Validation metrics:-

Classification?

① **Accuracy:-**

measures the proportion of correct predictions made by the model, out of all predictions.

$$Accuracy = \frac{No. \ of \ correct \ predictions}{Total \ no. \ of \ prediction}$$

$$= \frac{TP + TN \ (True \ positive + True \ Negative)}{TP + TN + FP + FN}$$

---

② **Precision:-** measures the accuracy of positive predictions.

$$precision = \frac{TP \ (True \ Positives)}{TP + FP \ (True \ positives + False \ positives)}$$

3. Recall or Sensitivity:-

measure model ability to correctly identify all positive cases.

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True positive (TP)} + \text{False Negatives (FN)}}$$

---

(4) F1 Score:-

measure used in classification tasks to provide a balance b/w precision

$$\text{F1 Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

F1 Score is the harmonic mean of precision and recall.

⑤ ROC - AUC Score:-

- Performance metric used in **Binary** **classification** problems.

- Measures the models ability to distinguish b/w positive and negative classes.

① ROC Curve?

* The ROC curve plots the True positive rate against False positive rate, at various threshold values.

True positive Rate (TPR) = $\dfrac{\text{True positives}}{\text{True positive + False Negatives}}$
(Recall or Sensitivity),

False positive Rate (FPR) = $\dfrac{\text{False positives}}{\text{False positives + True Negatives}}$

$$FPR = 1 - \text{Specificity}.$$

② Area Under Curve (ROC)

- It ranges from 0 to 1 and provides a single score that summarizes the model performance.

- An AUC of 1.0 indicates perfect discrimination b/w the classes, while an AUC of 0.5 means the model performs no better than random guessing.

---

⑥ . Specificity: (True Negative Rate):-

- models ability to correctly ~~measure~~ identify the negative instances.

$$Specificity = \frac{True\ Negatives}{True\ Negatives + False\ positives}$$

⑦. Confusion Matrix:

		Predicted:-	
		P	N
Actual	P	True Positive	False Negative
	N	False positive	True Negative

⑧. Log-loss:-

• performance metric for evaluating accuracy of a classification model that outputs probability predictions.

$$Logloss = -\frac{1}{N} \sum_{i=1}^{N} \left[y_i \log(p_i) + (1-y_i) \log(1-p_i) \right]$$

N is the no. of instances.

y_i is actual label.

p_i is predicted probability. of instance being in the class.