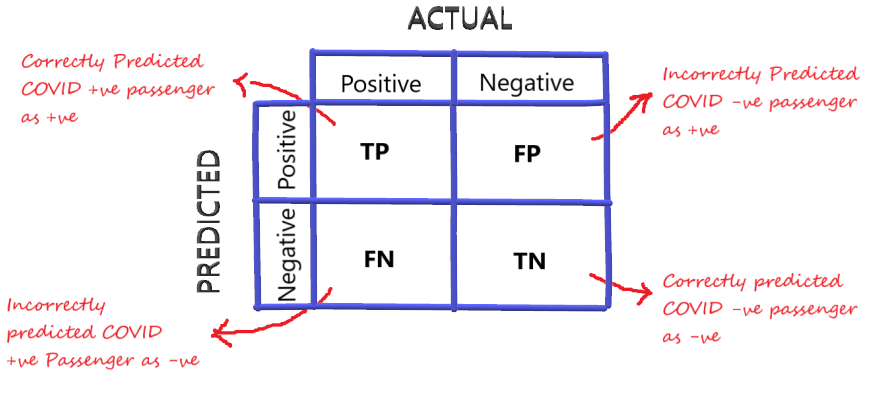
**Metrics for Classification**

**Confusion Matrix**

For better visualization of the performance of a model, these four outcomes are plotted on a confusion matrix.



**True Positive (TP):**When you predict an observation belongs to a class and it actually does belong to that class.

**True Negative (TN):**When you predict an observation does not belong to a class and it actually does not belong to that class.

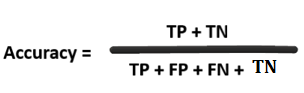
**False Positive (FP)**: When you predict an observation belongs to a class and it actually does not belong to that class.

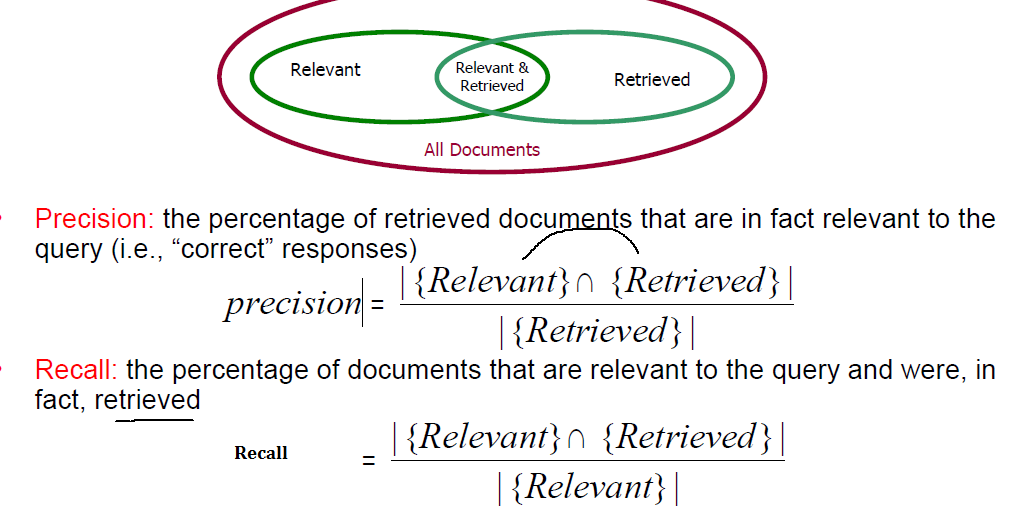
**False Negative(FN):**When you predict an observation does not belong to a class and it actually does belong to that class.

## Accuracy

Accuracy is one metric which gives the fraction of predictions our model got right. Formally, accuracy has the following definition:

Accuracy = Number of correct predictions / Total number of predictions.





## Recall (Sensitivity or True positive rate)

Recall gives the fraction you correctly identified as positive out of all positives.

Diagram, text

Description automatically generated

## Precision

Precision gives the fraction of correctly identified as positive out of all predicted as positives.

Diagram

Description automatically generated

## **Precision** is a metric that calculates the percentage of correct predictions for the positive class.

## **Recall** calculates the percentage of correct predictions for the positive class out of all positive predictions that could be made.

## Maximizing precision will minimize the false-positive errors, whereas maximizing recall will minimize the false-negative errors.

