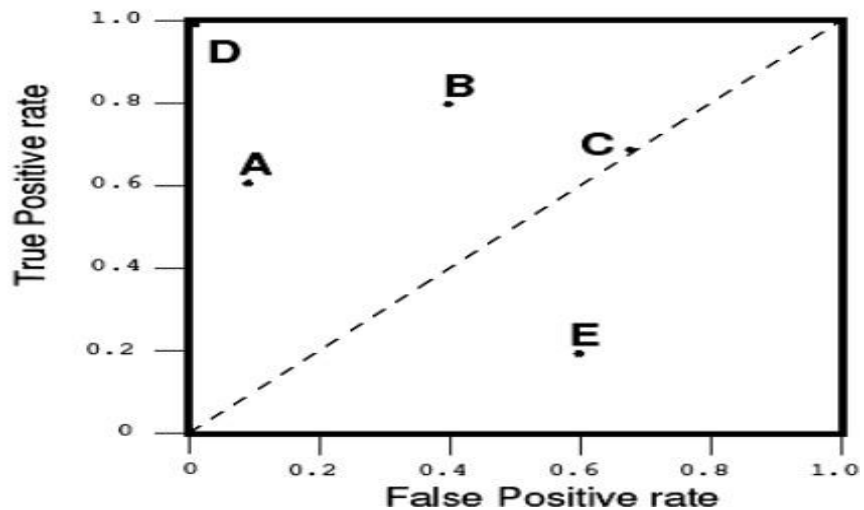
Model Evaluation

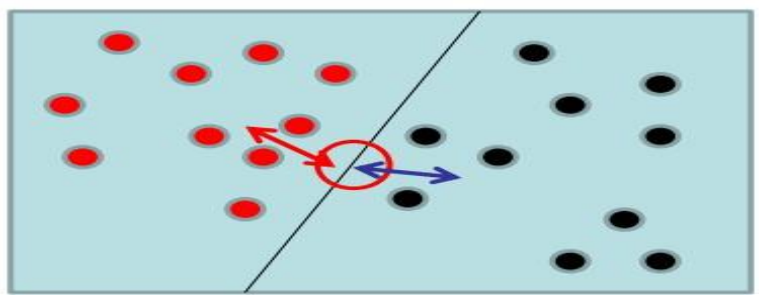
ROC Space

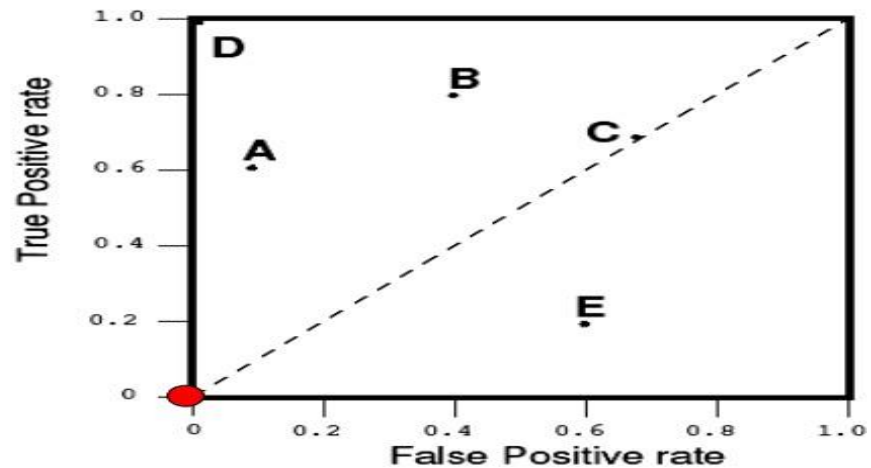
- Figure shows an ROC graph with five classifiers labeled A through E.
- A discrete classier is one that outputs only a class label.
- Each discrete classier produces a confusion matrix (fp rate, tp rate pair) corresponding to a single point in ROC space.
- Classifiers in figure are all discrete classifiers.



Special Points in ROC Space

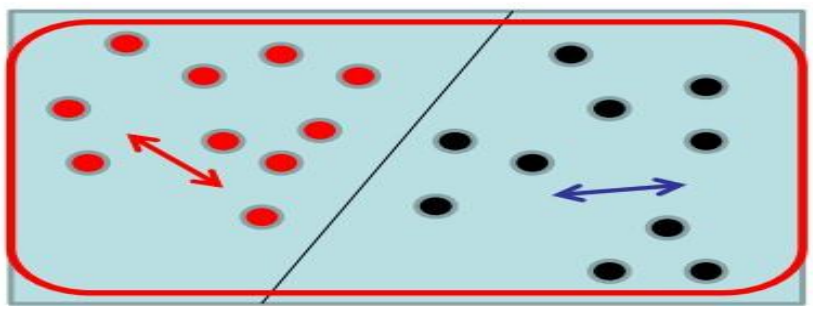
- Lower left point (0, 0) represents the strategy of never issuing a positive classification;
  - such a classier commits no false positive errors but also gains no true positives.

