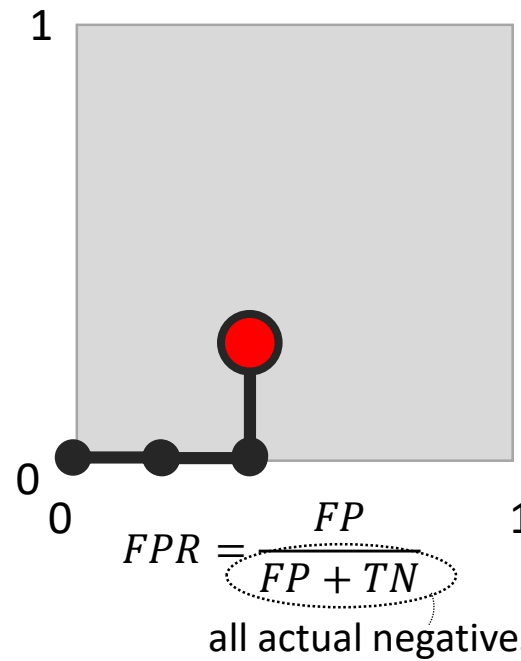
TPR: 0/4  
FPR: 2/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

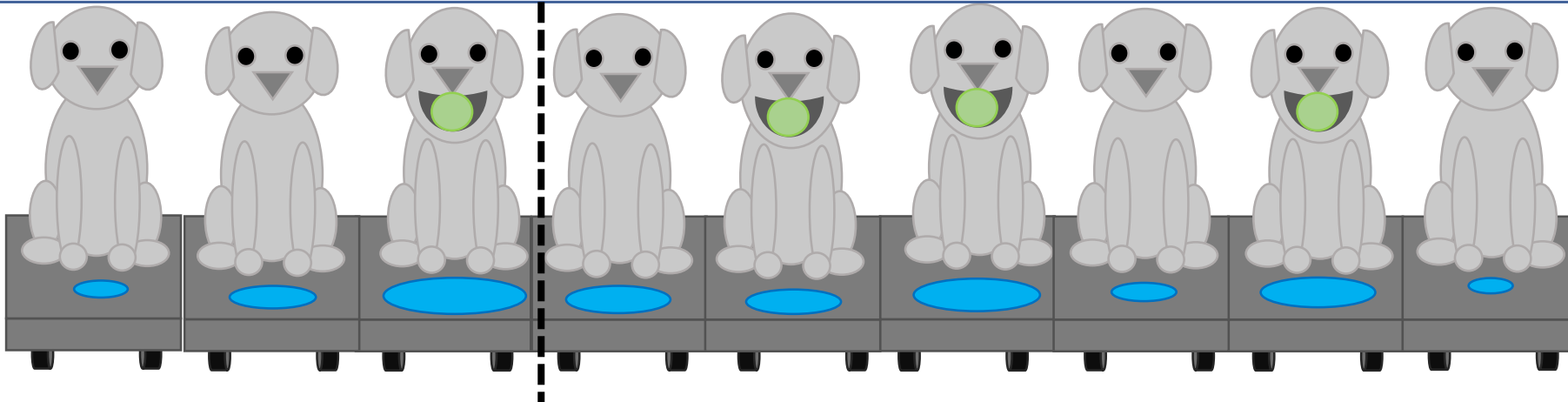


What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

TPR: 1/4

FPR: 2/5

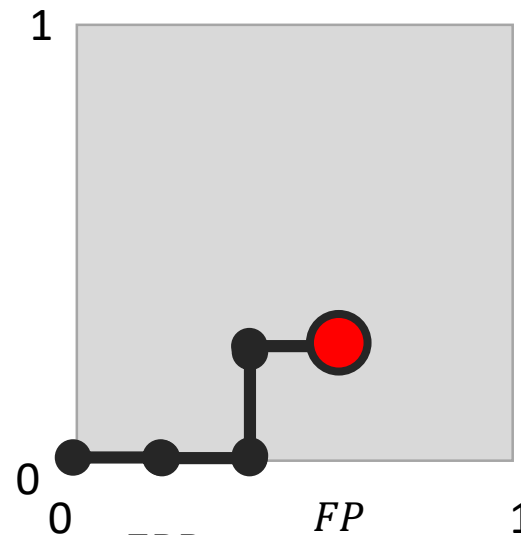
← Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

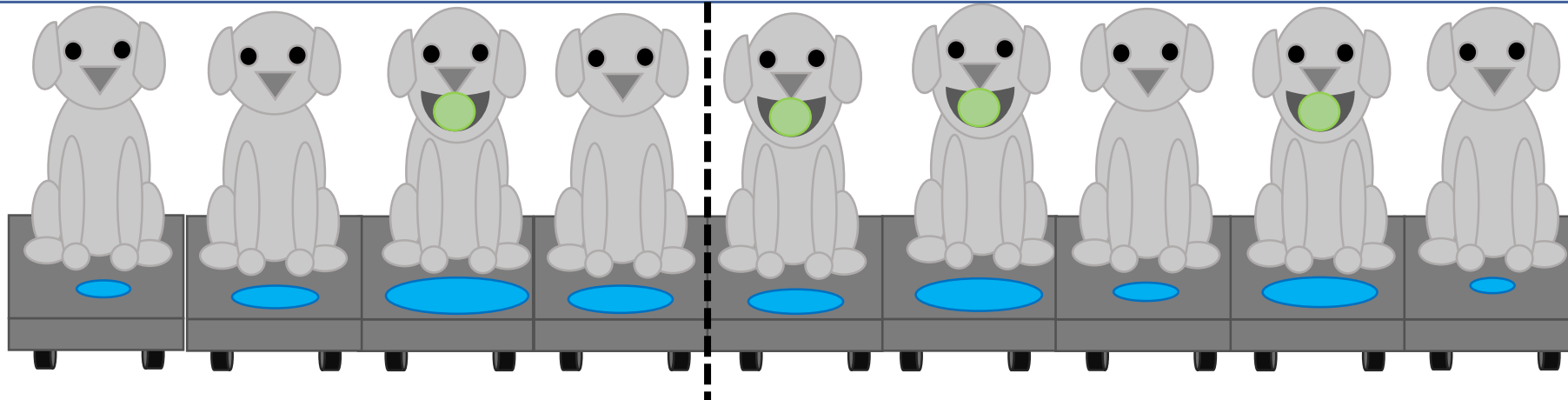


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

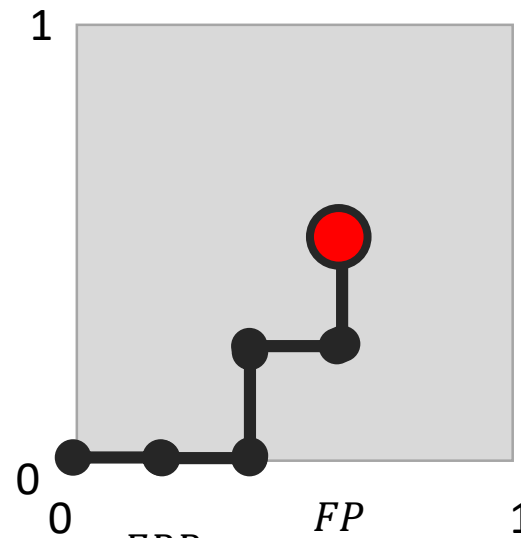
TPR: 1/4  
FPR: 3/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

