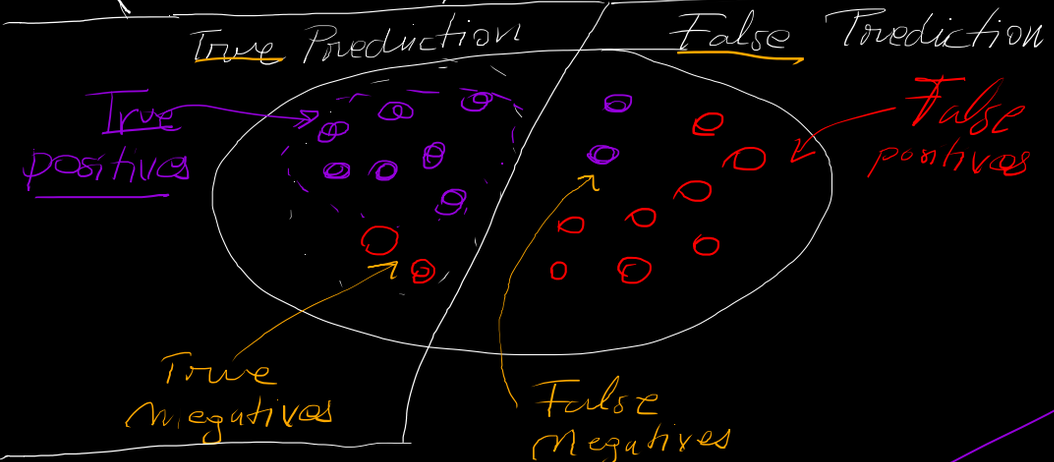
Python VII

1) Metrics: Accuracy (Binary Classifier)



→ Accuracy: $\frac{Tp}{total}$

→ Precision: $\frac{Tp}{Tp+Fp}$

→ Recall: $\frac{Tp}{Tp+Tn}$

True Positives / total

ex: $\frac{20}{100} = 0,2$

4 Data: ○○○○ 75% → 90%

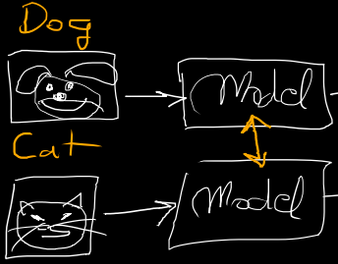
●●●●●●●●●● 50% → 10%

* Binario 2 } 50% // → 10%

True Positives / true posit + false pos

True Positive / true positive + true negative

2) Probability Threshold

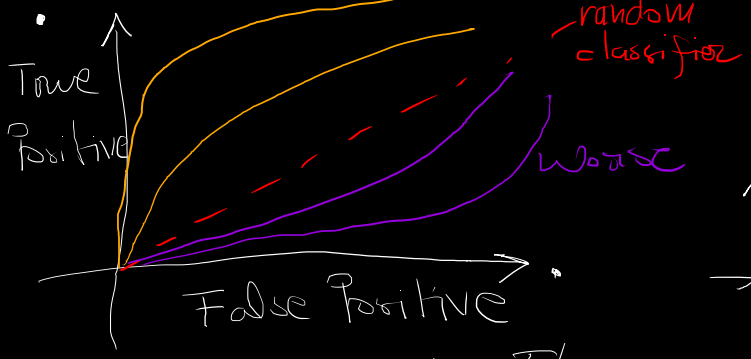


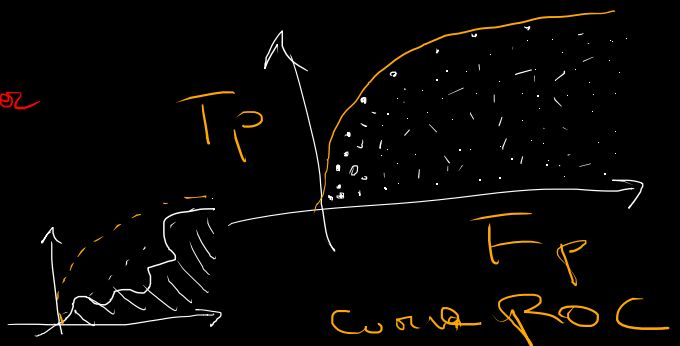
Probability Cats (Model!!)