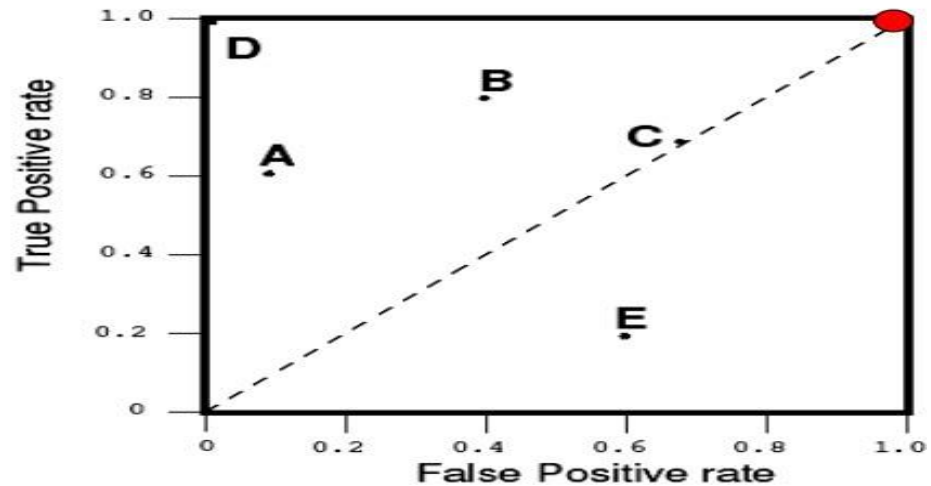




Special Points in ROC Space

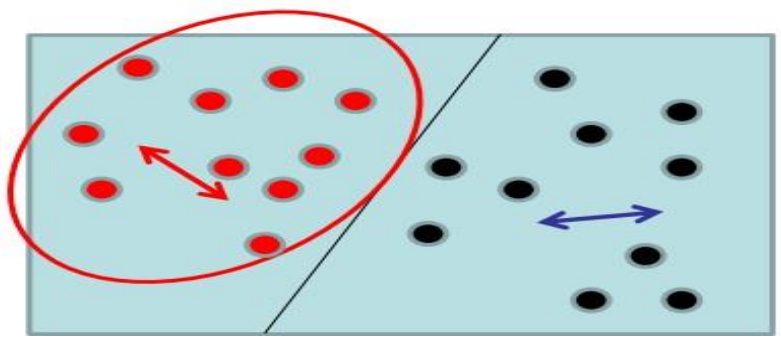
- Upper right corner (1, 1) represents the opposite strategy, of unconditionally issuing positive classifications.

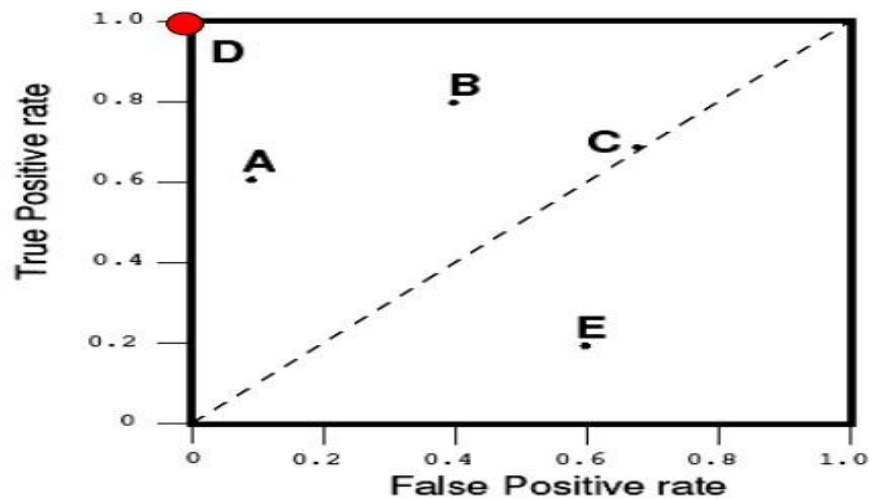




Special Points in ROC Space

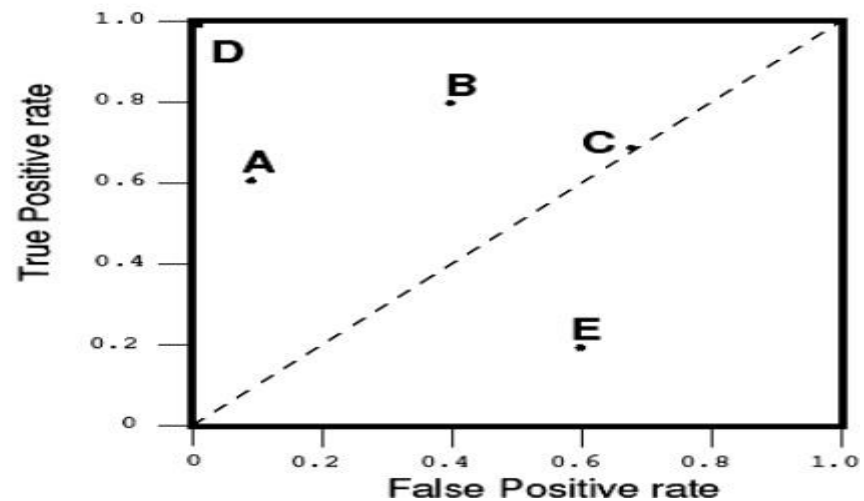
- Point (0, 1) represents perfect classification.
  - D's performance is perfect as shown.





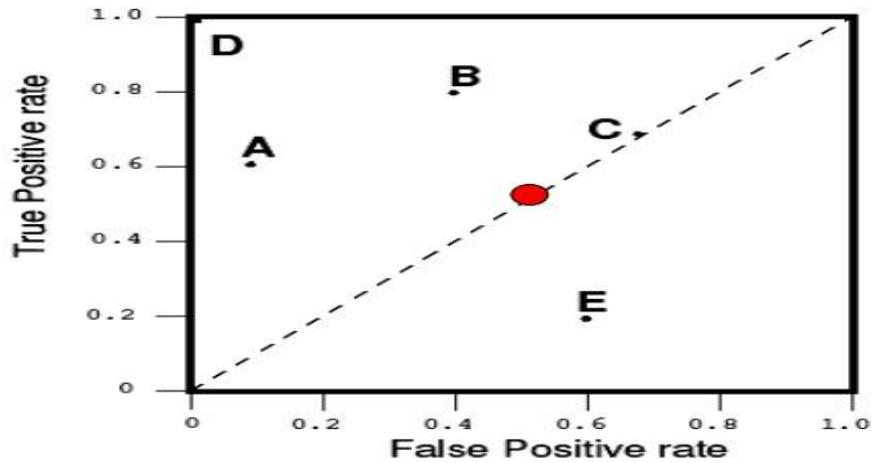
Point in ROC space

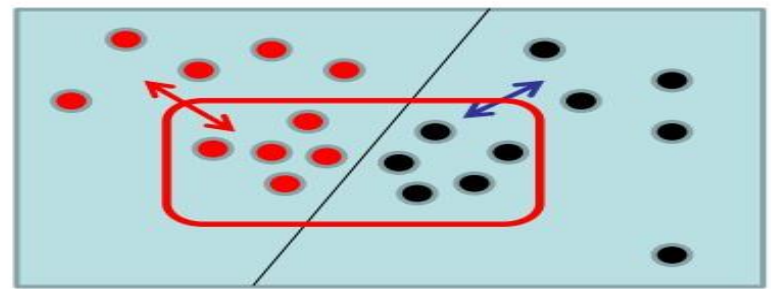
- Informally, one point in ROC space is better than another if it is to the northwest of the first
  - tp rate is higher, fp rate is lower, or both.