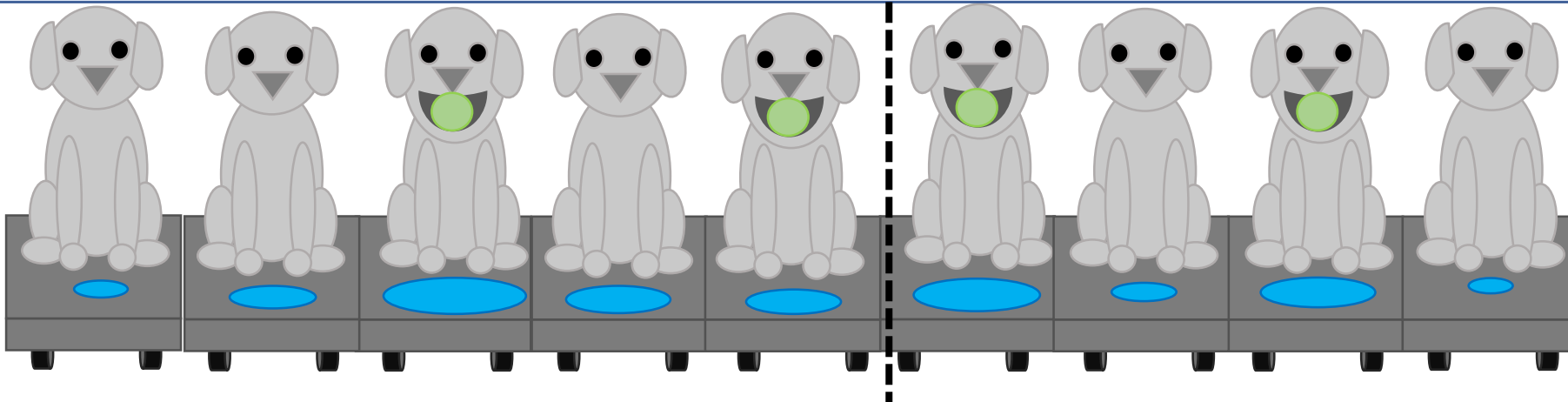


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

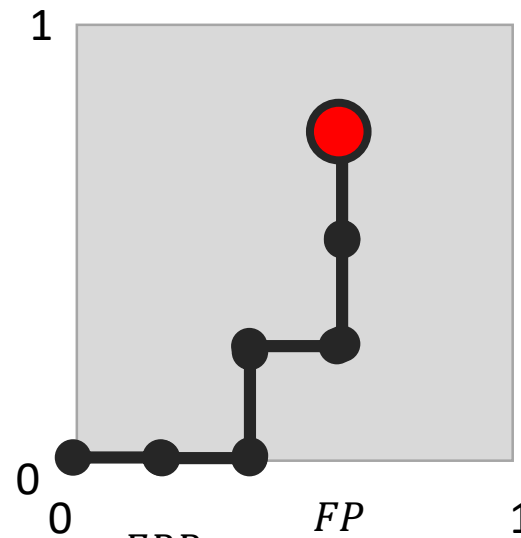
TPR: 2/4  
FPR: 3/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives



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

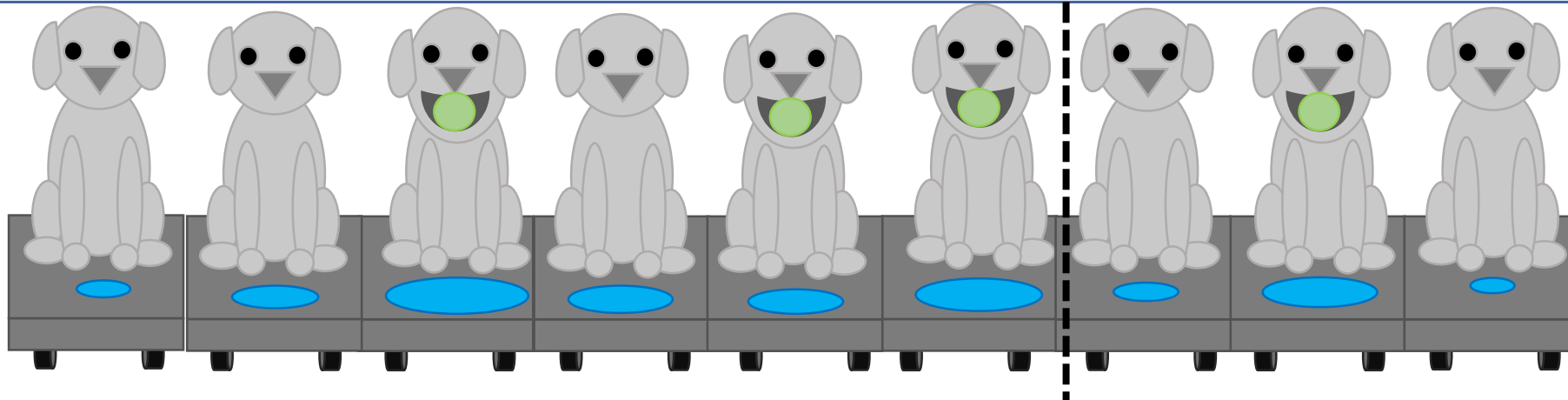
all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

TPR: 3/4

FPR: 3/5

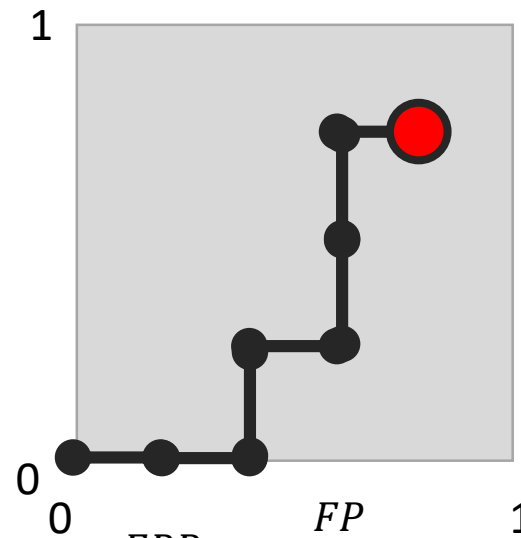
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

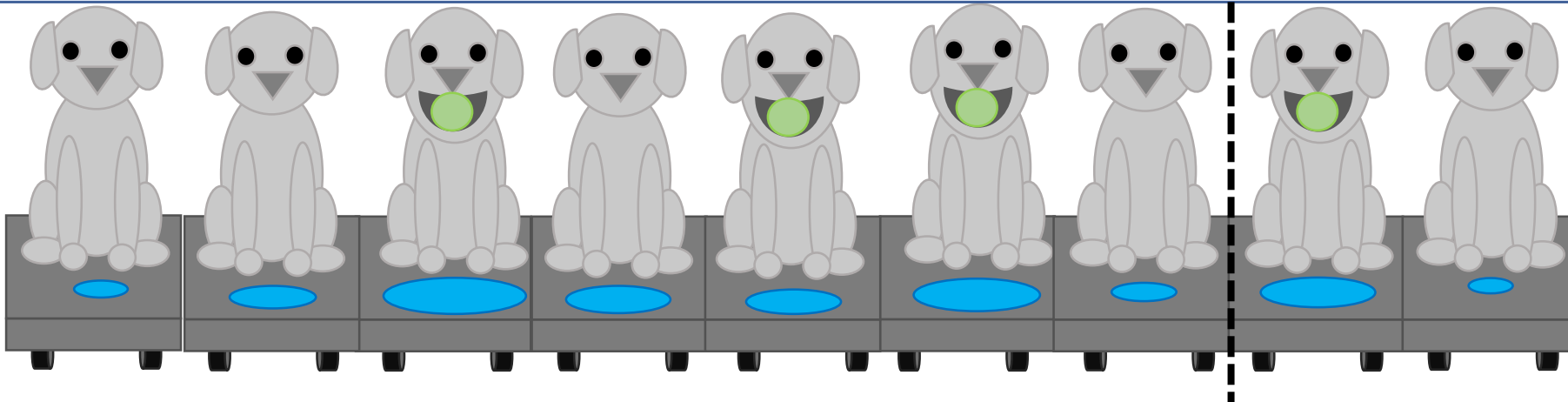


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

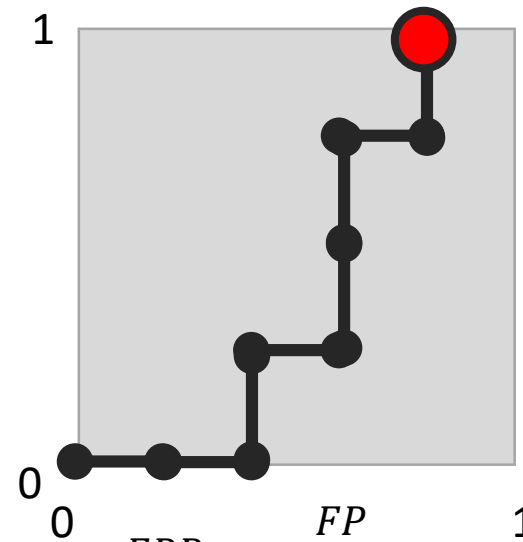
TPR: 3/4  
FPR: 4/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

