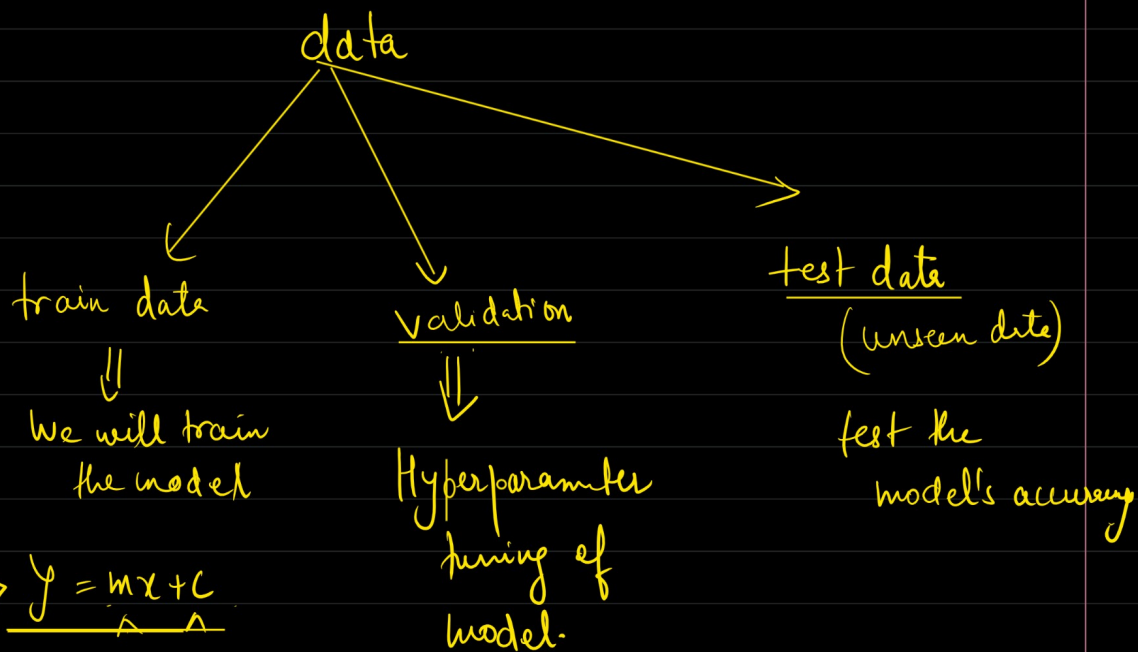


Train, Validation and test data

data

#Hours studies	#marks in Internal exam	Pass / fail.
5	80	1
3	70	1
1	40	0
2	45	1
0.5	20	?



$$m = \frac{y_2 - y_1}{x_2 - x_1}$$

$$\Rightarrow y = mx + c$$

dataset

1000 rows.

60%

↓
600 rows

↓
train the model

20%

↓
200 rows

↓
tune the hyperparameters

20%
(test data)

↓
200 rows

↓
to test the performance of model.

for training and validation phase, test data will never be used.

eg.