Random Classifiers

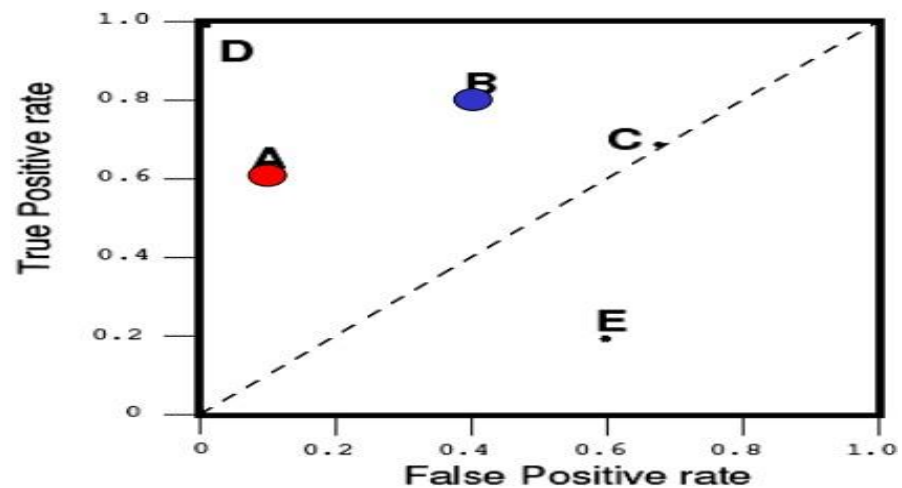
- The diagonal line y = x represents the strategy of randomly guessing a class.
- For example, if a classier randomly says “Positive” half the time (regardless of the instance provided), it can be expected to get half the positives and half the negatives correct;
  - this yields the point (0.5; 0.5)





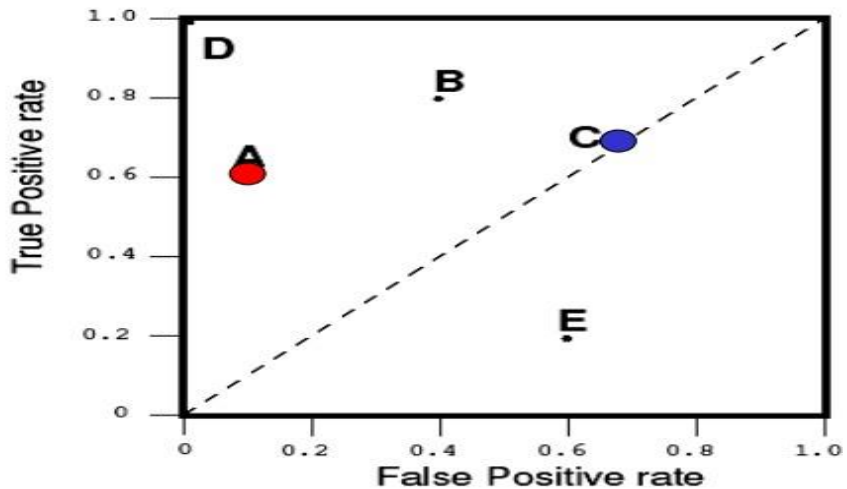
“Conservative” vs. “Liberal”

- Classifiers appearing on the left hand-side of an ROC graph, near the Y axis, may be thought of as “conservative”
  - they make positive classifications only with strong evidence so they make few false positive errors,
  - but they often have low true positive rates as well.
- In figure, A is more conservative than B.



“Conservative” vs. “Liberal”

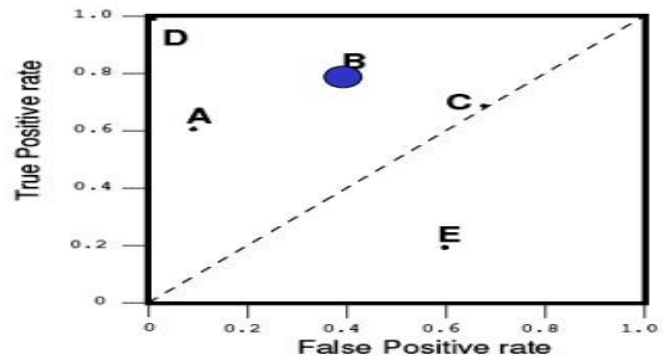
- Classifiers on the upper right-hand side of an ROC graph may be thought of as “liberal”
  - they make positive classifications with weak evidence so they classify nearly all positives correctly,
  - but they often have high false positive rates.
- In figure, C is more liberal than A.



Metric	**Conservative Classifier (A)**	**Liberal Classifier (C)**
True Positive (TP)	Low	High
False Positive (FP)	Low	High
True Negative (TN)	High	Low
False Negative (FN)	High	Low
TPR (Sensitivity)	Low	High
FPR	Low	High

