

DS.

1. Cross Validation.

It is a technique used to evaluate the performance of a machine learning model by splitting the dataset into multiple parts (folds), training the model on some parts and testing it on the remaining part.

It helps reducing overfitting and gives more reliable estimate of model performance compared to single train-test split.

Types of Cross Validation...

1. Hold out.

Dataset is split into train and test.

eg. 70% training 30% testing

2. K fold.

Dataset is divided into K equal folds, and tested on remaining fold.

3. Stratified K fold.

Similar to K fold but ensures each fold has same proportion.

1. man \rightarrow woman (FN)
2. man \rightarrow man (TP)
3. woman \rightarrow woman (TN)
4. man \rightarrow man (TP)
5. man \rightarrow woman (TN)
6. woman \rightarrow woman (TN)
7. woman \rightarrow man (FP)
8. man \rightarrow man (TP)
9. man \rightarrow woman (FN)
10. woman \rightarrow woman (TN)

Confusion Matrix. Predicted

Actual Man TP = 3

Actual Woman FP = 1

Predicted
woman

FN = 1

TP = 3

Metric

$$\text{Accuracy} = \frac{3+3}{3+3+1+3} = \frac{6}{10} = 0.6 (60\%) //$$

$$\text{Precision} = \frac{3}{3+1} = \frac{3}{4} = 0.75 //$$

$$\text{Recall} = \frac{3}{3+3} = \frac{3}{6} = 0.5 //$$

$$\text{Result} = [[3, 3], [1, 3]]$$