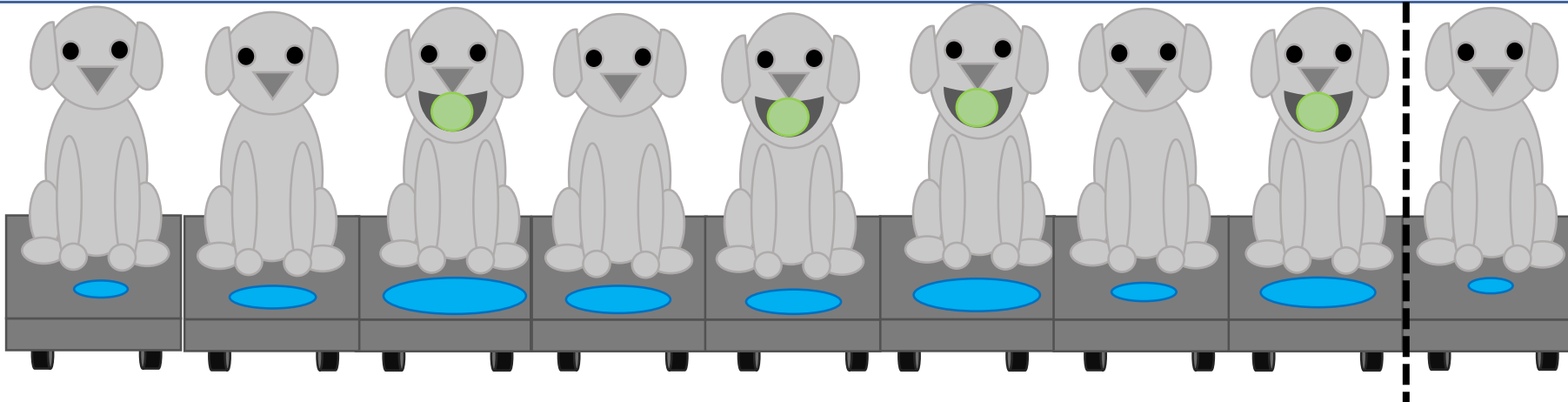


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

TPR: 4/4  
FPR: 4/5  
Included

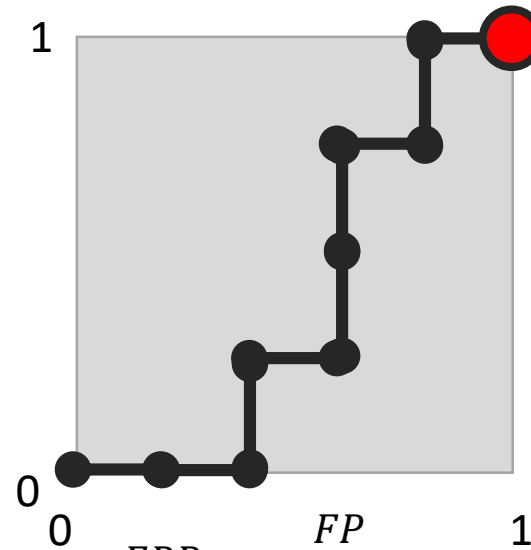




What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives



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

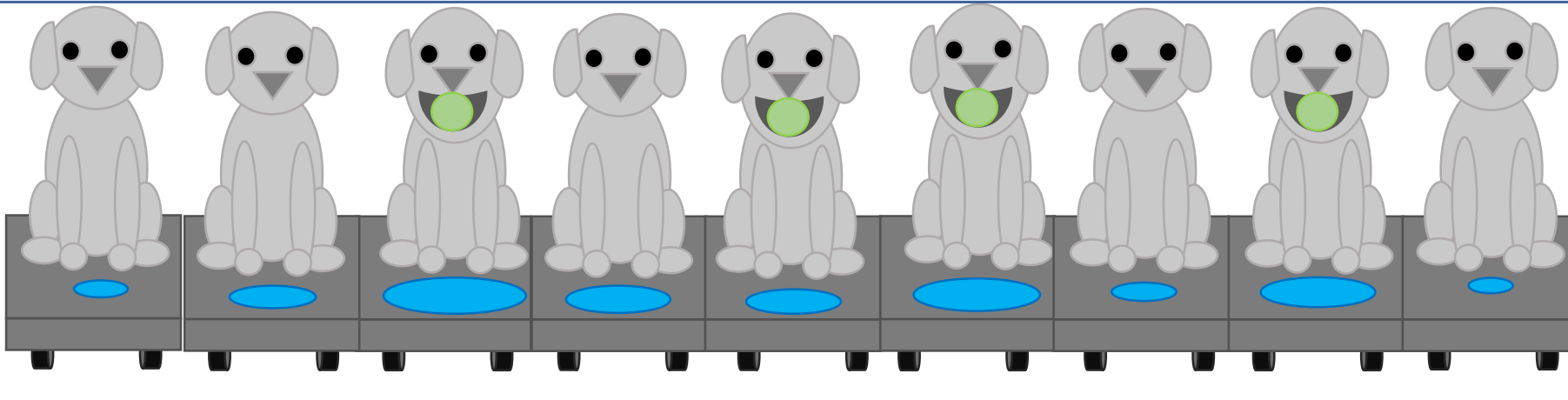
all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