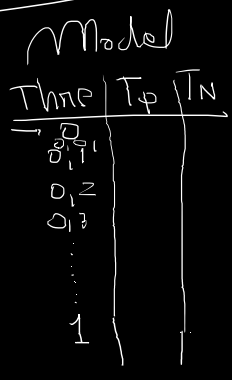
③ ROC (Receiving Operation Characteristic)

IN → Model → OUT
Prob

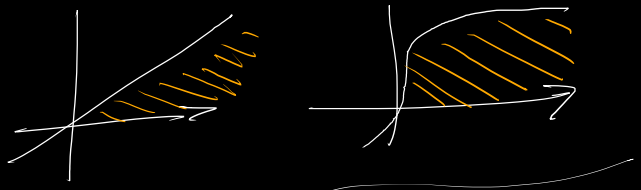
Variable: Threshold!



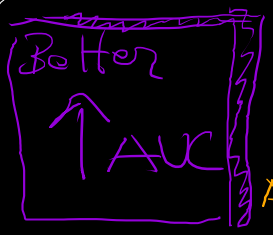


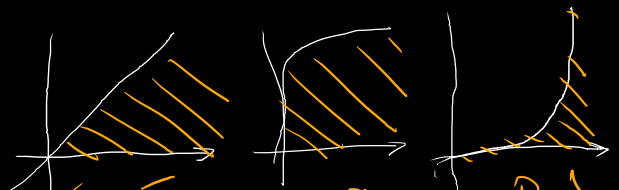


* Area under the curve:



AUC ROC





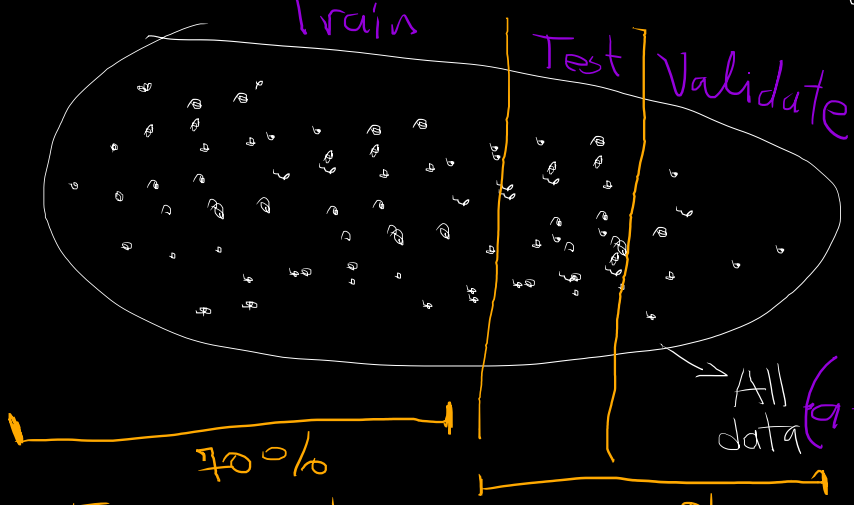
④ Real Life Data Set:

Dataset size: 5

Per category: >60.000

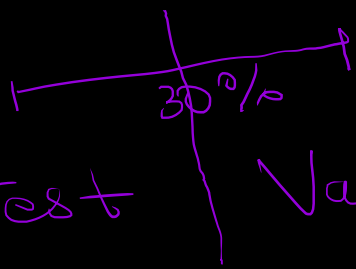
- Better Models, Few data
- How handle Big Data!