Curves and points in ROC space

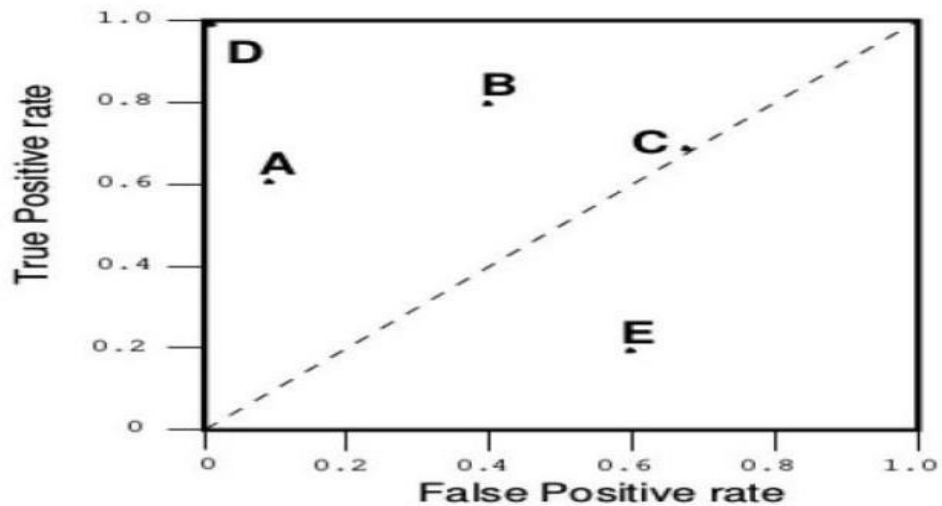
- Many classifiers, such as decision trees or rule learners, are designed to produce only a class decision, i.e., a Y or N on each instance.
  - When such a discrete classier is applied to a test set, it yields a single confusion matrix, which in turn corresponds to one ROC point.
  - Thus, a discrete classifier produces only a single point in ROC space.



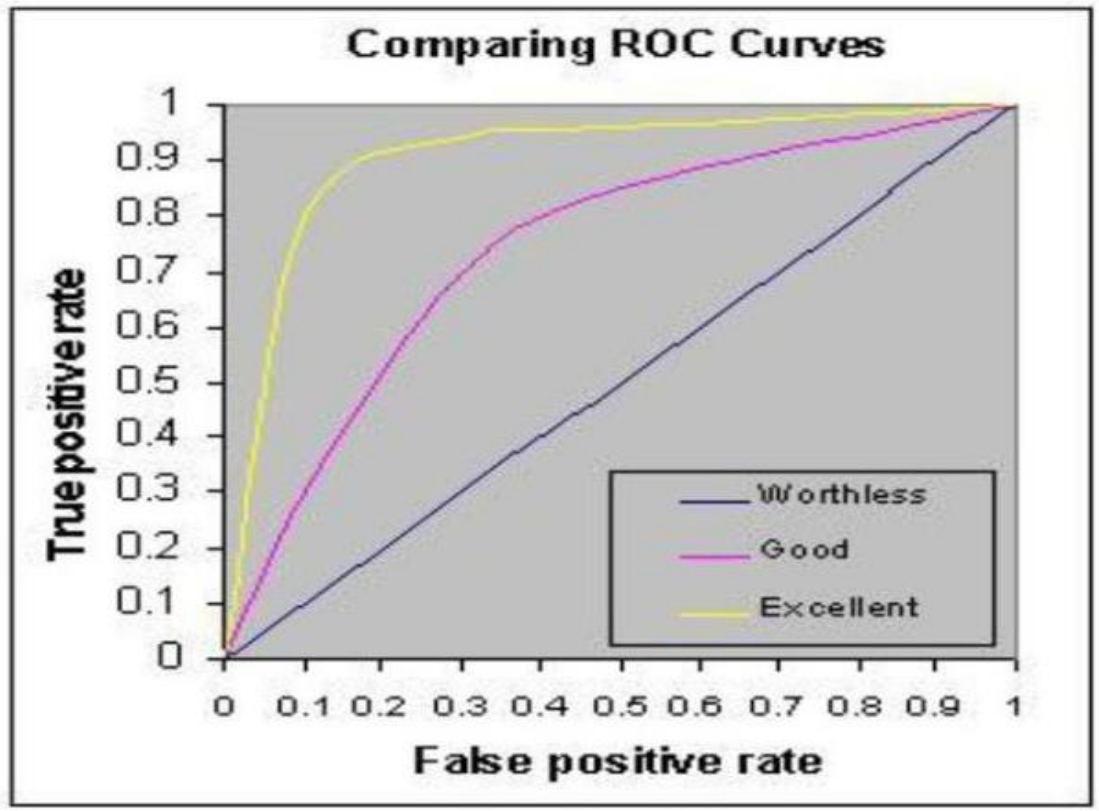
		Predicted class	
		Class +	Class -
Actual class	Class +	80	20
	Class -	40	60

