**Metrics**

## 1- Confusion Matrix

## 2- Classification Accuracy

Classification accuracy is perhaps the simplest metrics one can imagine, and is defined as the**number of correct predictions divided by the total number of predictions,**multiplied by 100**.**

## 3- Precision

**Precision= True\_Positive/ (True\_Positive+ False\_Positive)**

## 4- Recall

**Recall= True\_Positive/ (True\_Positive+ False\_Negative)**

## 5- F1 Score

**F1-score= 2\*Precision\*Recall/(Precision+Recall)**

## 6- Sensitivity and Specificity

Sensitivity= Recall= TP/(TP+FN)

Specificity= True Negative Rate= TN/(TN+FP)

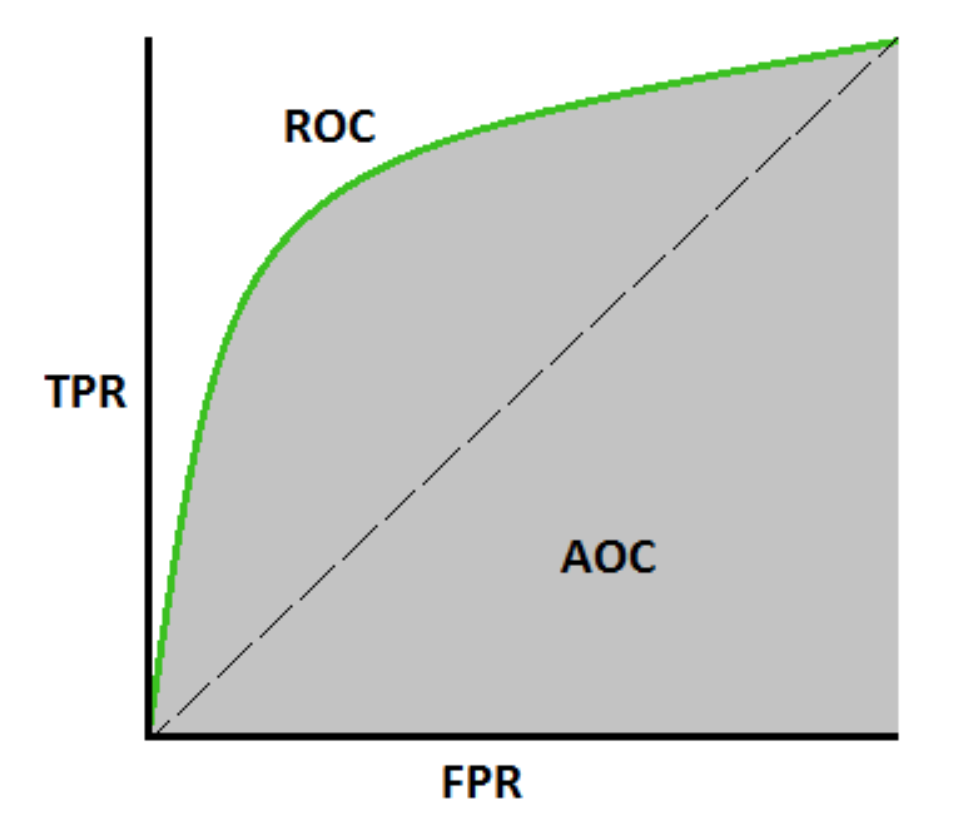
## 7- ROC Curve

The **receiver operating characteristic curve** is plot which shows the performance of a binary classifier as function of its cut-off threshold. **It essentially shows the true positive rate (TPR) against the false positive rate (FPR) for various threshold values.**

## 8- AUC

The **area under the curve** (AUC), is an aggregated measure of performance of a binary classifier on all possible threshold values (and therefore it is threshold invariant**)**.

AUC calculates the area under the ROC curve, and therefore it is between 0 and 1. One way of interpreting AUC is as the probability that the model ranks a random positive example more highly than a random negative example.



### 9. **Cohen’s kappa**

This measure is intended to compare labelings by different human annotators, not a classifier versus a ground truth.

The kappa score is a number between -1 and 1. Scores above .8 are generally considered good agreement; zero or lower means no agreement (practically random labels).

Kappa scores can be computed for binary or multiclass problems, but not for multilabel problems (except by manually computing a per-label score) and not for more than two annotators.

### 10. **Hamming loss**

The **[hamming\_loss](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.hamming_loss.html" \l "sklearn.metrics.hamming_loss" \o "sklearn.metrics.hamming_loss)** computes the average Hamming loss or [Hamming distance](https://en.wikipedia.org/wiki/Hamming_distance) between two sets of samples.

### 11. **Jaccard similarity coefficient score**

The **[jaccard\_score](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.jaccard_score.html" \l "sklearn.metrics.jaccard_score" \o "sklearn.metrics.jaccard_score)** function computes the average of [Jaccard similarity coefficients](https://en.wikipedia.org/wiki/Jaccard_index), also called the Jaccard index, between pairs of label sets.

Links:

1. <https://scikit-learn.org/stable/modules/model_evaluation.html#classification-metrics>
2. <https://machinelearningmastery.com/how-to-calculate-precision-recall-f1-and-more-for-deep-learning-models/>
3. <https://towardsdatascience.com/20-popular-machine-learning-metrics-part-1-classification-regression-evaluation-metrics-1ca3e282a2ce>