TPR: 4/4

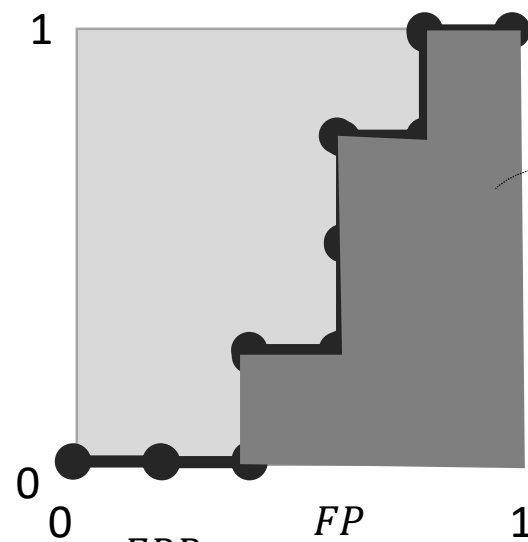
FPR: 5/5

Included



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

all actual positives

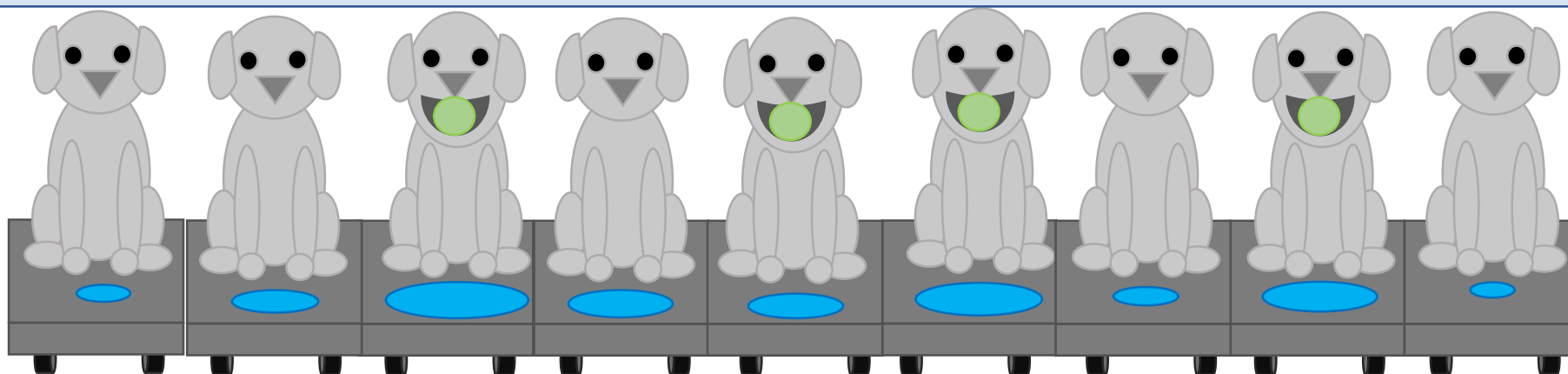


$$AUC = 8 * \left( \frac{1}{5} * \frac{1}{4} \right)$$

$$= 0.40$$

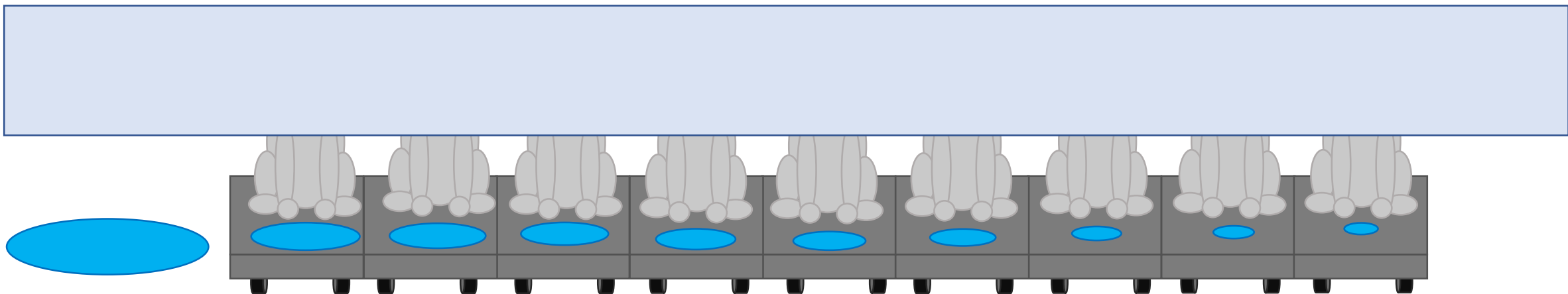
$$FPR = \frac{FP}{FP + TN}$$

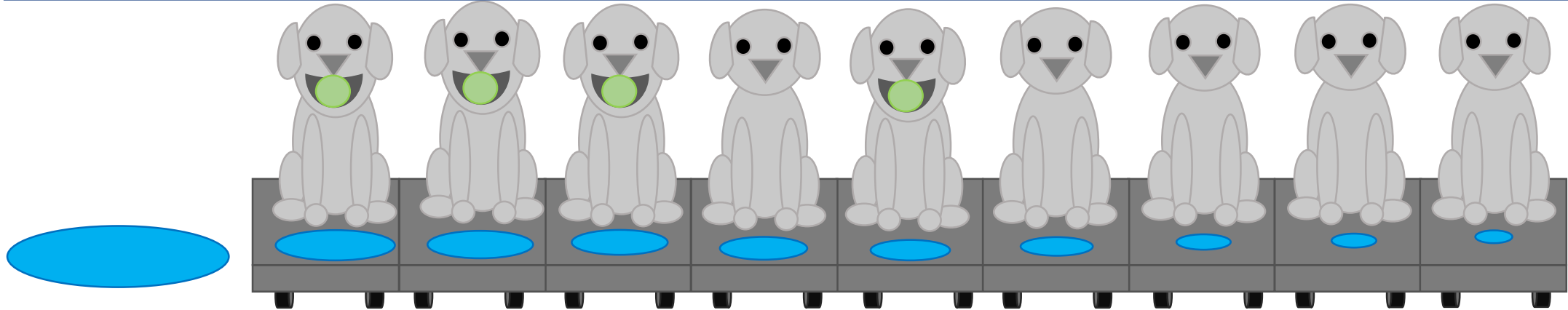
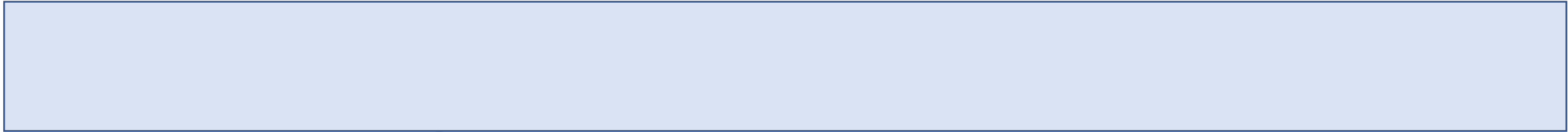
all actual negatives



### Model 2

We suppose that the more a dog has drooled,  
the more likely that dog is to have a tennis ball.