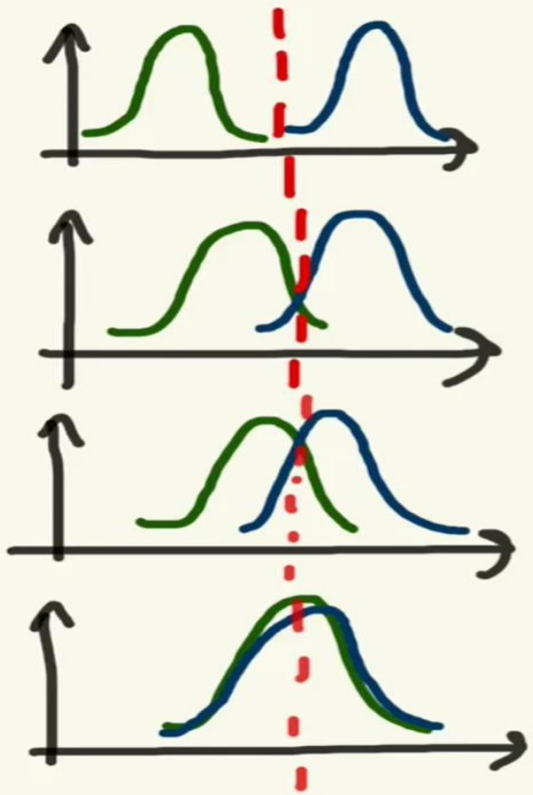
Confusion matrix for classifier B

ROC Graph

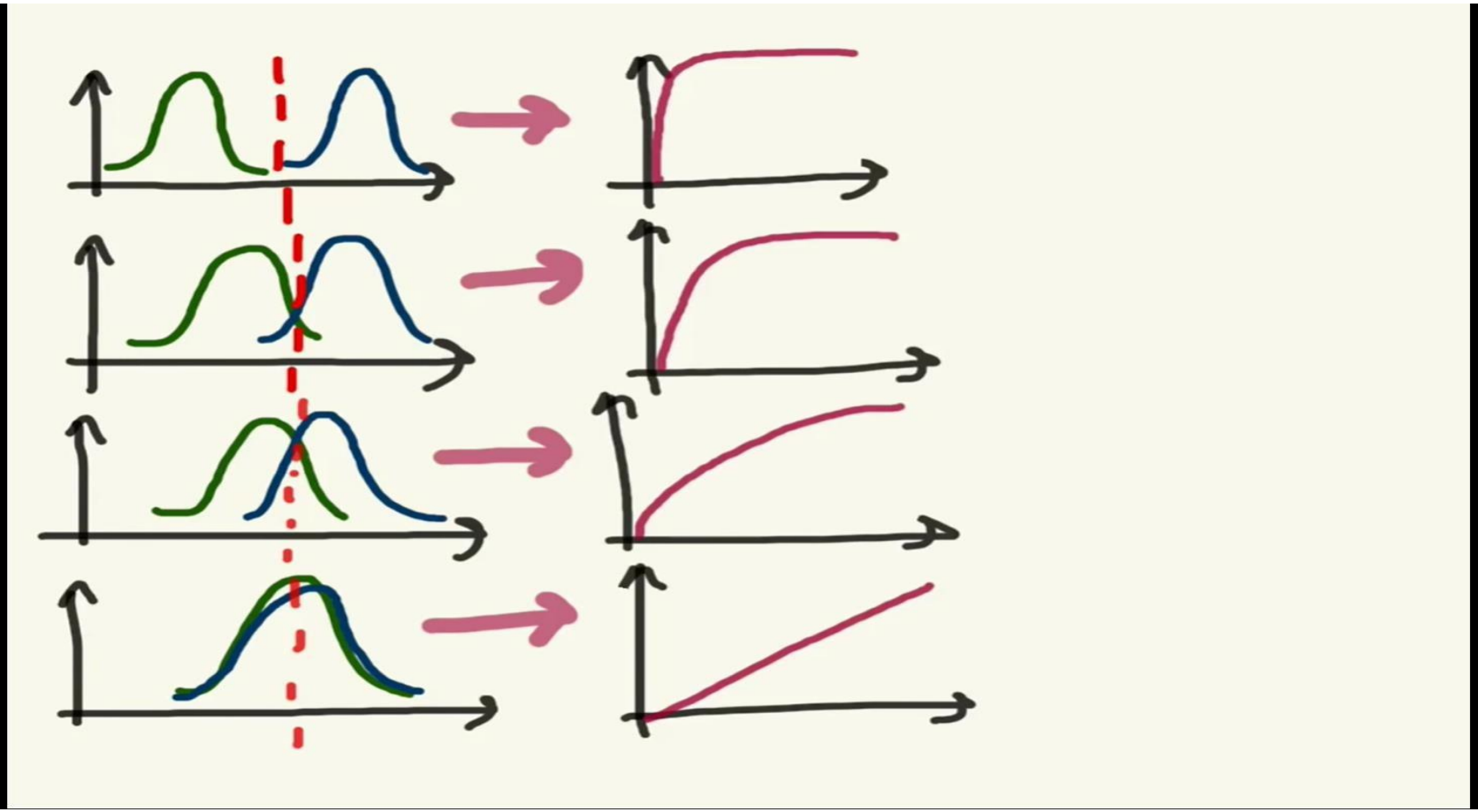


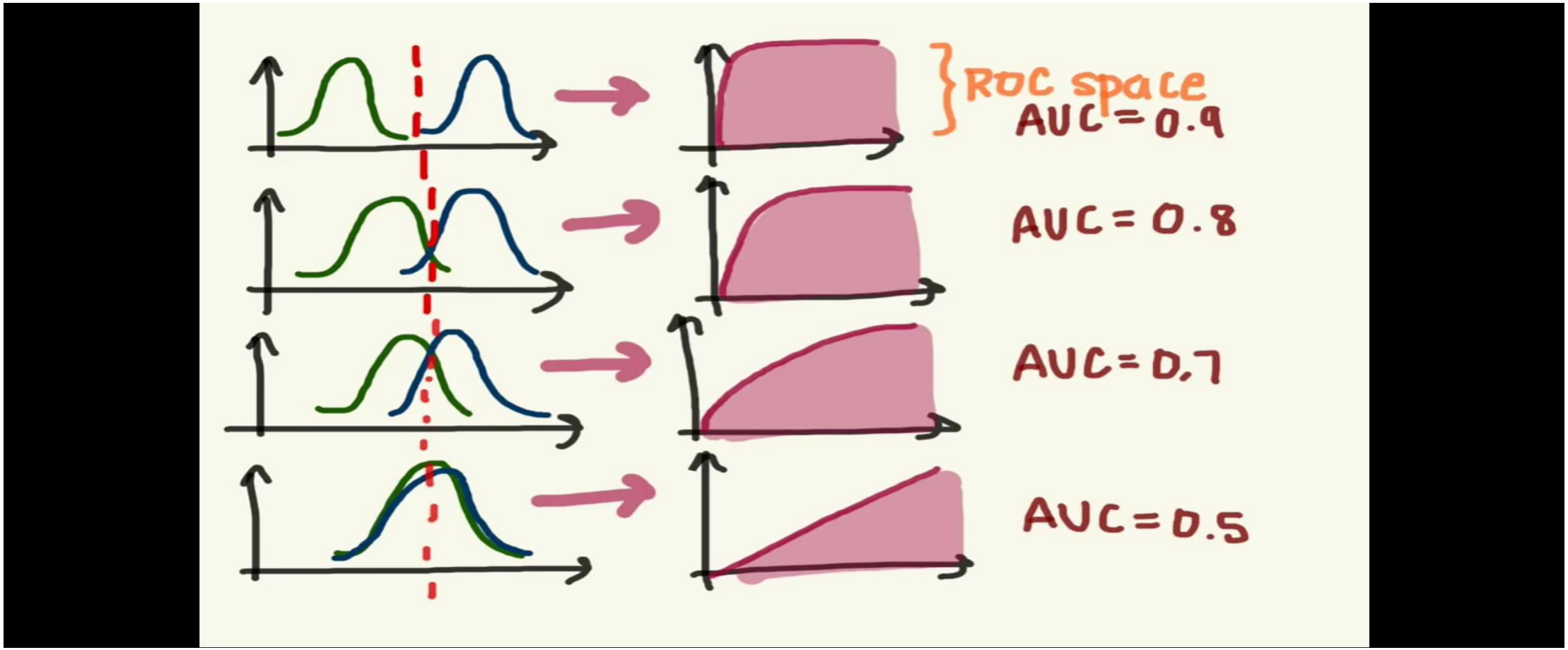
→ One point in ROC space is better than another if it is to the northwest (tp rate is higher, fp rate is lower, or both) of the first. (D is better than all other points)
→ Classifiers in the lower left part are known as conservative.
→ Classifiers in the upper right part are known as liberals.
→ the diagonal y=x line represents a random classifier.
→ Classifiers in the lower right triangle performs worse than random classifier.


