Measuring classifier performance:

It ~~predict~~ measures the predictive capabilities of machine learning models with metrics like accuracy, precision, recall & F1 score.

1. Performance Metrics for classification:

> The category or classes of data is identified based on training data.

> The Model learns from the given dataset and the classifies the new data into classes or groups based on the training.

> It predicts class labels as the o/p such as Yes or NO, 0 or 1, spam or Not spam.

1. Accuracy :

Formula :

$$\text{Accuracy} = \frac{\text{NO. of correct Predictions}}{\text{Total number of predictions.}}$$

2. Confusion Matrix :

A confusion matrix is a Tabular representation of prediction outcomes of any binary classifier, which is used to describe the performance of the classification model on a set of Test data when true values are known.

True class	Predicted class.		
	Positive	Negative.	
Positive	True positive 3	False Negative. 2	5
Negative	False positive 1	True Negative. 4	5
	4	6	

1. TP → The test correctly identifies a person who has the disease.

2. FP → The test incorrectly identifies a person as having the disease when they do not.

3. TN → The test correctly identifies a person who does not have the disease.

4. FN → ~~The test~~ A person has the disease but the test incorrectly identifies them as not having it.

3. Precision

It is used to overcome the limitation of Accuracy.

$$\text{Precision} = \frac{TP}{TP + FP}$$

~~(specificity)~~

4. Recall :

It aims to calulate the proportion of actual positive that was indentified incorrectly.

$$\text{Recall} = \frac{TP}{TP+FN}$$

5. F- Scores :

> It is a metric to evaluate a binary classification model on the basis of predictions that are made for the positive class.

$$F1\text{-}score = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}}$$

6. AUC - ROC

* True positive Rate

* False " "

$$TPR = \frac{TP}{TP+FN}$$ (recall)

$$FPR = \frac{FP}{FP+TN}$$

* True Negative Rate

* False " "

$$TNR = \frac{TN}{FP+TN}$$

$$FNR = \frac{FN}{FN+TP}$$ ~~TP+TN~~

7. Error rate $= \dfrac{FN + FP}{FN+FP+TN+TP}$

(or)

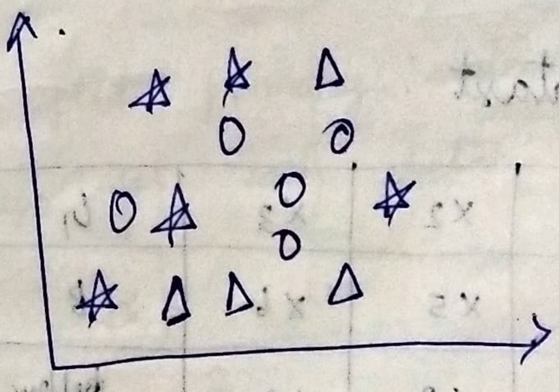
" " = 1 - (Accuracy rate).

Multiclass classification:

Task of classifying elements into different classes.

Eg:
* Books according to subject
* sentiment analysis.

Algorithms:
* KNN
* Decision tree
* Bayes
* Random Forest.



one vs All :

Training one vs Rest classifiers for our model involves three binary classifiers.

Medical diagram :

not ill, cold , fever

y = 1 2 3

eg :

class 1 : Green

" 2 : Blue

" 3 : yellow

Main dataset :

x1	x2	x3	G
x4	x5	x6	B
x7	x8	x9	yellow
features			classes

Training ① green

x1	x2	x3	+1	→ one vs rest.
x4	x5	x6	-1	→ Not taken
x7	x8	x9	-1	

same for ② & ③.

one vs one.

In OVO for n class dataset we have to generate

$$\frac{n*(n-1)}{2}$$ binary model.

We split primary dataset into binary for each data.

$$\frac{n*(n-1)}{2}.$$

$$n=3 = \frac{3*(3-1)}{2} = 3.$$

Eg :

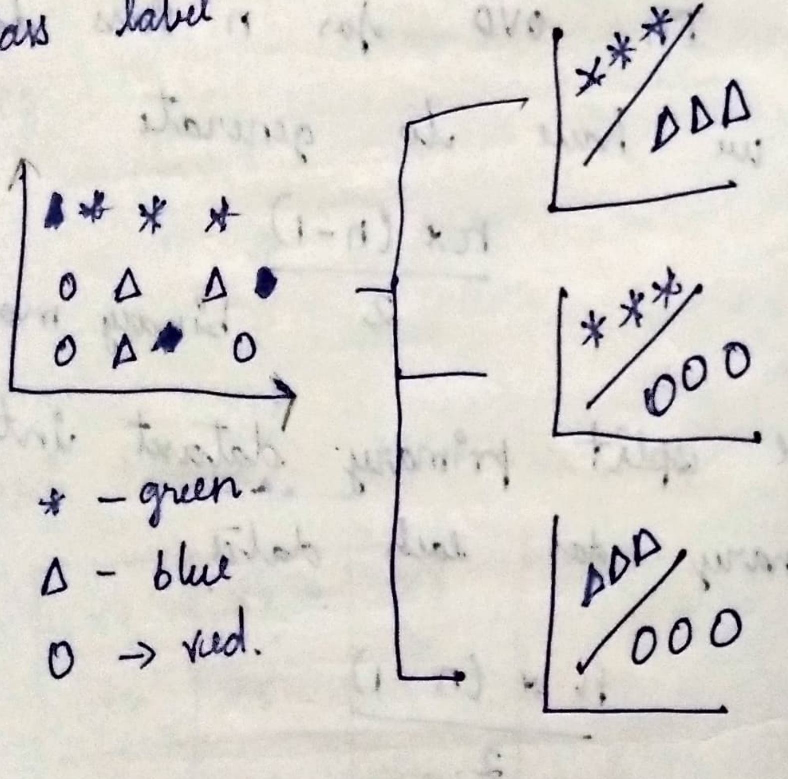
C1 = Green vs Blue

C2 = Green vs Red

C3 = Blue vs Red

$$\frac{N * (n-1)}{2} = 3.$$

Each binary classifier predict one class label.



