

# Courbes ROC et Precision Recall

Pour évaluer la qualité d'un modèle

Sensitivity et Specificity I:

Désion			
Appart		T+	T-
	w <sub>1</sub>	(TP)	(FN)
	w <sub>2</sub>	(FP)	(TN)

*homme + wataushah* (pointing to TP)

*homme + amek labo* (pointing to FN)

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad \text{Zphers}$$

$$\text{Specificity} = \frac{TN}{TN + FP}$$

1. True Positive + False Negative = Positive cases = #w<sub>1</sub>
2. True Negative + False positive = Negative cases = #w<sub>2</sub>



TPR = pourcentage de vrais positifs = sensitivity

TNR = pourcentage de vrais négatifs = specificity

FPR = pourcentage de faux positifs =  $\frac{\text{faux} +}{\text{cas} -}$

FNR = pourcentage de faux ~~-~~ =  $\frac{\text{faux} -}{\text{cas} +}$

$$\frac{\text{TPR}}{\text{PV}+} + \frac{\text{FNR}}{\text{PF}-} = \frac{\text{TNR}}{\text{V}-} + \frac{\text{FPR}}{\text{F}+} = 1$$

Exemple :

		T+	T	precision
1	w <sub>1</sub>	140 (T+)	60 (FN.)	200
0	w <sub>0</sub>	100 (F+)	900 (T-)	1000
Total		240	960	specificité

$$\text{TPR} = \frac{140}{200} = 0,7$$

$$\text{FNR} = \frac{60}{200} =$$

$$\text{TNR} = \frac{900}{1000} = 0,9$$

$$\text{FPR} = \frac{100}{1000}$$

Accuracy =



$$P(\omega_1) = 0,17 = \frac{200}{1200}$$

$$P(\omega_2) = \frac{1000}{1200}$$

$$\text{Precision} = \frac{TP + TN}{TP + TN + FP + FN} \quad \text{Interpret}$$

## Cas du Scoring :

Score  $S$  : à tout  $x$  incluse  $\rightarrow$  nb  $S(x)$

↓  
permet  
son affectation  
à l'une  
des classes

Distribution du score  $S$  conditionnellement  
aux classes

Si  $S \downarrow$  FPR et TPR i.e. (1-specificity)  
et sensibility  $\nearrow$

**Courbe ROC** (FPR, TPR) lorsque  $S$  varie

- Un score  $S$  est meilleur que un autre  $S'$   
si la courbe  $S$  est tjrs au dessus de celle  
de  $S'$

**- AUC** (aire sous la courbe), mesure les  
performances du score :  
coeff de Gini =  $2AUC - 1$   
 $AUC = P(X_1 > X_2)$

$X_1$  et  $X_2$  suivent la loi de score  $S$

AUC  $\rightarrow$  proba que le score d'un indiv  $x$  soit supérieur à celui d'un indiv  $y$ ,  
si  $x$  est tiré aléa parmi les  $+$  et  $y$   
est tiré aléa parmi les  $-$ .

## Precision et Recall

si on a tiré  
sur les classés  $+$

Recall

$$= \frac{\text{True } +}{T+ + F+}$$

Precision

$$= \frac{T+}{T+ + F+}$$

Precision : pourcentage des vrais  $+$   
parmi les prédits  $+$