ROC - curve

Example I - Logistic Regression for linearly separable test dataset!

In LR, we model $P(Y=1 \mid x)$ using a sigmoid function such that:

$$P(\hat{Y}=1 \mid x) = \sigma(f(x)) = \frac{1}{1+e^{-f(x)}}$$

In linear logistic regression:

$$f: \mathbb{R}^d \to \mathbb{R}$$
$$\bar{x} \to \langle \bar{w}, \bar{x} \rangle$$

→ Training logistic regression:
Using trainin dataset $D_n = (\bar{x}_i, y_i)_{i=1}^n$, we find the weight vector w_n^* that minimizes the logistic loss, or equivalently, maximizes the likelihood evaluated on the training data. (Lecture 7).

→ Evaluating the performance of a classifier!

- We assume that we have trained a LR classifier, which is fully defined by the function $f(\bar{x}) = \langle \bar{w}^*, \bar{x} \rangle$

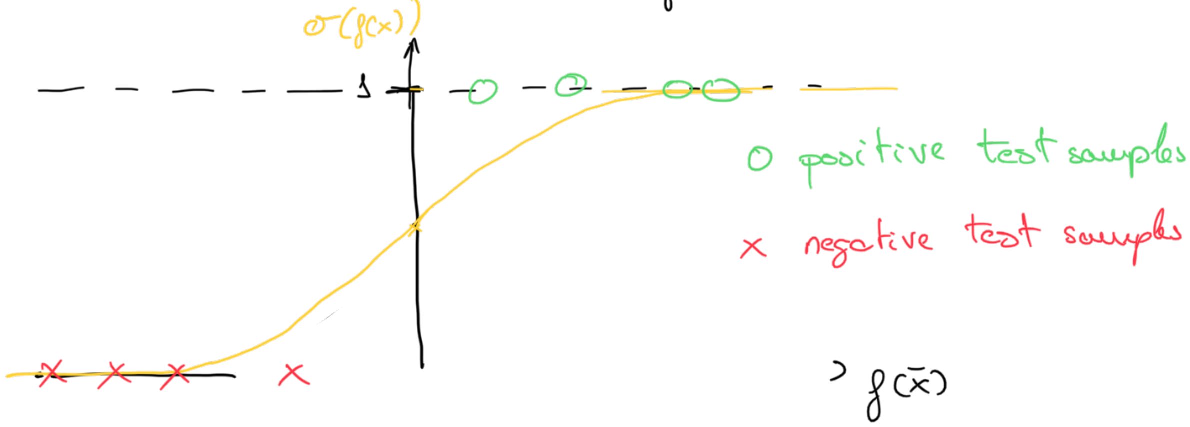
- We further assume access to a test dataset $D_{Test} = (\bar{x}_i, y_i)_{i=1}^m$, which contains m i.i.d samples of the measure P over $\mathcal{X} \times \mathcal{Y}$.

- So far, we assumed tha we classify as:

$$\hat{y}(\bar{x}) = \operatorname{sign} f(x)$$

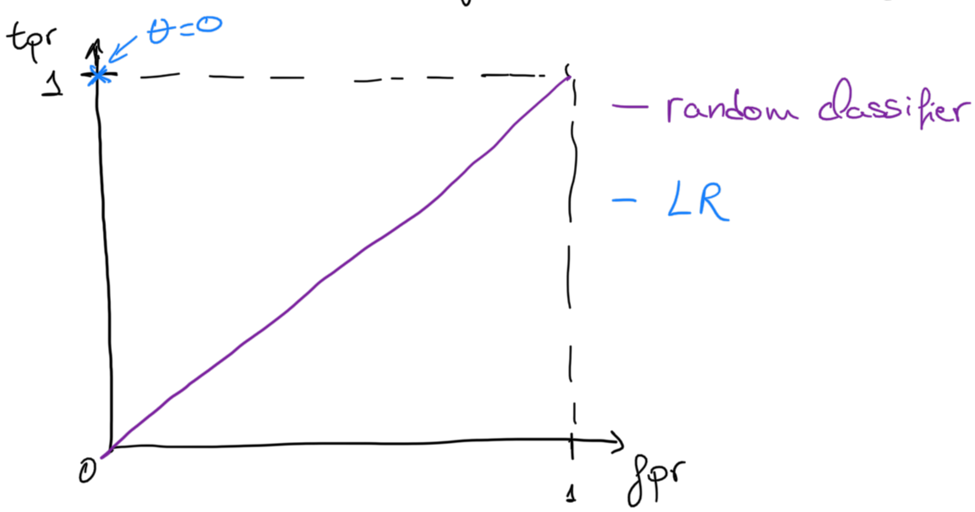
Note that the above classification rule is equivalent to

$$\hat{y}(\bar{x}) = \begin{cases} +1 & \text{if } \sigma(\bar{x}) \geq 0'5 \\ -1 & \text{if } \sigma(\bar{x}) < 0'5 \end{cases}$$



Assuming a classification threshold $\theta = 0$ (i.e. $\hat{y} = \text{sign}\, f(x)$) we obtain that Error $= 0 \Rightarrow D_{test}$ is <u>linearly separable</u>

Question: How do we get the ROC-curve?



For $\theta = 0 \Rightarrow fpr = 0$ & $tpr = 1$