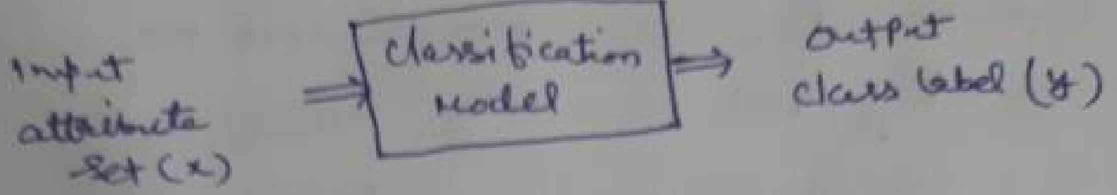
classification:- classification is the task of mapping an I/P attribute set x into its class label y.



Definition:- classification is the task of learning a target function f that maps each attribute set x to one of the predefined class label y.

The target function is also known as a classification model (informally)

A classification model is useful for the following purposes.

Descriptive Modeling:- A classification model can serve as an explanatory tool to distinguish b/w object of different classes.

Predictive modeling:- A classification model can also be used to predict the class label of unknown records.

General approach to Measure the Performance of classification problem:-

Evaluation of The Performance of a classification model is based on the counts of test records, which are predicted by the model correctly and incorrectly. These counts are tabulated in a table known as Confusion Matrix. Table depicts the confusion matrix for binary classification Problem.

Table:- Confusion Matrix for a 2-class Problem

		Predicted class	
		Class=1	Class=0
Actual class	Class=1	f_{11}	f_{10}
	Class=0	f_{01}	f_{00}

		Predicted class		total
		Yes	No	
Actual class	Yes	TP	FN (Type-2 error)	P
	No	FP	TN	N
	total	P'	N'	

Type-1 error

For Ex:-

		Predicted class		
		Yes	No	
Actual class	Yes	100	5	105
	No	10	50	60
		110	55	

what can be learn from This Matrix

There are Two Possible Predicted class "yes" or "No". If we were predicting the presence of a disease, for Ex:- "yes would mean They have The disease, and no would mean They don't have The disease.

The classifier made a total of 165 Predictions (e.g 165 patients were tested for The Presence of that disease)

Out of These 165 cases, The classifier Predicted "yes" 110 times and "no" 55 times.

in Reality, 105 patients in the sample have the disease and 60 Patient do not.

Basic term in confusion matrix :- (whole numbers Not Rate)

① True Positive (TP) :- These are cases in which we predicted yes (They have the disease), and They do have The disease.

② True Negative (TN) :- We predicted No and They don't have the disease.

③ False Positive (FP) :- we predicted yes, but They don't actually have the disease (Also known as type-1) error.

④ False Negative (FN) :- we predicted no, but they actually do have the desease.

this is the list of rates that are often computed from a confusion matrix for a binary classifier

(i) Accuracy :- overall, how often is the classifier correct ?

$$\text{Accuracy} : \frac{TP+TN}{TP+TN+FP+FN} = \frac{TP+TN}{\text{total}} = \frac{100+50}{165}$$

$\rightarrow 0.91$

* This is also called recognition rate

(ii) Error rate (Misclassification rate) = $\frac{FP+FN}{P+N}$

overall, how often it is wrong $\rightarrow \frac{10+5}{165} = 0.09$

It is equivalent to 1 - accuracy

also known as error rate

(iii) True positive rate (TPR) :- when its actually yes, how often does it predict yes.

$$\text{Sensitivity / recall} :- \frac{TP}{TP+FN} = \frac{TP}{P} = \frac{TP}{\text{actually ye}}$$

$\rightarrow \frac{100}{105} = 0.95$

$\rightarrow \frac{TP}{TP+FN}$

(iv) Specificity / True negative Rate (TNR) :→ when it's actually no, how often does it Predict no?

$$* \frac{TN}{TN+FP} = \frac{TN}{N} = \frac{50}{50+10} = \frac{50}{60} = 0.83$$

It is equivalent to 1 minus false Positive rate

(v) false Positive Rate : $\frac{FP}{\text{actual No}} = \frac{FP}{TN+FP} = \frac{10}{60} = 017$

when its actually No, how often does it Predict no?

(vi) Precision :- Precision can be thought of as a Measure of exactness

$$\text{Precision} = \frac{TP}{TP+FP}$$: when it is Predicted yes How often, it is correct

$$\Rightarrow \frac{TP}{\text{Predicted yes}} \Rightarrow \frac{100}{110} \Rightarrow 0.91$$

* what Percentage of tuples labeled as Positive as actually such.

(vii) Prevalence :- How often does the yes condition actually occur in our Sample

$$* \frac{\text{actual yes}}{\text{total yes}} = \frac{TP+FN}{TP+FN+TN+FP} = \frac{105}{165} = 0.64$$

For Example :-

class	yes	NO
yes	90	210
No	140	9560

Confusion matrix for class cancer = yes and no

(i) Sensitivity $= \frac{TP}{TP+FN} = \frac{90}{90+210} = \frac{90}{300} = 30\%$

(ii) Specificity :- $\frac{TN}{TN+FP} = \frac{9560}{9560+140} = \frac{9560}{9700} = 98.56\%$

(iii) Accuracy :- $\frac{TP+TN}{TP+TN+FP+FN} = \frac{9650}{10,000} = 96.50\%$

(iv) Precision $= \frac{TP}{TP+FP} = \frac{90}{90+140} = \frac{90}{230} = 39.13\%$

(v) false Positive rate $= 1 - \text{specificity}$

$\Rightarrow 1 - 98.56\% \Rightarrow 0.0144$

$\Rightarrow 1 - 0.9856 \Rightarrow 1\%$

Harmonic Mean :-

This is weighted Average of the True +ve rate (recall) and Precision, which is also called f. score

$$f\text{-Score} = \frac{2 * \text{Precision} * \text{recall}}{\text{Precision} + \text{recall}}$$

$$= \frac{2 * 0.91 * 0.95}{0.91 + 0.95} = \frac{1.729}{1.86} = 0.93$$ Ans

Matthews co-rrelation coefficient :-

It is used in machine learning as a Measure of The Quality of binary class classification, which is Introduced by biochemist Brian D. matthews in 1975. It takes ~~into~~ true and false Positive and negatives into its account and provide a balanced Measure, which can be used even if the classes are of very different in sizes.

The MCC is in essence a correlation coefficient b/w The observed and predicted binary classifications; it returns a value b/w -1 and +1. A coefficient of +1 represent a perfect Prediction, 0 no better Than random Prediction and -1 indicates total disagreement b/w Prediction and observation. MCC is one of the best Measure when Two classes are in very different in size. Accuracy Measure is also fail in case of class Imbalance.

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$