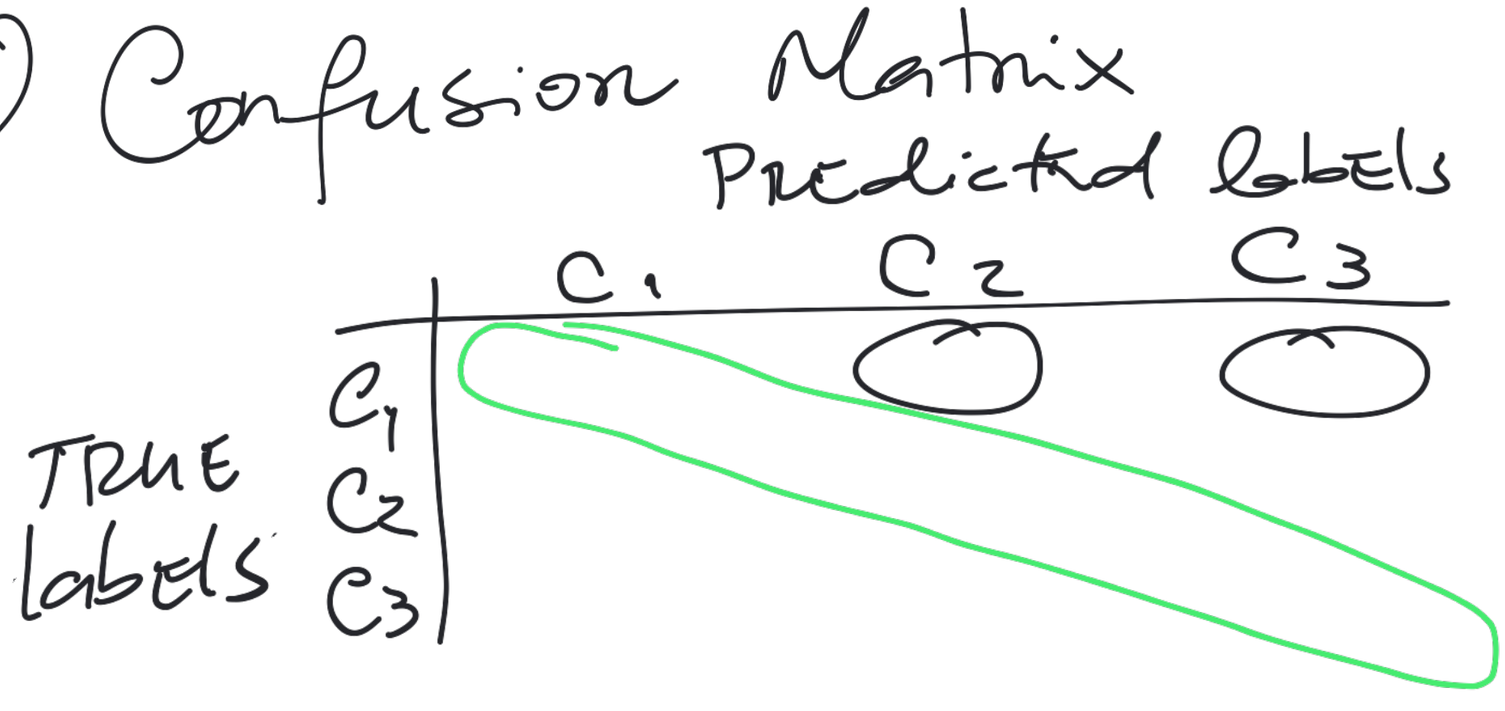
Performance Metrics for classification:

(1) Accuracy OR ERROR

$$\frac{N_{correctly\ classified}}{N}, \quad \frac{N_{misclassified}}{N}$$

Drawback: do not inform about individual class performance.

(2) Confusion Matrix



Each class has:

③ • TRUE positives (TP)
• TRUE NEgatives (TN)
• FALSE positives (FP)
• FALSE NEgatives (FN)

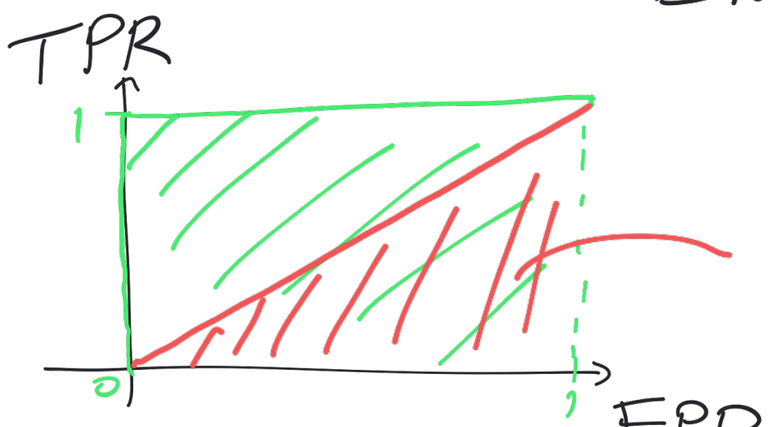
NOTE: ROC CURVE is computed PER class.

