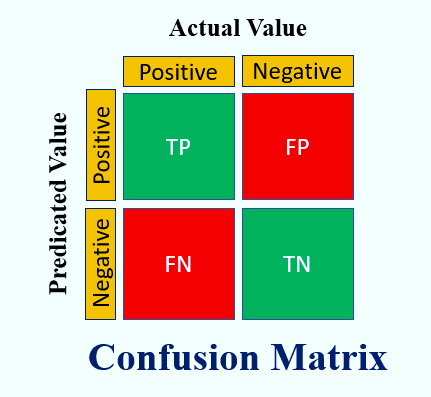
**What is ROC and Auc Curve**

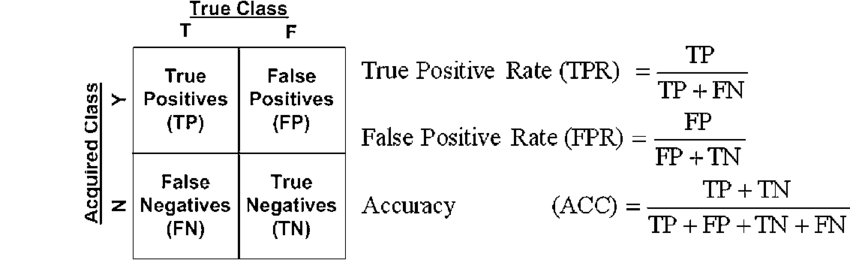
**ROC (Receiver Operating Characteristic) curve** is a graphical representation of the performance of a binary classification model across different threshold values. It plots the true positive rate (TPR) against the false positive rate (FPR) for various threshold settings.

Here's how it works:



**True Positive Rate (TPR)**,

also known as sensitivity or recall, is the proportion of actual positive cases that are correctly identified by the model. It is calculated as

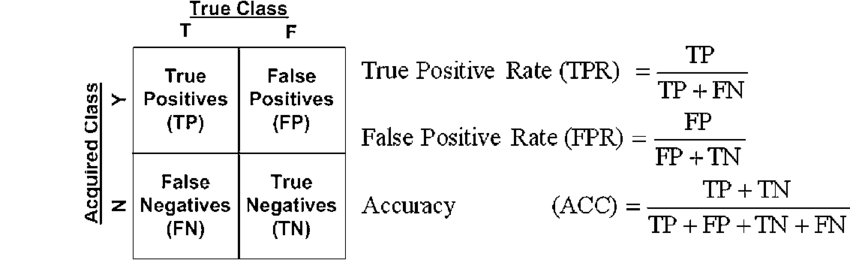


where TP is the number of true positives (correctly predicted positive cases) and

FN is the number of false negatives (positive cases incorrectly predicted as negative).

**False Positive Rate (FPR)**

is the proportion of actual negative cases that are incorrectly classified as positive by the model. It is calculated as



where FP is the number of false positives (negative cases incorrectly predicted as positive) and

TN is the number of true negatives (correctly predicted negative cases).

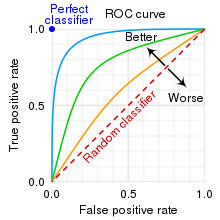
The ROC curve is generated by plotting TPR on the y-axis against FPR on the x-axis for different threshold values used by the classification model. Each point on the ROC curve represents a specific threshold setting. The diagonal line (y = x) represents the performance of a random classifier.

A perfect classifier would have an ROC curve that passes through the upper left corner (TPR = 1, FPR = 0), indicating high sensitivity (true positive rate) and low false positive rate. The area under the ROC curve (AUC-ROC) is a commonly used metric to quantify the overall performance of a binary classification model. A higher AUC-ROC value (closer to 1) indicates better discrimination ability of the model across all threshold settings.

In summary, the ROC curve is a useful tool for evaluating the performance of binary classification models and comparing different models based on their ability to trade-off between true positive and false positive rates across various threshold values.

**AUC stands for Area Under the ROC Curve.**

It is a metric used to evaluate the performance of a binary classification model based on its Receiver Operating Characteristic (ROC) curve.

The ROC curve plots the True Positive Rate (sensitivity) against the False Positive Rate (1 - specificity) for various threshold settings of the classification model. The AUC-ROC is then calculated as the area under this curve. 

Here's what the AUC represents and how it's interpreted:

**AUC < 0.5: Inverted classifier.**

The model performs worse than random guessing. This could happen when the model predictions are consistently opposite to the true labels.

**AUC= 0.5: Random classifier. TPR=0.5 and FPR: 0.5**

The model performs no better than random guessing. The ROC curve follows the diagonal line (y = x), indicating that the model's true positive rate is equal to its false positive rate across all thresholds.

**AUC = 1: Perfect classifier, TPR=1 and FPR: 0**

The model achieves a perfect separation between positive and negative classes, with a TPR of 1 and an FPR of 0 across all threshold settings.

In summary, AUC is a single scalar value that summarizes the overall performance of a binary classification model. A higher AUC indicates better performance in distinguishing between positive and negative classes, making it a widely used metric for evaluating and comparing different classification models.