Comparing ROC Curves
True positive rate
False positive rate
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Worthless
Good
Excellent

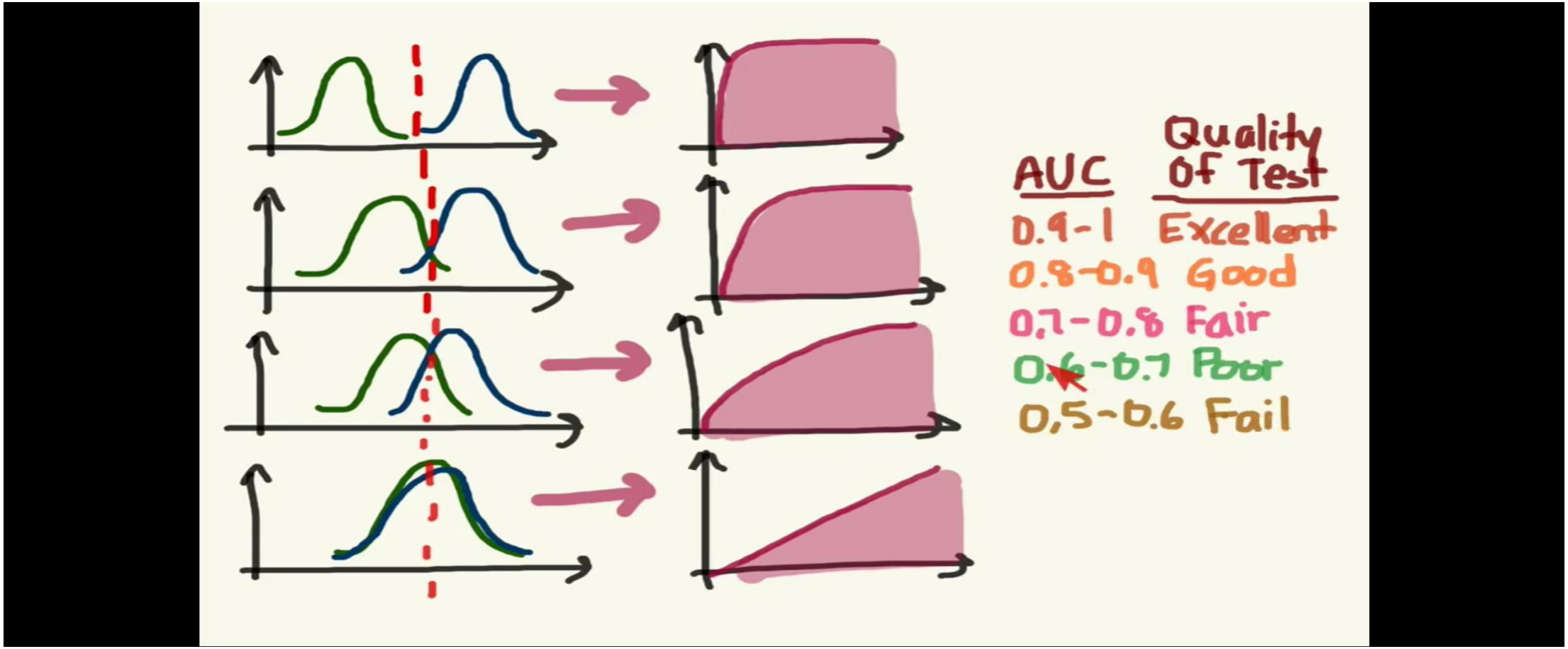


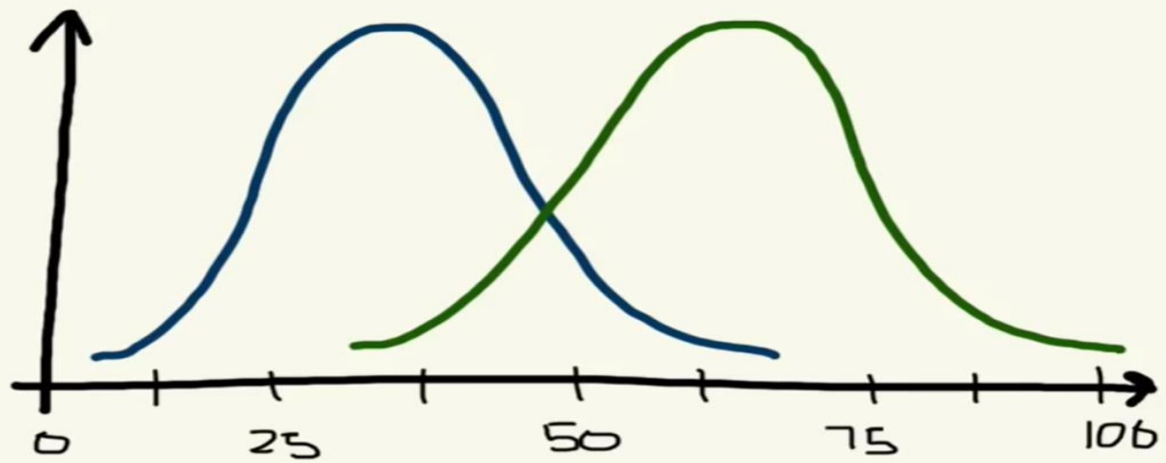
AUC is a measure of how well a test can distinguish between different classes.



