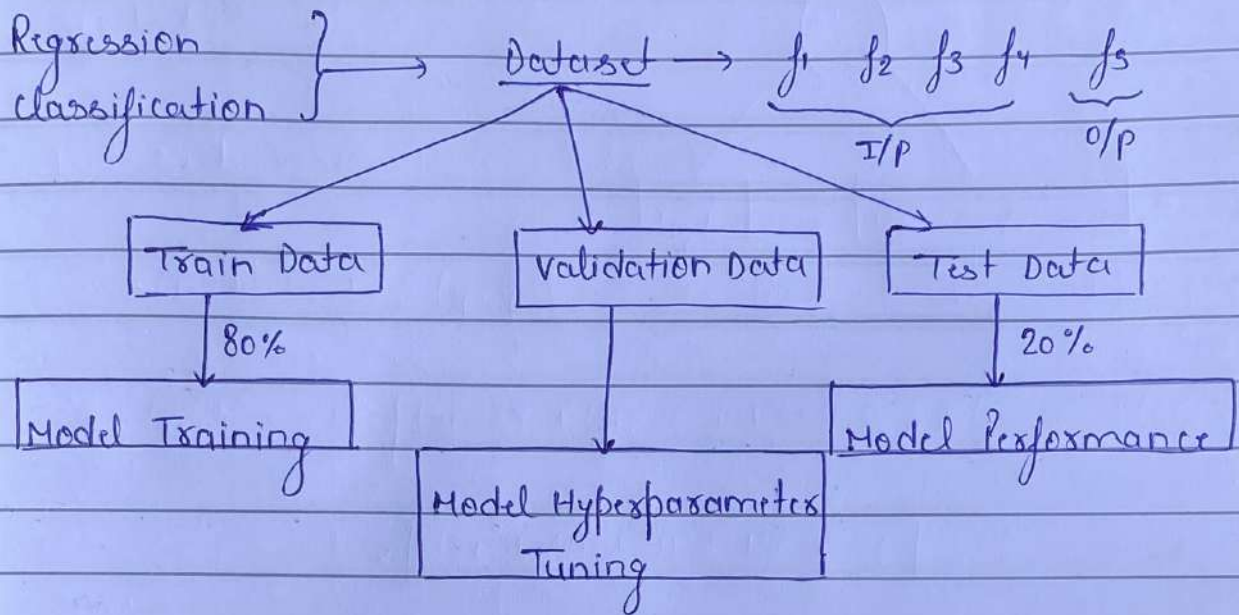


## Training Data, Test Data and Validation Data



\* Cross Validation :  $\rightarrow$  Hyperparameters tuning

Dataset : { 10,000 records }

80%  
Training Data  
8000

20%  
Test Data  
2000

let  $CV = 5$

$$\frac{8000}{5} = 1600$$

cv1 | 1600 | 6400 |  $\rightarrow Acc_1$

cv2 | | 1600 |  $\rightarrow Acc_2$

cv3 | | | 1600 |  $\rightarrow Acc_3$

cv4 | | | 1600 |  $\rightarrow Acc_4$

cv5 | | | 1600 |  $\rightarrow Acc_5$

Every cross validation

$$Train = (8000 - 1600) = 6400$$

$$Test = 1600$$

$\uparrow$  validation

$$Accuracy = \frac{Acc_1 + Acc_2 + Acc_3 + Acc_4 + Acc_5}{5}$$