ROC stands for Receiver Operating Characteristic .The purpose of using ROC is for better charting out the results of many confusion matrices with many different results and different thresholds .

Following is the structure of a confusion matrix with 2 classes having results as positive or negative

It is mainly used in order to show the relationship between sensitivity and specificity . It plots the True Positive Rate which is sensitivity against the false positive rate which is (1- specificity ).

In order to determine the sensitivity (true positive rate ) and the false positive rate (1-specificity) . it is required to use the below formulas

ROC curve characteristics

* The ROC curve does not provide information about the actual values of thresholds used for the classifier.
* Performance of different classifiers can be compared using the AUC of different Classifier. The larger the AUC, the better the classifier.

In order to proceed to understand the practical implementation we are making use of an example that has classification of people as healthy or unhealthy on the basis of absence of a vitamin .Tthe mean of the parameter with healthy and not healthy datapoints are close and the distributions overlap, as shown  below . On setting the threshold to 0.5 the confusion matrix based on the 500 observations is shown below

This tells us that the accuracy is only 88 % and tells us that this is the ROC curve for the worst possible classifier. Therefore, by comparing the obtained ROC curve with the diagonal, we see how much better our classifer is from random guessing. We then understand that the more further from the diagonal the ROC curve is the better the classifier is

References :

Rohitmishra. (2019, April 9). rohitmishra. Retrieved from <https://data-science-blog.com/blog/2019/04/09/introduction-to-roc-curve/>.

Precision replaces false positive rate