Binary PERFORMANCE metric

④ ROC CURVE

$$FPR = \frac{FP}{FP+TN}$$

$$TPR = \frac{TP}{TP+\underbrace{FN}_{=0}}$$

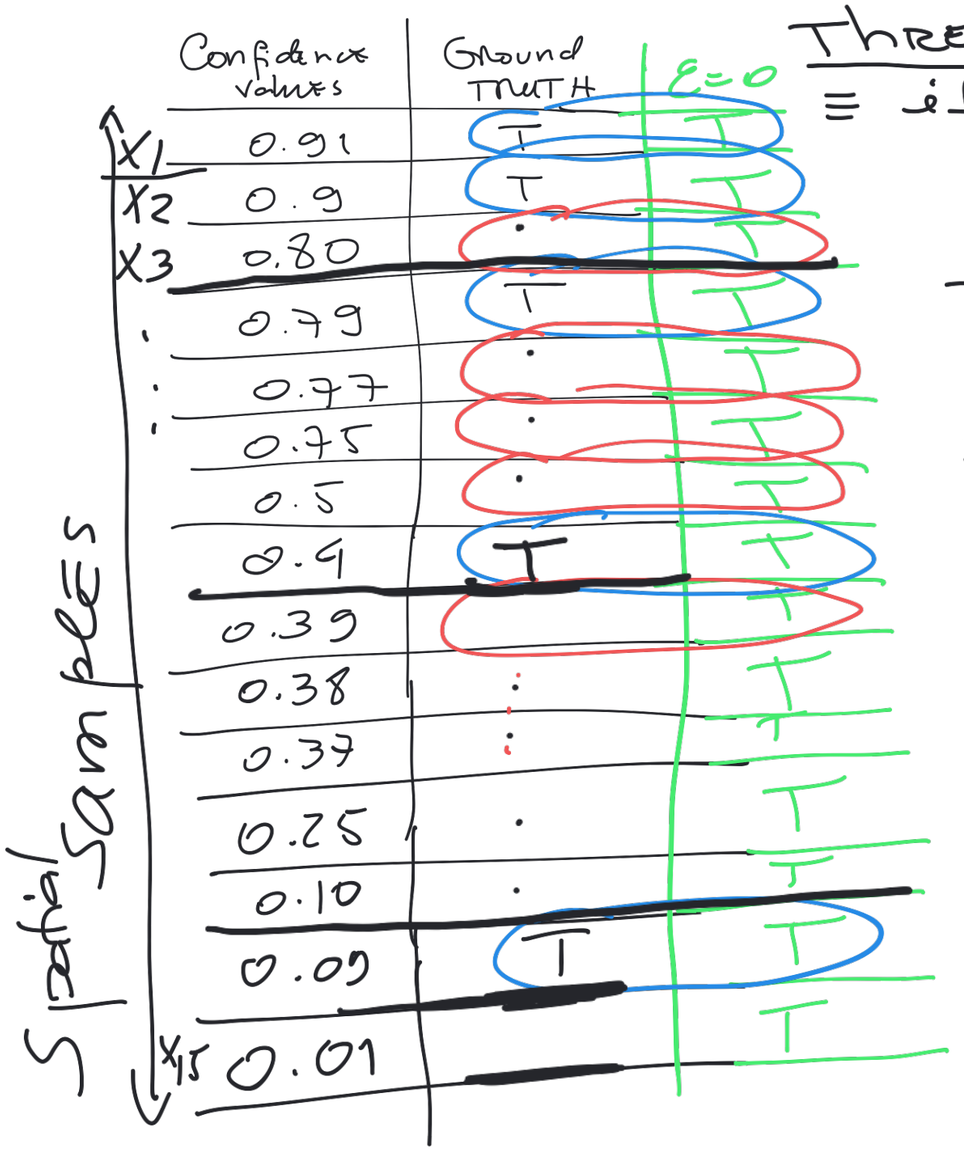


$$A = \frac{1}{2} = \frac{b \times h}{2}$$

⑤ AREA UNDER ROC CURVE.