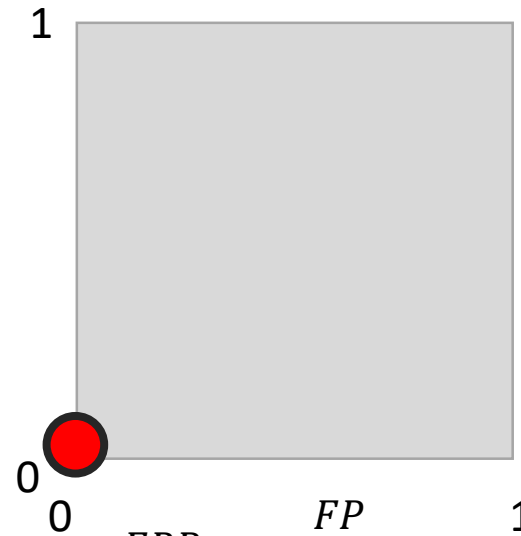




What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

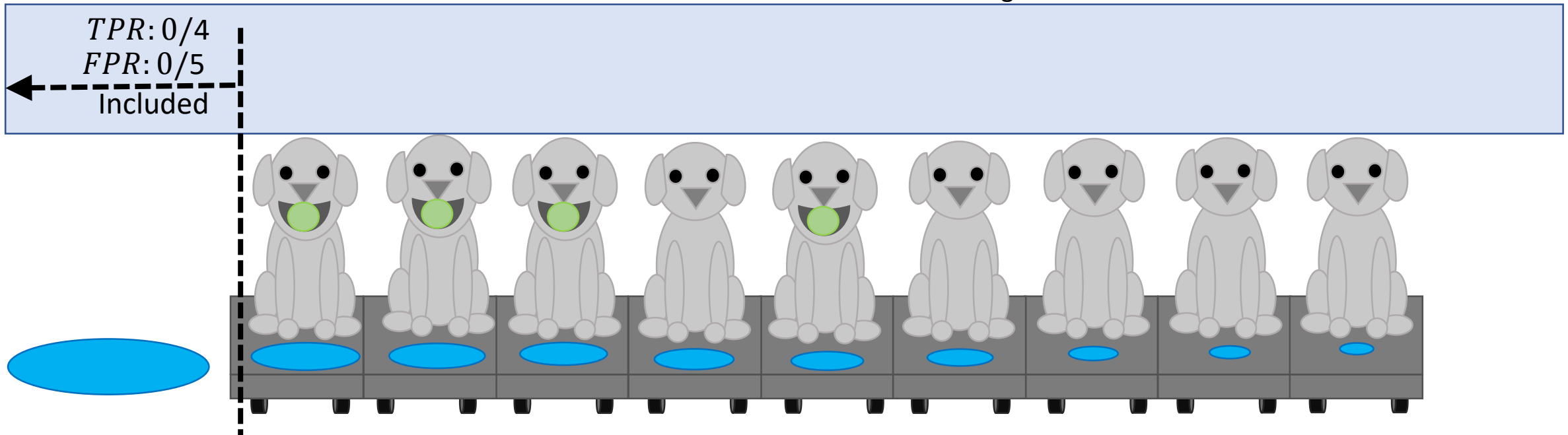
all actual positives



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

all actual negatives

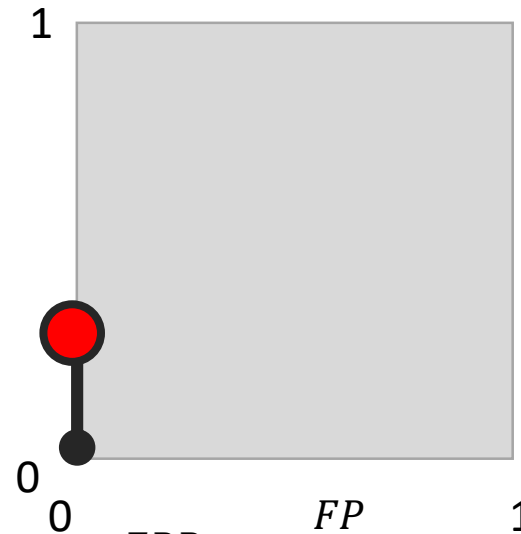
What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

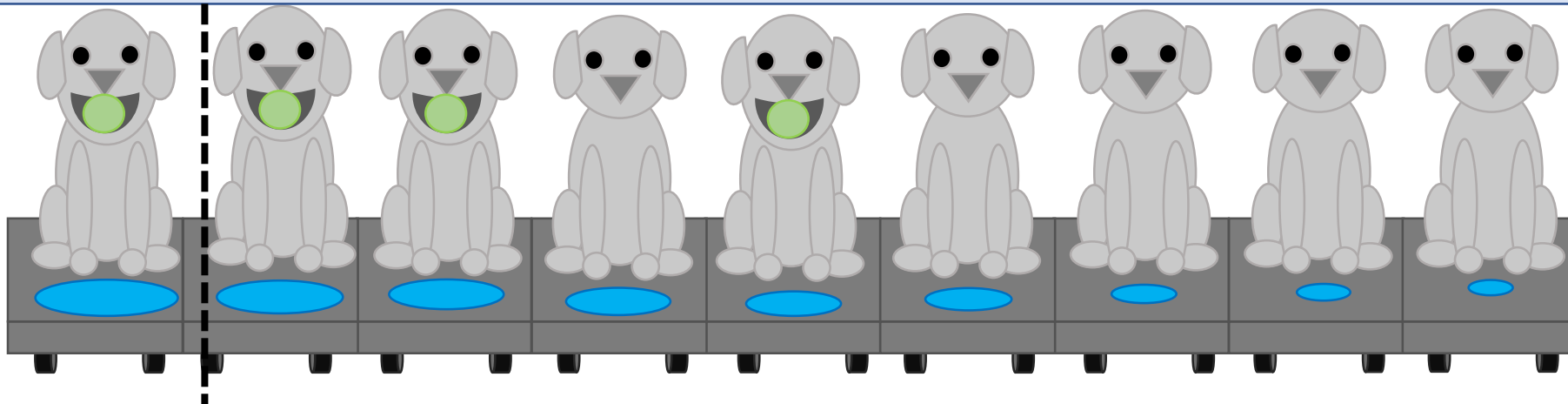


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

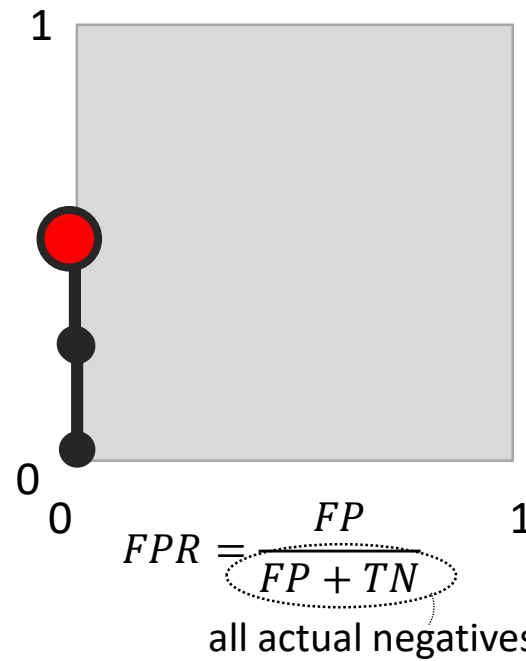
TPR: 1/4  
FPR: 0/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

