

AUC-ROC Curve in Machine Learning

In Machine Learning, only developing an ML model is not sufficient as we also need to see whether it is performing well or not. It means that after building an ML model, we need to evaluate and validate how good or bad it is, and for such cases, we use different Evaluation Metrics. *AUC-ROC curve is such an evaluation metric that is used to visualize the performance of a classification model.* It is one of the popular and important metrics for evaluating the performance of the classification model. In this topic, we are going to discuss more details about the AUC-ROC curve.

Note: For a better understanding of this article, we suggest you first understand [the Confusion Matrix](#), as AUC-ROC uses terminologies used in the Confusion matrix.

What is AUC-ROC Curve?

AUC-ROC curve is a performance measurement metric of a classification model at different threshold values. Firstly, let's understand ROC (Receiver Operating Characteristic curve) curve.

ROC Curve

ROC or Receiver Operating Characteristic curve represents a probability graph to show the performance of a classification model at different threshold levels. The curve is plotted between two parameters, which are:

- True Positive Rate or TPR
- False Positive Rate or FPR

In the curve, TPR is plotted on Y-axis, whereas FPR is on the X-axis.

TPR:

TPR or True Positive rate is a synonym for Recall, which can be calculated as:

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

FPR or False Positive Rate can be calculated as:

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

Here, **TP**: True Positive

FP: False Positive

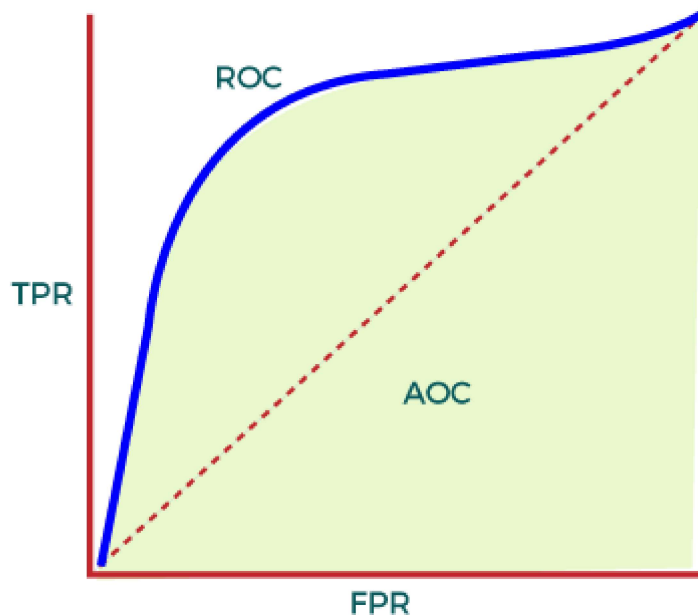
TN: True Negative

FN: False Negative

Now, to efficiently calculate the values at any threshold level, we need a method, which is AUC.

AUC: Area Under the ROC curve

AUC is known for **Area Under the ROC curve**. As its name suggests, AUC calculates the two-dimensional area under the entire ROC curve ranging from (0,0) to (1,1), as shown below image:



In the ROC curve, AUC computes the performance of the binary classifier across different thresholds and provides an aggregate measure. The value of AUC ranges from 0 to 1, which means an excellent model will have AUC near 1, and hence it will show a good measure of Separability.

When to Use AUC-ROC

AUC is preferred due to the following cases:

- AUC is used to measure how well the predictions are ranked instead of giving their absolute values. Hence, we can say AUC is **Scale-Invariant**.
- It measures the quality of predictions of the model without considering the selected classification threshold. It means AUC is **classification-threshold-invariant**.

When not to use AUC-ROC

- AUC is not preferable when we need to calibrate probability output.
- Further, AUC is not a useful metric when there are wide disparities in the cost of false negatives vs false positives, and it is difficult to minimize one type of classification error.

How AUC-ROC curve can be used for the Multi-class Model?

Although the AUC-ROC curve is only used for binary classification problems, we can also use it for multiclass classification problems. For multi-class classification problems, we can plot N number of AUC curves for N number of classes with the One vs ALL method.

For example, if we have three different classes, X, Y, and Z, then we can plot a curve for X against Y & Z, a second plot for Y against X & Z, and the third plot for Z against Y and X.

Applications of AUC-ROC Curve

Although the AUC-ROC curve is used to evaluate a classification model, it is widely used for various applications. Some of the important applications of AUC-ROC are given below:

1. Classification of 3D model

The curve is used to classify a 3D model and separate it from the normal models. With the specified threshold level, the curve classifies the non-3D and separates out the 3D models.

2. Healthcare

The curve has various applications in the healthcare sector. It can be used to detect cancer disease in patients. It does this by using false positive and false negative rates, and accuracy depends on the threshold value used for the curve.

3. Binary Classification

AUC-ROC curve is mainly used for binary classification problems to evaluate their performance.