

Les données numériques: toutes des nombres!

Métriques de classification

Diane Lingrand



2022 - 2023

- sounds classification into 3 classes : 'cat', 'dog' and 'bird' from the Speech Commands Dataset
 - 200 sounds per class for training, 200 other sounds per class for testing
 - 1 sound is described by a single MFCC vector (duration of the sound)
 - classification algorithm is kNN
 - which k value?
- How to evaluate the classification ?
 - qualitative evaluation
 - quantitative evaluation : compute metrics on the test set

Evaluation d'une classification binaire : positifs et négatifs

vrais positifs (VP) : données positives calculées comme positives

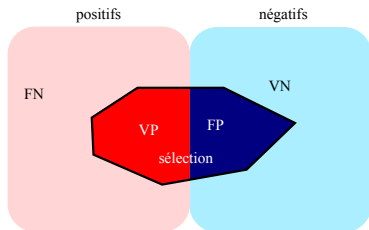
vrais négatifs (VN) : données négatives calculées comme négatives

faux positifs (FP) : données négatives calculées comme positives

faux négatifs (FN) : données positives calculées comme négatives

Matrice de confusion :

		classes estimées	
		positif	négatif
classes réelles	positif	9	1
	négatif	1	9



Vrais faux positifs négatifs

vrais positifs (VP) : données positives calculées comme positives

vrais négatifs (VN) : données négatives calculées comme négatives

faux positifs (FP) : données négatives calculées comme positives

faux négatifs (FN) : données positives calculées comme négatives

précision (*accuracy*) : $\frac{VP+VN}{P+N} = \frac{VP+VN}{VP+FP+VN+FN}$

sensibilité, rappel (*recall*, taux de VP : $\frac{VP}{VP+FN}$

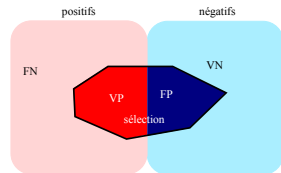
spécificité, sélectivité, taux de VN : $\frac{VN}{VN+FP}$

précision (*precision*), v. de préd. pos. : $\frac{VP}{VP+FP}$

valeur de prédiction négative : $\frac{VN}{VN+FN}$

F-mesure (*F1 score*) :

$$2 \frac{\text{précision} \cdot \text{rappel}}{\text{précision} + \text{rappel}} = \frac{2VP}{2VP+FP+FN}$$



- précision *accuracy* : fraction des prédictions correctes
- mesure F1 : plusieurs options sont disponibles, entre autres :
 - 'micro' : calcul global des VP, FP, FN
 - None : une valeur de score F1 est calculée pour chaque classe
 - classe considérée comme positive, les autres classes étant alors considérées comme négatives
 - 'macro' : calcul du score pour chaque classe et moyenne de ces scores
- rappel *recall* : mêmes options que pour F1

Quelques ordres de grandeur : cas des classes équilibrées

- classes équilibrées : chaque (vraie) classe contient le même nombre de données
- classification binaire
 - classification aléatoire : 1 chance sur 2 d'être positif

Quelques ordres de grandeur : cas des classes équilibrées

- classes équilibrées : chaque (vraie) classe contient le même nombre de données
- classification binaire
 - classification aléatoire : 1 chance sur 2 d'être positif
 - $VP = FP = VN = VP = 0.5 =$ toute métrique

- exemple de classification :

		classes estimées	
		positif	négatif
classes réelles	positif	80	20
	négatif	10	90

Quelques ordres de grandeur : cas des classes équilibrées

- classes équilibrées : chaque (vraie) classe contient le même nombre de données
- classification binaire
 - classification aléatoire : 1 chance sur 2 d'être positif
 - $VP = FP = VN = VP = 0.5 =$ toute métrique

- exemple de classification :

		classes estimées	
		positif	négatif
classes réelles	positif	80	20
	négatif	10	90

- $VP = 80, VN = 90, FN = 20, FP = 10$

Quelques ordres de grandeur : cas des classes équilibrées

- classes équilibrées : chaque (vraie) classe contient le même nombre de données
- classification binaire
 - classification aléatoire : 1 chance sur 2 d'être positif
 - $VP = FP = VN = VP = 0.5 =$ toute métrique

- exemple de classification :

		classes estimées	
		positif	négatif
classes réelles	positif	80	20
	négatif	10	90

- $VP = 80, VN = 90, FN = 20, FP = 10$
- $\text{rappel} = 0.8, \text{accuracy} = 0.85, F1 = 16/19 = 0.84$

- classification multi-classes
 - exemple : cas de 3 classes équilibrées
 - classification aléatoire : 1 chance sur 3 d'être de la classe 0

- classification multi-classes
 - exemple : cas de 3 classes équilibrées
 - classification aléatoire : 1 chance sur 3 d'être de la classe 0
 - $\text{accuracy} = 1/3$, $\text{rappel}(\text{micro}) = 30/90 = 1/3$, $\text{F1}(\text{micro}) = 1/3$