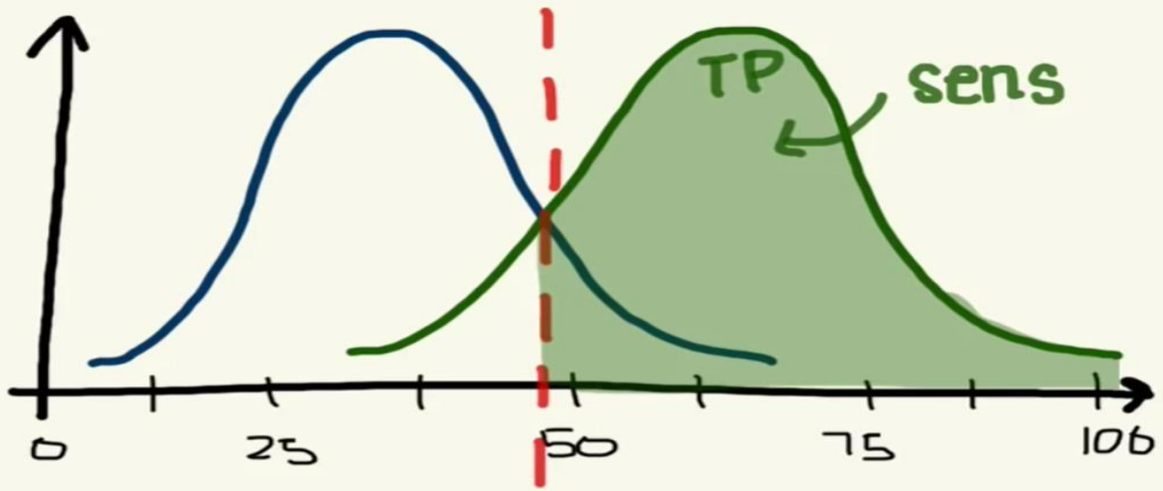

ROC space
AUC = 0.9
AUC = 0.8
AUC = 0.7
AUC = 0.5

AUC is a measure of how well a test can distinguish between different classes.

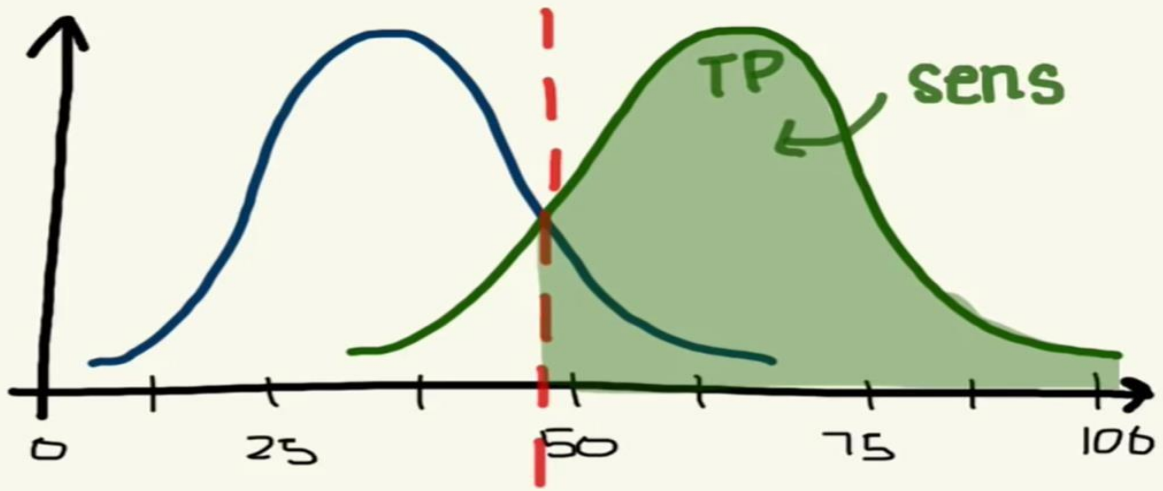




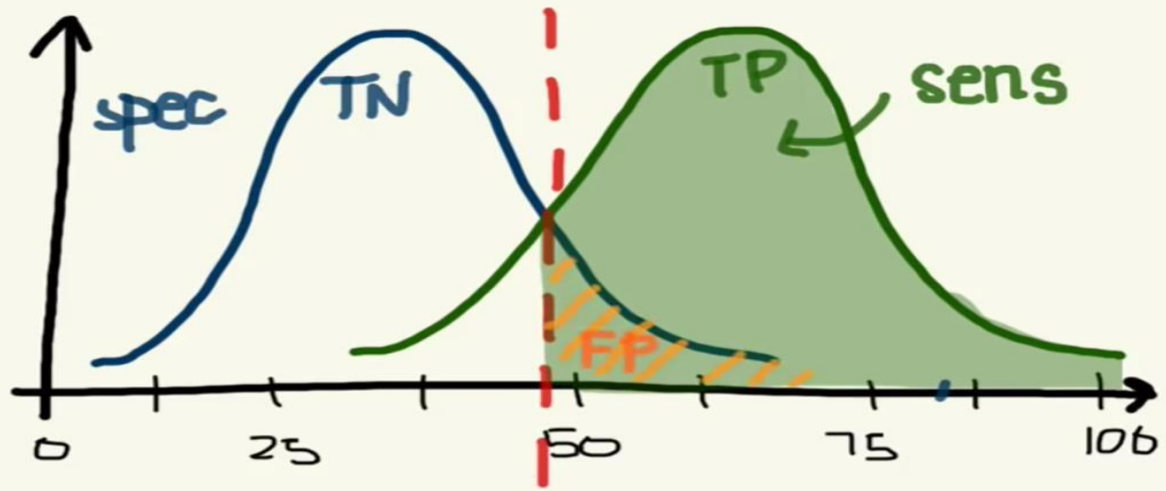
"why (1-spec)?"