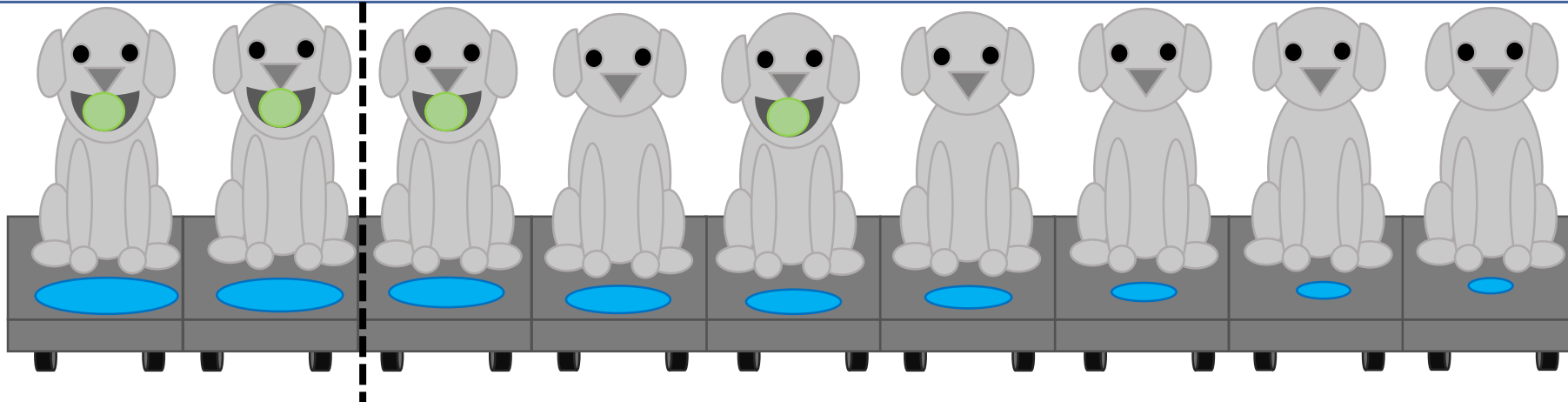
$$TPR = \frac{TP}{TP + FN}$$

all actual positives



What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

TPR: 2/4  
FPR: 0/5  
← Included

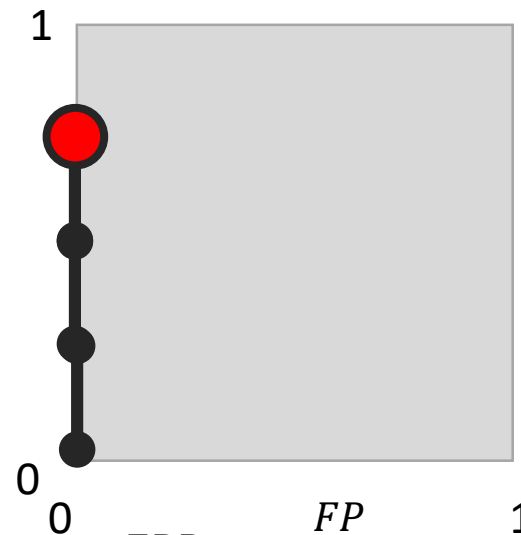




What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

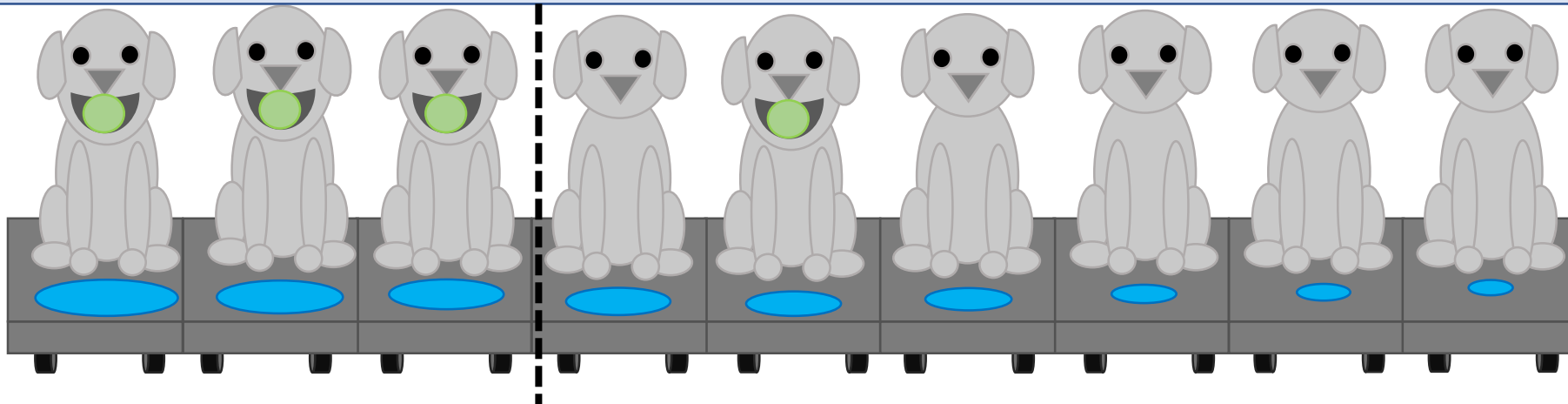


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

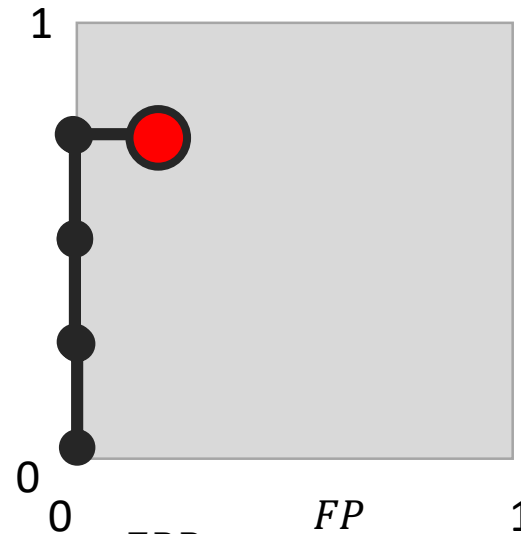
TPR: 3/4  
FPR: 0/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

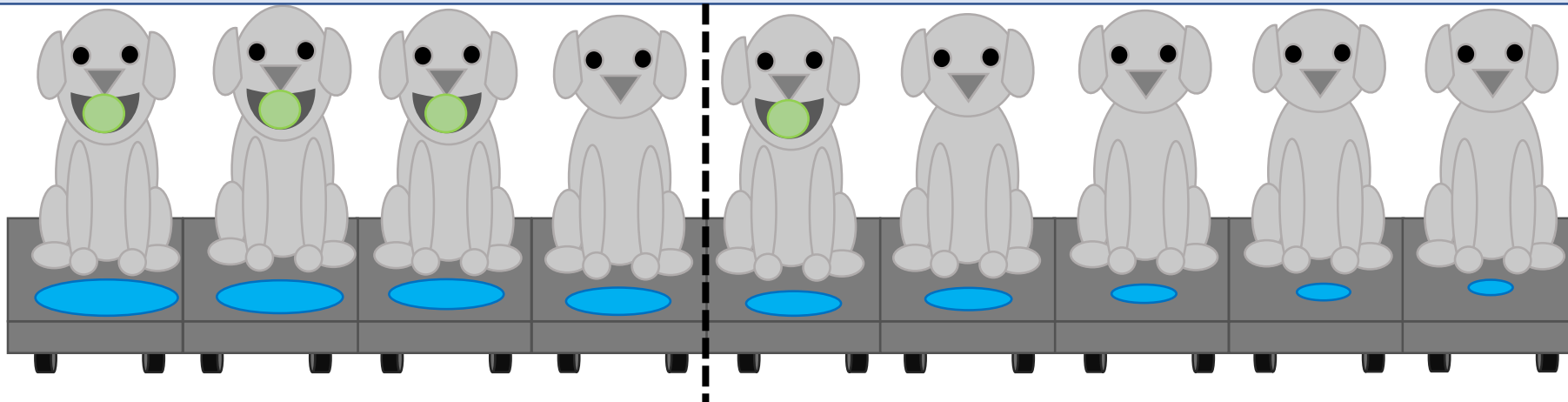


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

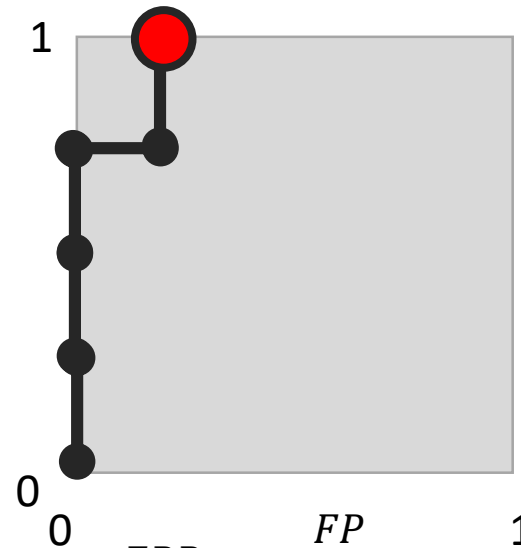
TPR: 3/4  
FPR: 1/5  
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives

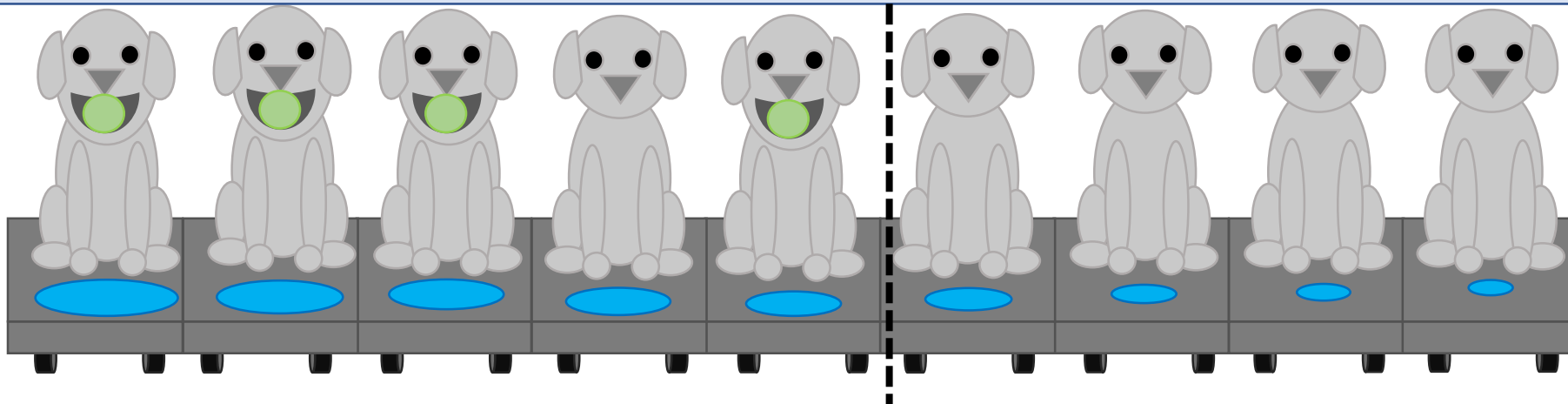


$$FPR = \frac{FP}{FP + TN}$$

all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

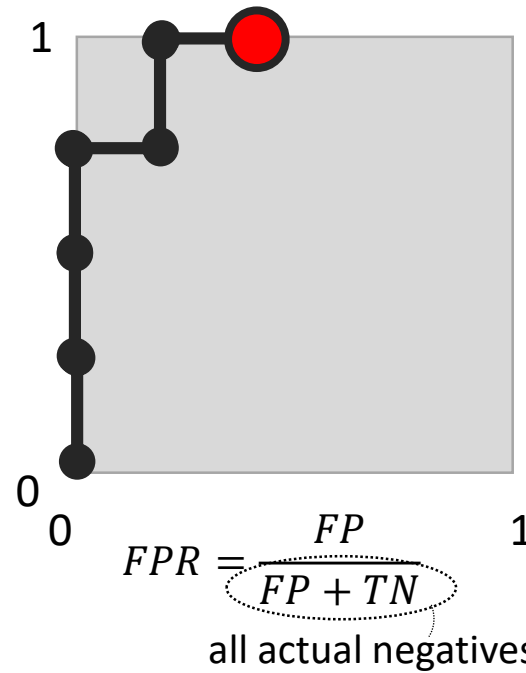
TPR: 4/4  
FPR: 1/5  
Included



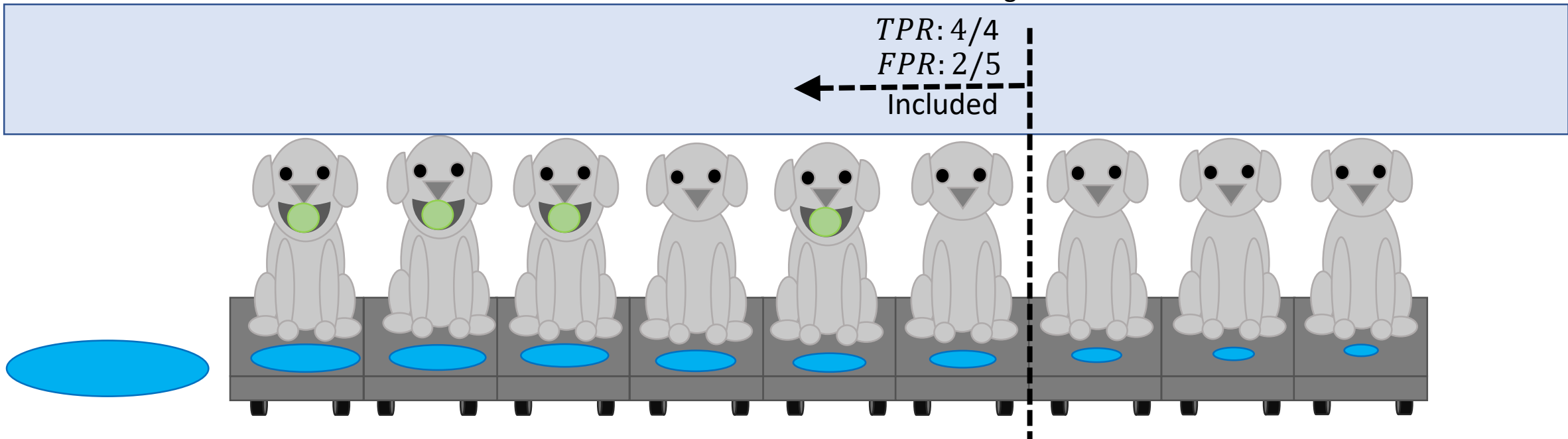
What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives



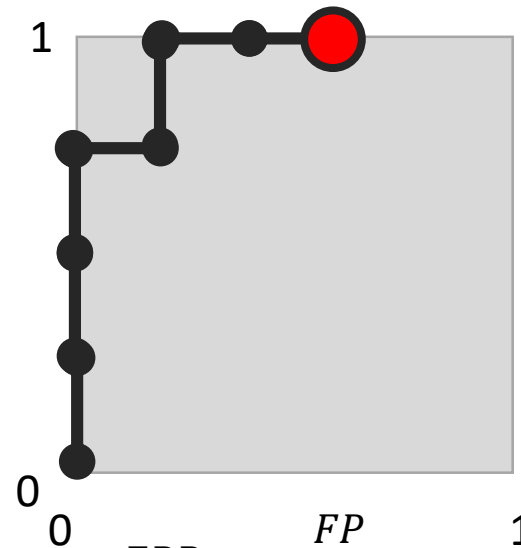
What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

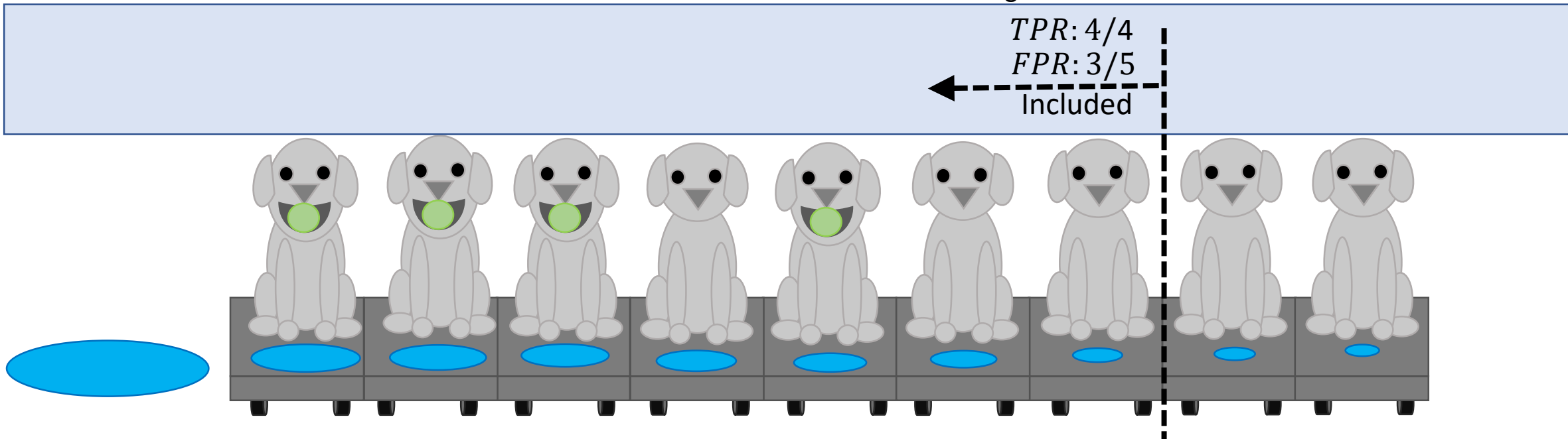
all actual positives



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

all actual negatives

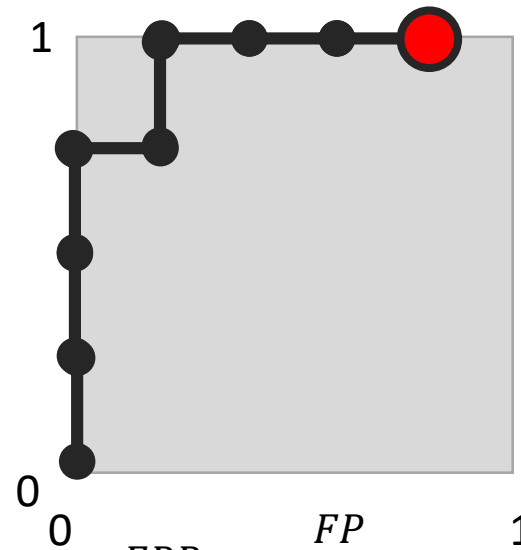
What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives



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

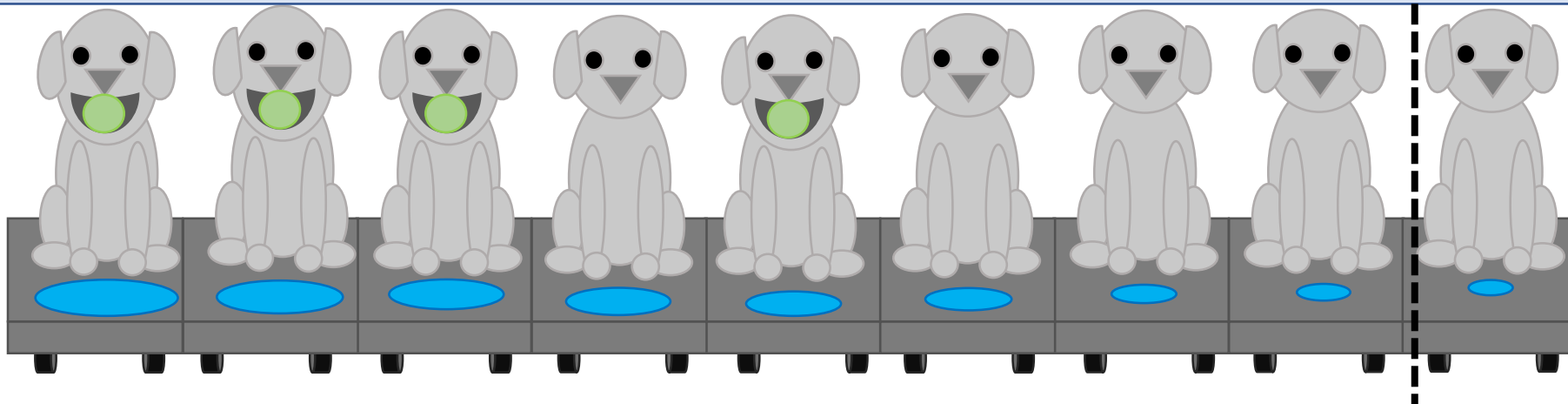
all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

$TPR: 4/4$

$FPR: 4/5$

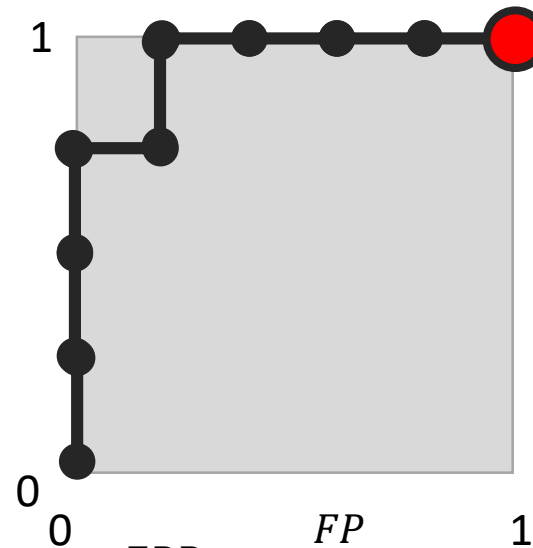
Included



What fraction of the dogs with tennis balls we've included in the set of dogs we've predicted to have tennis balls.

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

all actual positives



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

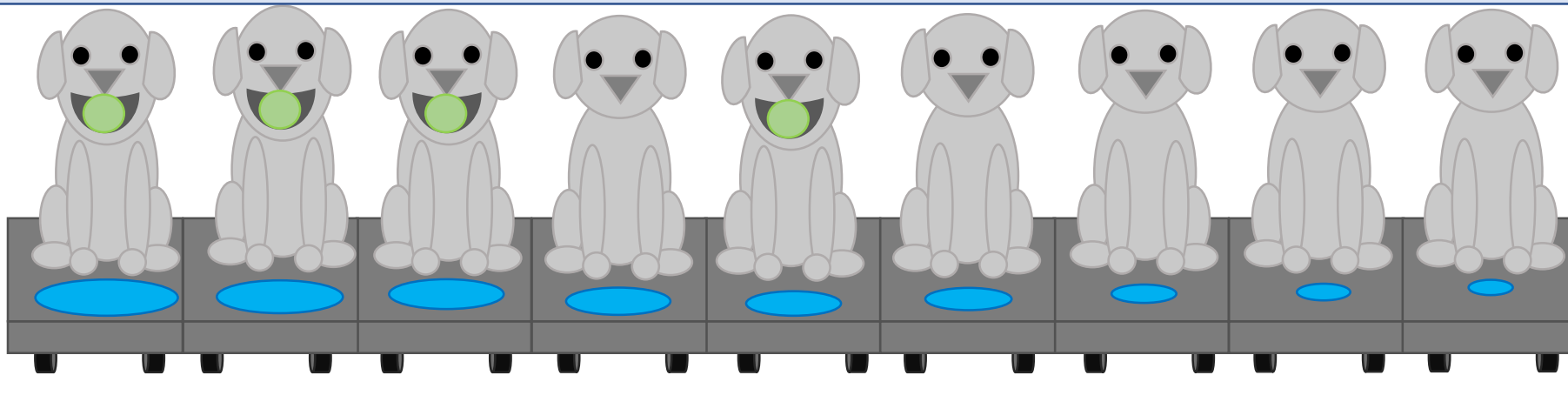
all actual negatives

What fraction of the dogs without tennis balls we've included in the set of dogs we've predicted to have tennis balls.

$TPR: 4/4$

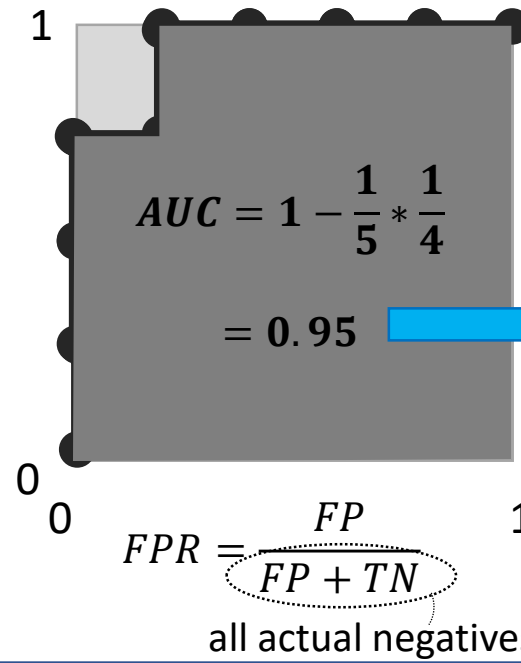
$FPR: 5/5$

Included



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

all actual positives



Model 1's AUC: 0.40

Model 2's AUC: 0.95

Drool pool size is a much better classifier here.

