

ROC Curve

- An ROC curve (Receiver Operating Characteristic) is a graph that shows the performance of a classification model at different level of decision thresholds.
  - Each point on the ROC curve represents a sensitivity/specificity pair corresponding to a particular decision threshold.
  - A test with perfect discrimination (no overlap in the two distributions) has a ROC curve that passes through the upper left corner (100% sensitivity, 100% specificity).
  - Therefore the closer the ROC curve is to the upper left corner, the higher the overall accuracy of the test (Zweig & Campbell, 1993).

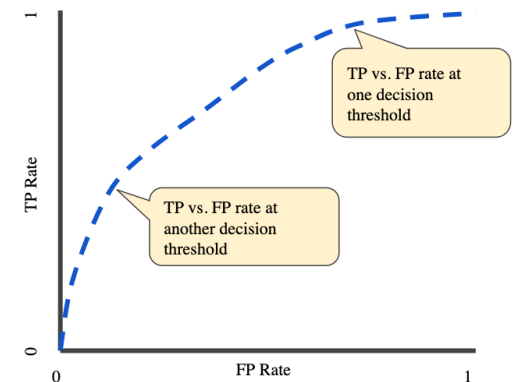
- This curve represents two parameters:

➤ True positive rate

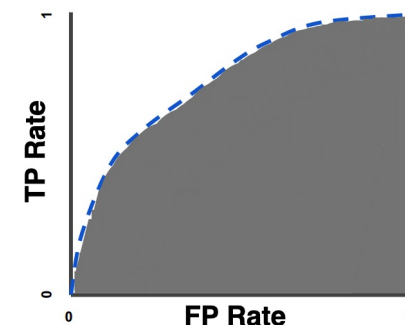
$$TPR = \frac{TP}{TP + FN}$$

➤ False positive rate

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



- **AUC** means «Area under the ROC Curve»
- One way of interpreting the AUC is the probability that the model ranks a random positive example more highly than a random negative example
- The AUC is between 0 and 1.
  - A model whose predictions are completely wrong has an AUC of 0.0.
  - A model whose predictions are correct has an AUC of 1.0.
- The AUC is desirable for the following two reasons:
  - The AUC is invariant. It measures the ranking of predictions, rather than their absolute values.
  - The AUC is independent of decision thresholds. It measures the quality of the model's predictions independently of the classification threshold chosen.



- For more information on ROC and AUC:  
<https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5>