AUC-ROC Curve

(This note is generated from Gen AI tools and augmented with extra information).

The Area Under the Receiver Operating Characteristic (AUC-ROC) curve is a measure of a classification model's performance across all possible classification thresholds. An AUC of 1.0 means perfect classification, while 0.5 means the model performs no better than random guessing.

Here is a step-by-step worked example for calculating the AUC and plotting the ROC curve for a simple binary classification model.

Example Data: Model Scores

Imagine a classifier predicts the probability (score) of an instance belonging to the positive class (Class = 1). We have 6 instances, where the actual true class is known:

Instance	True Class	Predicted Score (Probability of Class 1)
A	1	0.9
B	0	0.8
C	1	0.7
D	1	0.4
E	0	0.3
F	0	0.2

Step 1: Sorting and Calculating Confusion Matrix Points

We sort the data by the predicted score in descending order. The ROC curve is generated by treating **every unique predicted score** as a potential **threshold** and calculating the True Positive Rate (TPR) and False Positive Rate (FPR) at each threshold.
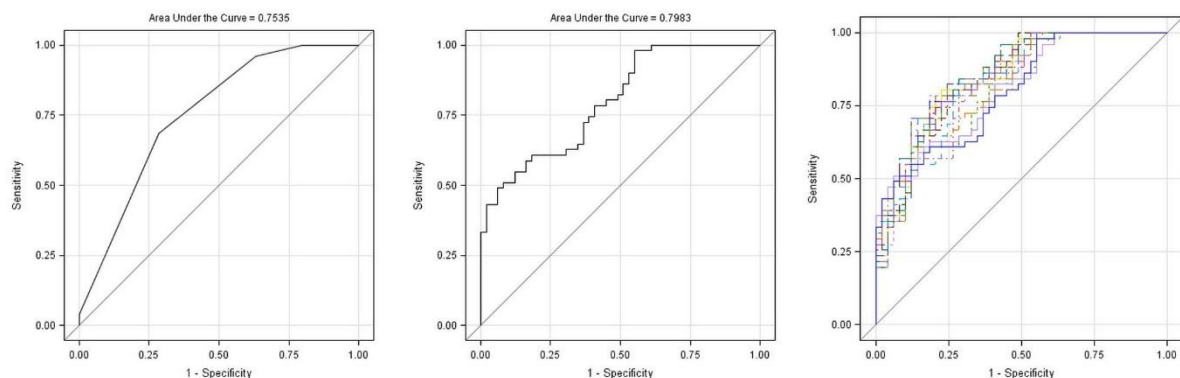
- P: Total number of Positive instances (True Class = 1) = 3 (A, C, D)

- N: Total number of Negative instances (True Class = 0) = 3 (B, E, F)

Threshold	Instances Predicted Positively	TP (True Positives)	FP (False Positives)	TPR (TP/P)	FPR (FP/N)	ROC Plot Point (FPR, TPR)
Start (Threshold=1.0)	None	0	0	0/3 = **0.00**	0/3 = **0.00**	**(0.00, 0.00)**
0.9 (A's score)	A	1	0	1/3 = **0.33**	0/3 = **0.00**	**(0.00, 0.33)**
0.8 (B's score)	A, B	1	1	1/3 = **0.33**	1/3 = **0.33**	**(0.33, 0.33)**
0.7 (C's score)	A, B, C	2	1	2/3 = **0.67**	1/3 = **0.33**	**(0.33, 0.67)**
0.4 (D's score)	A, B, C, D	3	1	3/3 = **1.00**	1/3 = **0.33**	**(0.33, 1.00)**
0.3 (E's score)	A, B, C, D, E	3	2	3/3 = **1.00**	2/3 = **0.67**	**(0.67, 1.00)**
0.2 (F's score)	All	3	3	3/3 = **1.00**	3/3 = **1.00**	**(1.00, 1.00)**

Note: TPR = Recall/Sensitivity; FPR = 1 - Specificity.

Step 2: Plotting the ROC Curve

We plot the coordinates (FPR, TPR) found in Step 1. The curve starts at (0, 0) and ends at (1, 1).

Step 3: Calculating AUC (Area Under the Curve)

The AUC is the area under the stepped curve generated by connecting these points. We can calculate this using the **Trapezoidal Rule** (summing the areas of the rectangles/trapezoids under each segment).

The area of a trapezoid is given by: Area = Height x Width

Where:

- **Height:** The average of the current and previous TPR.

- **Width:** The difference between the current and previous FPR.

$$\text{AUC} = \sum_{i=1}^{m} (\text{TPR}_i + \text{TPR}_{i-1}) \cdot (\text{FPR}_i - \text{FPR}_{i-1}) \cdot \frac{1}{2}$$

Segment	Previous Point	Current Point	Width (ΔFPR)	Avg Height (Avg TPR)	Area Contribution
1	(0.00, 0.00)	(0.00, 0.33)	0.00 - 0.00 = **0.00**	(0.00+0.33)/2 = 0.165	$0.00 \times 0.165 =$ **0.000**
2	(0.00, 0.33)	(0.33, 0.33)	0.33 - 0.00 = **0.33**	(0.33+0.33)/2 = 0.330	$0.33 \times 0.330 =$ **0.109**
3	(0.33, 0.33)	(0.33, 0.67)	0.33 - 0.33 = **0.00**	(0.33+0.67)/2 = 0.500	$0.00 \times 0.500 =$ **0.000**
4	(0.33, 0.67)	(0.33, 1.00)	0.33 - 0.33 = **0.00**	(0.67+1.00)/2 = 0.835	$0.00 \times 0.835 =$ **0.000**
5	(0.33, 1.00)	(0.67, 1.00)	0.67 - 0.33 = **0.34**	(1.00+1.00)/2 = 1.000	$0.34 \times 1.000 =$ **0.340**
6	(0.67, 1.00)	(1.00, 1.00)	1.00 - 0.67 = **0.33**	(1.00+1.00)/2 = 1.000	$0.33 \times 1.000 =$ **0.330**
Total					**0.779**

The final AUC for this model is approximately **0.779**. This indicates a good ability to distinguish between the positive and negative classes.