A = 1 for optimal ROC CURVE

We want $A(ROC) > \frac{1}{2}$

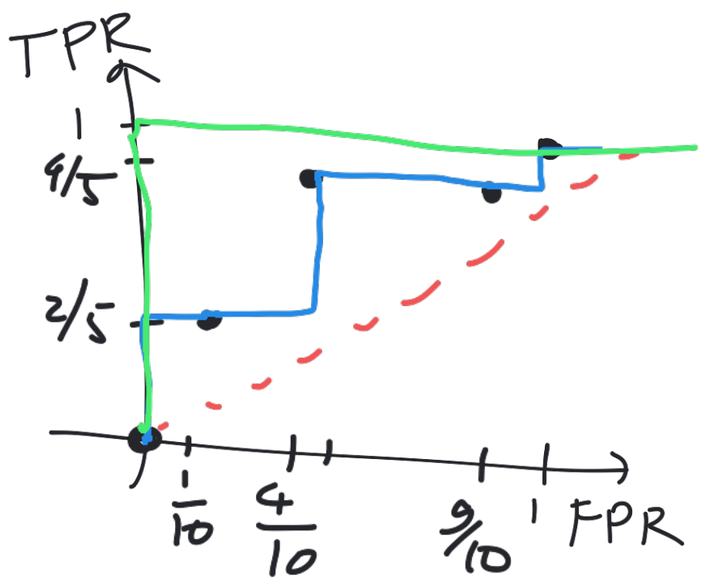


Threshold = 0

≡ if confidence value > 0 then classify as T

$$TPR_{\varepsilon=0} = \frac{5}{5+0} = 1$$

$$FPR_{\varepsilon=0} = \frac{FP}{FP+TN} = \frac{10}{10+0} = 1$$



$\varepsilon = 0.1$

$TPR = \frac{4}{5}$

$FPR = \frac{9}{10}$

$\varepsilon \geq 0.4$

$TPR = \frac{4}{5}$

$FPR = \frac{4}{10}$

$\varepsilon \geq 0.8$

$TPR = \frac{2}{5}$

$FPR = \frac{1}{10}$

$\varepsilon \geq 1$

$TPR = 0$

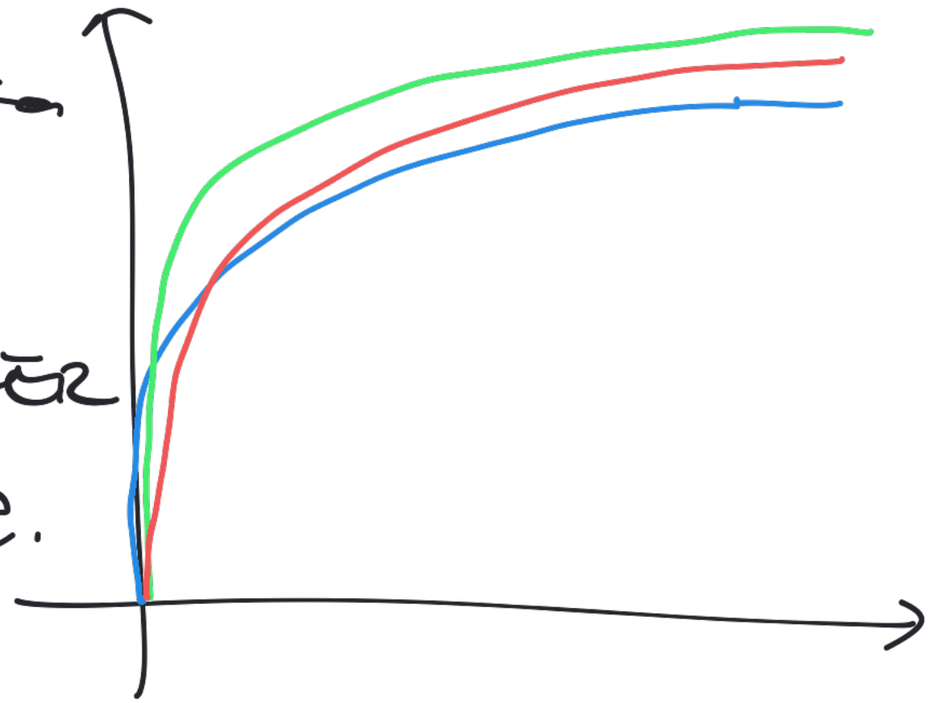
$FPR = 0$

ROC Curve:

