

Evaluating Classification model Performance

1

Confusion Matrix

DEFINITION

A confusion matrix is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known

	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

Example

	Predicted: NO	Predicted: YES
Actual: NO	TN = 50	FP = 10
Actual: YES	FN = 5	TP = 100

Parameter

	Predicted: NO	Predicted: YES
Actual: NO	TN = 50	FP = 10
Actual: YES	FN = 5	TP = 100

Evaluation

	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

Total no. of Patient : 165

Cancer - Binary prediction (Yes / No)

ACTUAL

Cancer Yes - 105 Patient | Cancer No- 60 Patient

OUR MODEL PREDICTED

Cancer Yes - 110 Patient | Cancer No- 55 Patient

True positives (TP)These are cases in which we predicted yes (they have the cancer), and they have the cancer

True negatives (TN)We predicted no, and they don't have the cancer

False positives (FP)/Type I errorWe predicted yes, but they don't actually have the cancer

False negatives (FN)/Type II errorWe predicted no, but they actually have the cancer

AccuracyOverall correct
 $(TP+TN)/total = (100+50)/165 = 0.91$

Misclassification RateOverall wrong | Error Rate
 $(FP+FN)/total = (10+5)/165 = 0.09$

Sensitivity or RecallTrue Positive Rate
 $TP/actual\ yes = 100/105 = 0.95$

False Positive Rate
 $FP/actual\ no = 10/60 = 0.17$

SpecificityTrue Negative Rate
 $TN/actual\ no = 50/60 = 0.83$

Precision
 $TP/predicted\ yes = 100/110 = 0.91$

Prevalence
 $actual\ yes/total = 105/165 = 0.64$

2

Accuracy Paradox

Accuracy is defined as the freedom from mistake or error

Accuracy is not a reliable metric to determine a model performance for imbalanced data

That's why it's called a Paradox because, intuitively, you'd expect a Model with a higher Accuracy to have been the best Model but Accuracy Paradox tells us that this, sometimes, isn't the case

SOLUTION

F1 Score $2*((precision*recall)/(precision+recall))$

3

Receiver Operating Characteristic (ROC) Curve

True Positive Rate (TPR) is plot against False Positive Rate (FPR) for the probabilities of the classifier predictions. Then, the area under the plot is calculated

Area Under the Curve (AUC) is the measure of the ability of a classifier to distinguish between classes and is used as a summary of the ROC curve

The higher the AUC, the better the performance of the model at distinguishing between the positive and negative classes

$sensitivity = \frac{count\ TP}{count\ TP + count\ FN}$ $specificity = \frac{count\ TN}{count\ TN + count\ FP}$

Eqn, 1

4

Cumulative Accuracy Profile (CAP) Curve

It tries to analyse how to effectively identify all data points of a given class using minimum number of tries

It helps us to understand and conclude about the robustness of the classification model

CAP Analysis using Plot

Difference B/W ROC & CAP Curve - Easy Formula

X axis:

Gains	ROC
$\frac{count\ TP + count\ FP}{count\ all\ observations}$	$\frac{count\ FP}{count\ TN + count\ FP}$

Y axis:

$\frac{count\ TP}{count\ TP + count\ FN}$	$\frac{count\ TP}{count\ TP + count\ FN}$
---	---

Curve Graph