# Types of datasets

Training Dataset
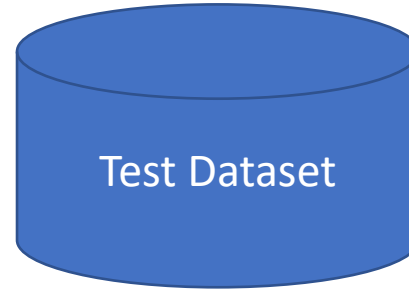
Validation Dataset

Test Dataset

**All these three datasets are labelled datasets.**

**Training dataset** is used for building the models.

**Validation dataset** is used for evaluating the intermediate model, while building the model.

**Testing dataset** is used for evaluating the final model.

# Evaluation

Assumptions:

No of samples : N

No of classes : +ve and –ve

No of samples in +ve class: P

No of samples in -ve class : N-P

|  | Observed +ve | Observed -ve |  |
|---|---|---|---|
| Actual +ve | TP | FN | P |
| Actual -ve | FP | TN | N - P |
|  | p | N-p |  |

TP: True Positive
FN: False Negative
TN: True Negative
FP: False Positive

# Evaluation

|  | Observed +ve | Observed -ve |  |
|---|---|---|---|
| Actual +ve | TP | FN | P |
| Actual -ve | FP | TN | N - P |
|  | p | N-p |  |

$$precision = \frac{TP}{TP + FP}$$

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

$$F1\ measure = \frac{2.Precision.Recall}{Precision + Recall}$$

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

.  .