## Precision-Recall Tradeoff

Chart, line chart

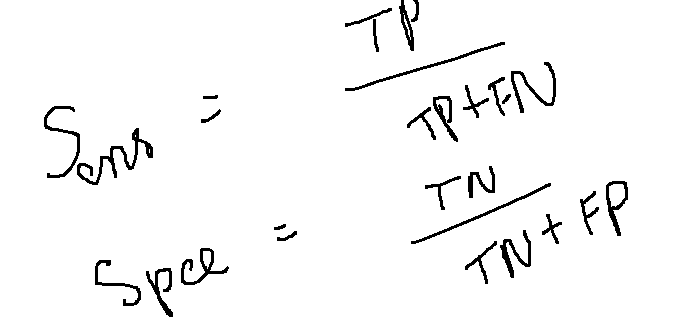
Description automatically generated

## F1 Score

It is defined as the harmonic mean of the model’s precision and recall.

Graphical user interface, text

Description automatically generated



**Sensitivity and Specificity**

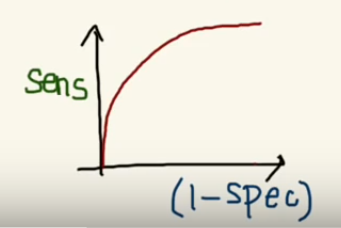
As **Sensitivity** ⬇️ **Specificity** ⬆️

As **Specificity** ⬇️ **Sensitivity** ⬆️

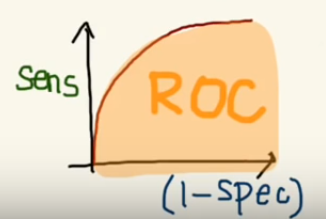


**Receiver Operating Characteristic (ROC)**

To plot ROC curve, instead of Specificity we use (1 — Specificity) and the graph will look something like this:

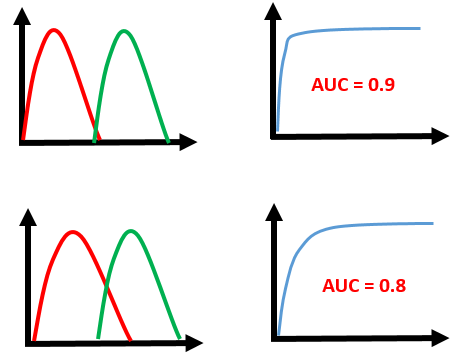


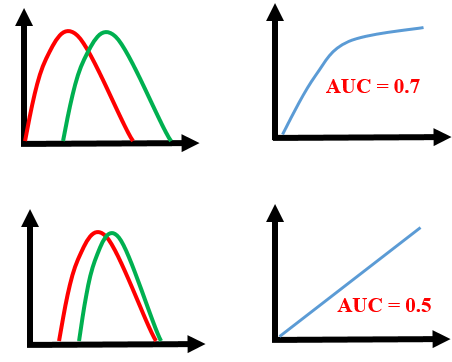
So now, when the sensitivity increases, (1 — specificity) will also increase. This curve is known as the ROC curve.



**Area Under the Curve (AUC)**

The AUC is the area under the ROC curve. This score gives us a good idea of how well the model performances.



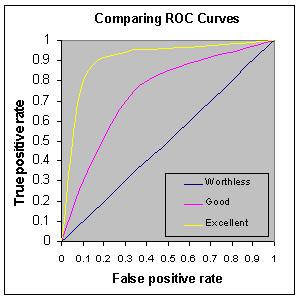


As we see, the first model does quite a good job of distinguishing the positive and the negative values. Therefore, there the AUC score is 0.9 as the area under the ROC curve is large.

Whereas, if we see the last model, predictions are completely overlapping each other, and we get the AUC score of 0.5. This means that the model is performing poorly, and it is predictions are almost random.

Specificity gives us the True Negative Rate and (1 — Specificity) gives us the False Positive Rate.

So, the sensitivity can be called as the “***True Positive Rate***” and (1 — Specificity) can be called as the “***False Positive Rate***”.



A rough guide for classifying the accuracy of a diagnostic test is the traditional academic point system:

* .90-1 = excellent (A)
* .80-.90 = good (B)
* .70-.80 = fair (C)
* .60-.70 = poor (D)
* .50-.60 = fail (F)