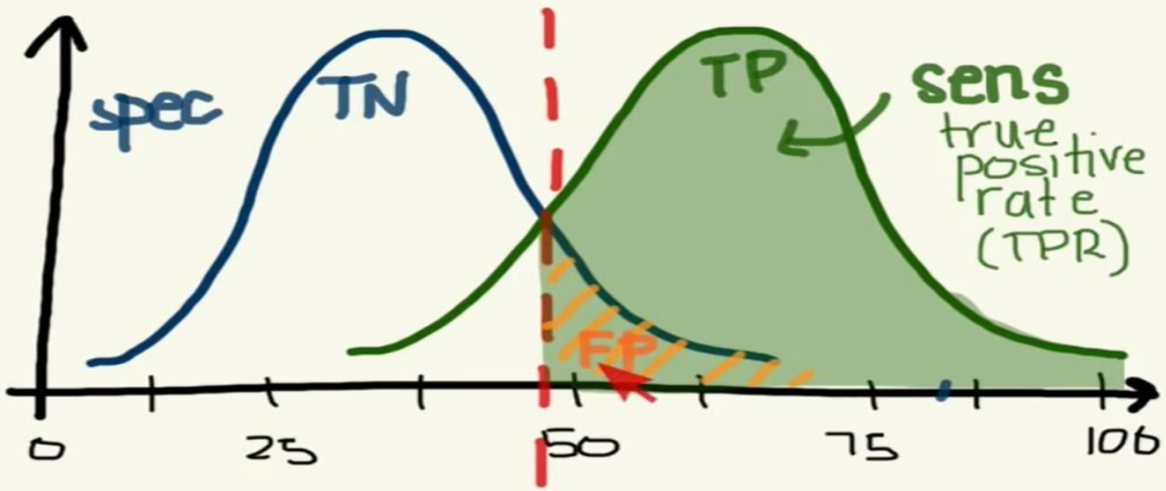
"why (1-spec)?"



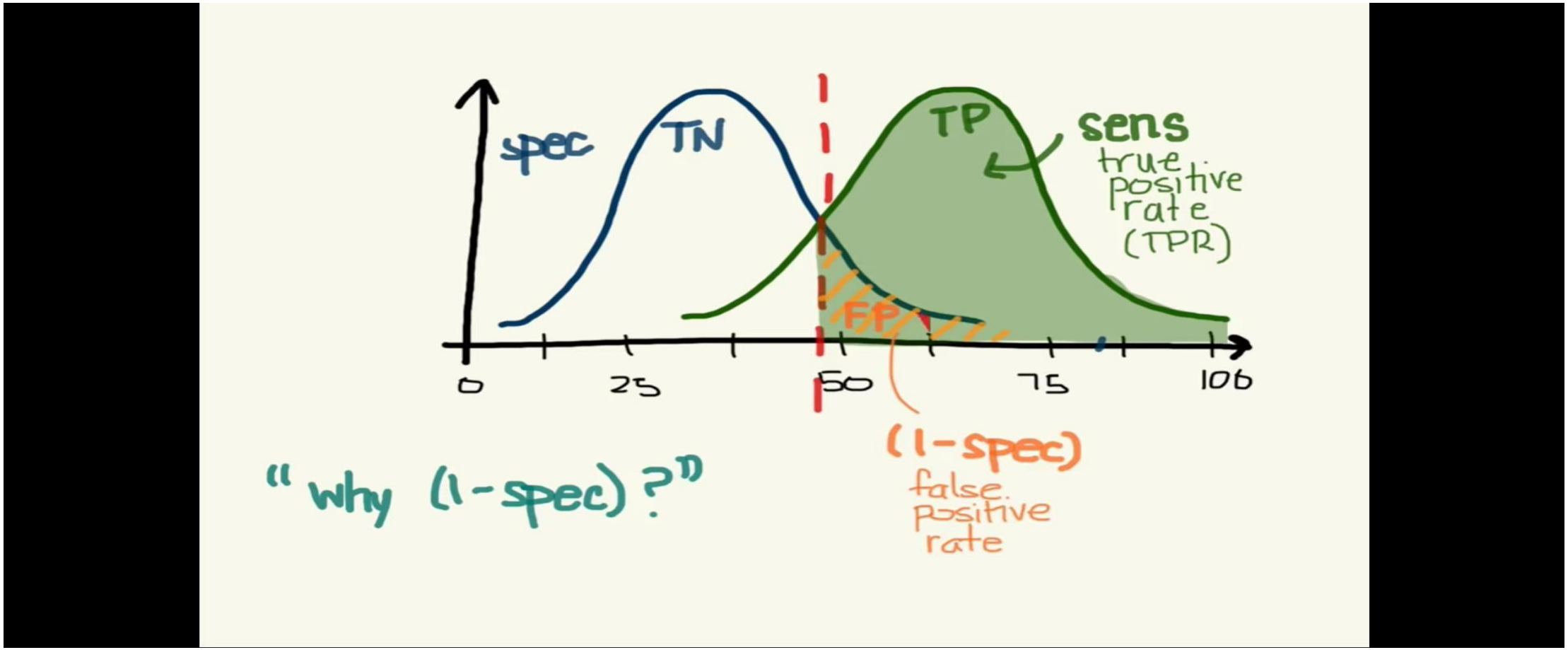
"why (1-spec)?"



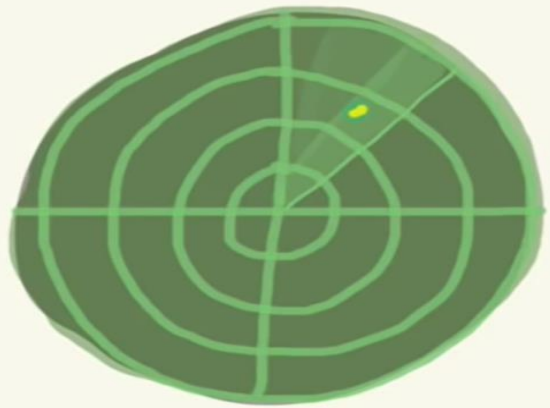
"why (1-spec)?"



"why (1-spec)?"


spec
TN
TP
sens
true positive rate (TPR)
FP
0
25
50
75
100
(1-spec)
false positive rate
"why (1-spec)?"

WW II - Signal Detection Theory



true positive rate

false positive rate

ROC