Big Data = BD

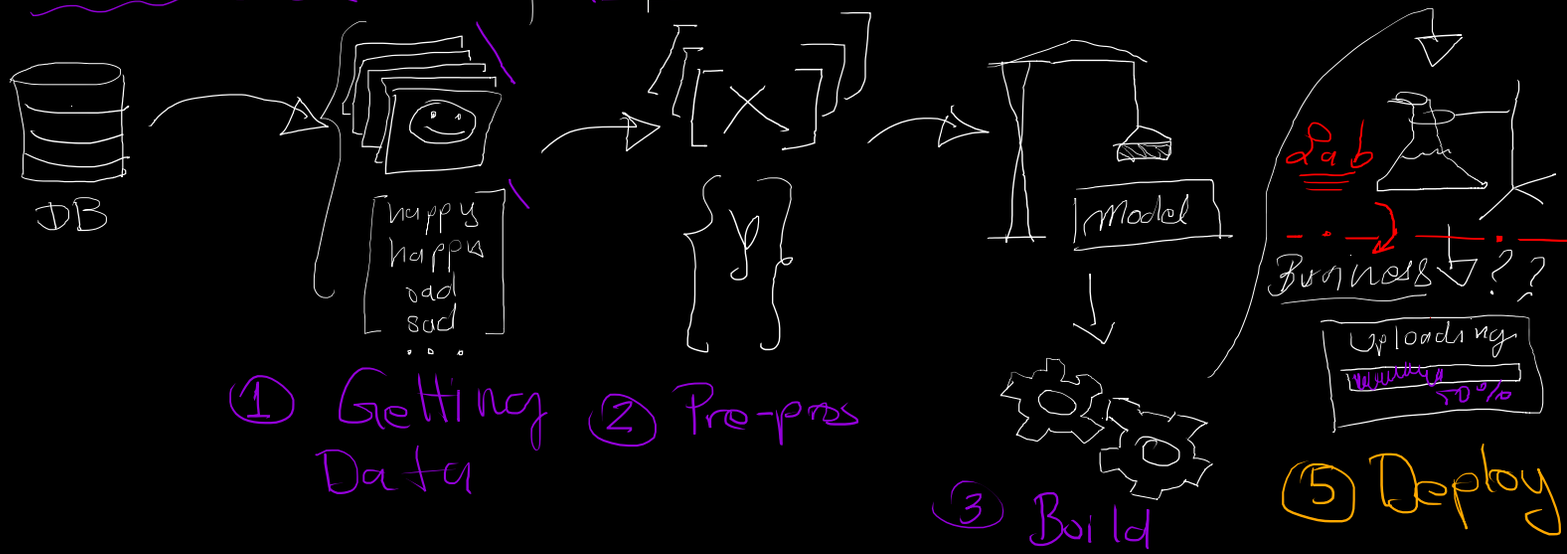


Train Data

Test Data



*Practice : Real Life Pipeline AI/ML



① Getting Data ② Pre-pros ③ Build ④ Train ⑤ Deploy

* MNIST database: [1] [2] [3] → DB handwritting digits (Echo)

→ Use Tensorflow Mnist

→ Using other images

[0] scalar 0-D

[0000] vector 1-D

Matrix 2-D

3D

4D

→ Tensorflow !!!

* Train Data:

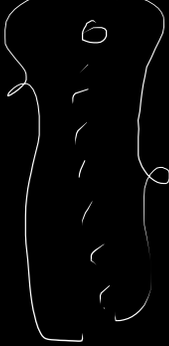
[0] [1] [2] ....

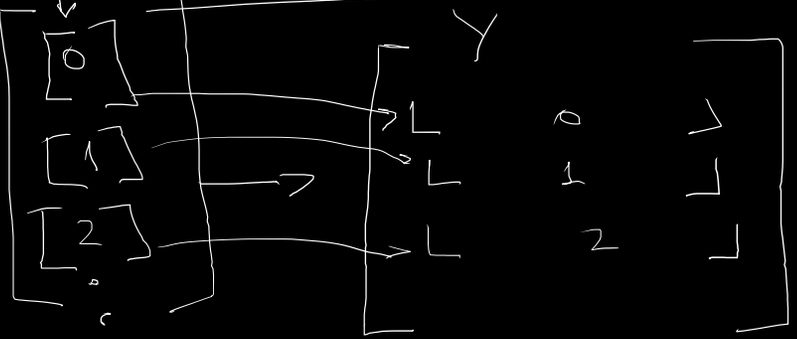
* Labels

[0, 1, 2, 9, ...]

X fation for [1, 0, 0, 0 ....]

[6] =





$2^N = 2, 4, 8, 16, 32 \ldots \ldots 1024 \ldots$

DFT

FFT 2^k