- classification multi-classes
 - exemple : cas de 3 classes équilibrées
 - classification aléatoire : 1 chance sur 3 d'être de la classe 0
 - $\text{accuracy} = 1/3$, $\text{rappel}(\text{micro}) = 30/90 = 1/3$, $\text{F1}(\text{micro}) = 1/3$
 - n classes équilibrées

- classification multi-classes
 - exemple : cas de 3 classes équilibrées
 - classification aléatoire : 1 chance sur 3 d'être de la classe 0
 - $\text{accuracy} = 1/3$, $\text{rappel}(\text{micro}) = 30/90 = 1/3$, $F1(\text{micro}) = 1/3$
 - n classes équilibrées
 - $\text{accuracy} = 1/n$, $\text{rappel}(\text{micro}) = 1/n$, $F1(\text{micro}) = 1/n$

- classification multi-classes
 - exemple : cas de 3 classes équilibrées
 - classification aléatoire : 1 chance sur 3 d'être de la classe 0
 - $\text{accuracy} = 1/3$, $\text{rappel}(\text{micro}) = 30/90 = 1/3$, $F1(\text{micro}) = 1/3$
 - n classes équilibrées
 - $\text{accuracy} = 1/n$, $\text{rappel}(\text{micro}) = 1/n$, $F1(\text{micro}) = 1/n$
 - exemple de classification :

		classes estimées		
		classe 0	classe 1	classe 2
classes réelles	classe 0	70	10	20
	classe 1	10	90	0
	classe 2	5	15	80

Quelques ordres de grandeur : cas des classes déséquilibrées

- Vu que la classe 'George W Bush' est majoritaire, il peut être tentant de construire un classifieur qui réponde toujours 'George W Bush'.

Dans ce cas, quelles seraient les valeurs de chaque métrique ?

classe	numéro de classe	nombres d'images
Colin Powell	0	236
Donald Rumsfeld	1	121
George W Bush	2	530
Gerhard Schroeder	3	109
Tony Blair	4	144

Quelques ordres de grandeur : cas des classes déséquilibrées

- Vu que la classe 'George W Bush' est majoritaire, il peut être tentant de construire un classifieur qui réponde toujours 'George W Bush'. Dans ce cas, quelles seraient les valeurs de chaque métrique ?

classe	numéro de classe	nombres d'images
Colin Powell	0	236
Donald Rumsfeld	1	121
George W Bush	2	530
Gerhard Schroeder	3	109
Tony Blair	4	144

- précision (*accuracy*) = $530 / (236 + 121 + 530 + 109 + 144) = 0.4649$
- rappel (micro) : $\frac{VP}{VP+FN}$
 - $VP = 530$ et $FP = (236 + 121 + 109 + 144) = FN$
 - 'micro' : $\text{rappel} = 530 / (236 + 121 + 530 + 109 + 144) = F1 = 0.4649$
- None :
 - pour chaque classe sauf 'Bush' : $VP=0, FP=0, FN=\text{nb.images}$
 - pour la classe 'Bush' : $VP = 530, FP = (236 + 121 + 109 + 144), FN = 0$
donc $\text{recall}=1$ et $F1 = 2 * 0.4649 / (0.4649 + 1) = 0.6347$
- 'macro' : $0.6347 / 5 = 0.1269$
- 'weighted' : $0.6347 * 530 / (236 + 121 + 530 + 109 + 144) = 0.2951$

Quelques ordres de grandeur (2)

- Et que donnerait un classifieur qui tire au hasard les classes ?
 - En réalisant 1000 tirages aléatoires des prédictions de chaque données et en calculant la moyenne des performances, on obtient :
 - F1 micro : 0.2005
 - F1 None : [0.2037 0.1389 0.2800 0.1300 0.1550]
 - F1 macro : 0.1815
 - F1 weighted : 0.2191
 - Peut-on expliquer ce résultat ? (en supposant que chaque donnée a autant de chances d'être classée dans chacune des classes)

Quelques ordres de grandeur (2)

- Et que donnerait un classifieur qui tire au hasard les classes ?
 - En réalisant 1000 tirages aléatoires des prédictions de chaque données et en calculant la moyenne des performances, on obtient :
 - F1 micro : 0.2005
 - F1 None : [0.2037 0.1389 0.2800 0.1300 0.1550]
 - F1 macro : 0.1815
 - F1 weighted : 0.2191
 - Peut-on expliquer ce résultat ? (en supposant que chaque donnée a autant de chances d'être classée dans chacune des classes)
 - Probabilité qu'une donnée soit classée en classe i : $1/5 = 0.2$ avec $0 \leq i \leq 4$
 - Calcul pour F1 None, classe 0 'Colin Powell' : $P = 236$; $N = 904$
 - Parmi les positifs, 1/5e seront classés positifs : $VP = 236*0.2$, $FN = 236*0.8$
 - Parmi les négatifs, 1/5e seront classés positifs : $FP = 904*0.2$, $VN = 904*0.8$
 - $F1 = 2VP/(2VP+FN+FP) = 236*0.4/(236*0.4+236*0.8+904*0.2)=0.2035$
 - Il faudra faire mieux que $F1('micro')=0.4649$

Résultat possible du TP de classification de sons