T- Test :

1) Single Mean = $\dfrac{\bar{x} - \mu}{s / \sqrt{n}}$

dof = n-1.

2) Two mean = $\dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n1} + \dfrac{s_2^2}{n2}}}$

$\dfrac{\bar{x}_1 - \bar{x}_2}{sp \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

dof = n1 + n2 - 2.

3) paired sample t-test (or) correlation t-test

$t = r \sqrt{\dfrac{n-2}{1-r^2}}$

dof = n-2

1. We want to test if the average weight of apples is 150 grams. We take sample of 5 apples with weights: 148, 152, 149, 150, 151.

given,

$n = 5.$

$\mu = 150$

$$t = \frac{\bar{x} - \mu}{s / \sqrt{n}}.$$

$$\bar{x} = \frac{148 + 152 + 149 + 150 + 151}{5}$$

$$= \frac{750}{5}$$

$$= 150$$

B O D M A S

() 2^2 / × + −

$$s = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$$

$$= \sqrt{\frac{(148-150)^2 + (152-150)^2 + (149-150)^2 + (150-150)^2 + (151-150)^2}{5-1}}.$$

$$= \sqrt{\frac{(-2)^2 + 2^2 + (-1)^2 + 0^2 + 1^2}{4}}$$

$$S = \sqrt{\frac{10}{4}} = \sqrt{2.5} \approx 1.58$$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{150 - 150}{1.58/\sqrt{5}} = 0$$

2.132

Reject null hypothesis.

(2) ⇒ We want to compare the average heights of two groups of plants.

Group A : 18, 20, 22, 19, 21

" B : 17, 18, 20, 19, 18

soln:

$$\bar{x}_1 = \frac{18 + 20 + 22 + 19 + 21}{5} = 20.$$

$$\bar{x}_2 = \frac{17 + 18 + 20 + 19 + 18}{5} = 18.4$$

$s_1^2 = \sqrt{2.5} \approx 1.58$

$s_2^2 = \sqrt{1.3} \approx 1.14$

$$t = \frac{20 - 18.4}{\sqrt{\frac{(1.58)^2}{5} + \frac{(1.14)^2}{5}}}$$

$\therefore 1.58^2 = 2.4964$

$1.14^2 = 1.2996$

$$= \frac{1.6}{\sqrt{0.49928 + 0.25992}}$$

$$= \frac{1.6}{\sqrt{0.7592}}$$

$$= \frac{1.6}{0.8713}$$

$$\sqrt{\frac{(n_1-1).s_1^2 + (n_2-1).s_2^2}{n_1+n_2-2}}$$

$$sp = \sqrt{\frac{(5-1).(1.58)^2 + (5-1).(1.42)^2}{5+5-2}}$$

$$= \sqrt{1.91} \approx 1.38$$

$$t = \frac{\bar{x_1} - \bar{x_2}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{20 - 18.4}{1.38\sqrt{\frac{1}{5} + \frac{1}{5}}}$$

$$= \frac{1.6}{1.38\sqrt{0.4}}$$

$$= \frac{1.6}{1.38\,(0.632)}$$

$$= \frac{16}{0.872}$$

$$\approx 1.83$$

$$dof = n_1 + n_2 - 2$$

$$= 5 + 5 - 2$$

$$= 8.$$

Table value = 2.306.

$2.306 > 1.83$.

Accepts H0.

③ ⇒ we want to test if there is a significant correlation b/w study hours and test scores for 5 students. The correlation coefficient r is calculated to be 0.85

$\therefore dof = n-2$
$= 5-2 = 3$

soln:

$r = 0.85$ $\quad r^2 = 0.7225.$

$n = 5.$

$$t = r\sqrt{\frac{n-2}{1-r^2}}$$

$$= 0.85\sqrt{\frac{5-2}{1-0.7225}}$$

$$= \frac{0.85\sqrt{3}}{0.528.}$$

$$= \frac{0.85 \cdot 1.732}{0.528.}$$

reject H0 //

$$= \frac{1.472}{0.528} \approx 2.79$$

MC Nemar's test :

> Non parametric test for paired Nomial data :

> compares performance of two classifiers on n items from single test set.

df = 1

eg :

pairs matched on age & sex .

Diabets	M. Infection		Total
	yes	NO	
yes	10 (A)	10 (B)	20
NO	10 (C)	30 (D)	40
Total	20	40	80

$$Q = \frac{(|B-C| - 1)^2}{B+C}$$

$$= \frac{(|10-10| - 1)^2}{20} = \frac{1}{20} = 0.05$$

T.V = $\boxed{3.841}$ > 0.05 => Accept H0.