→ Binary measure

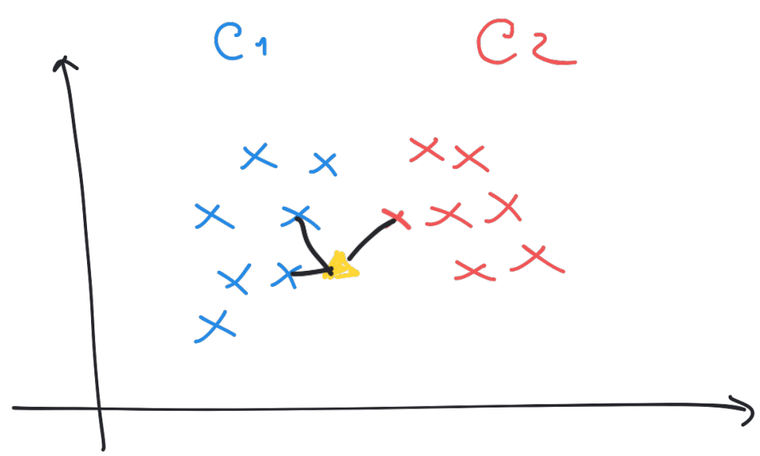
↳ it measures one-vs-all.

→ Compute the area under each ROC curve and consider the avg. area ROC.



K - NEAREST NEIGHBOR (KNN) Algorithm

- non-parametric approach for classification



$\{x_i, t_i\}_{i=1}^{N}$

↑ 2-Dimensional

▲ : test point

$K = 3$

→ Finds labels for K neighbors.

→ neighbors: is such that it has smallest distance

↳ Distance metric can vary.

→ Label for ▲ is based <u>majority vote</u>

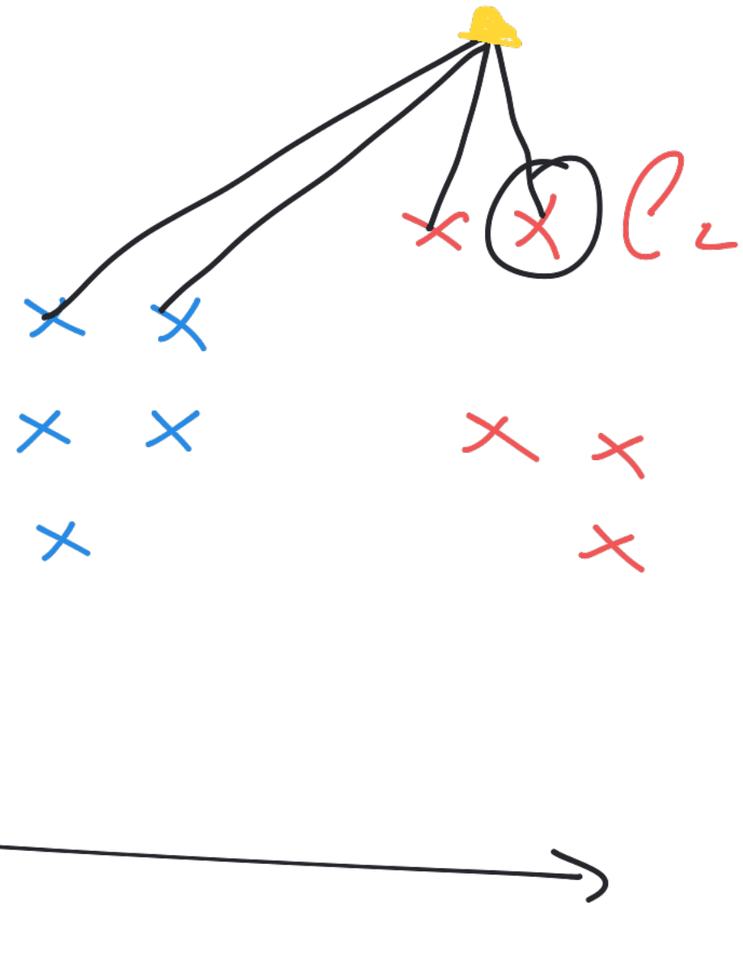
▲ has 2 neighbors for C_1 → ▲ $\in C_1$
1 neighbor for C_2

When we have ties?

① "Flip a coin".

② Decide class with closest neighbor

③ Use <u>odd value</u> for K.



→ We can also consider different distance metrics.

Q: Is K-NN sensitive to data scaling?

YES!

We need scale/normalize data.