

IN104 - Classifieur de sons

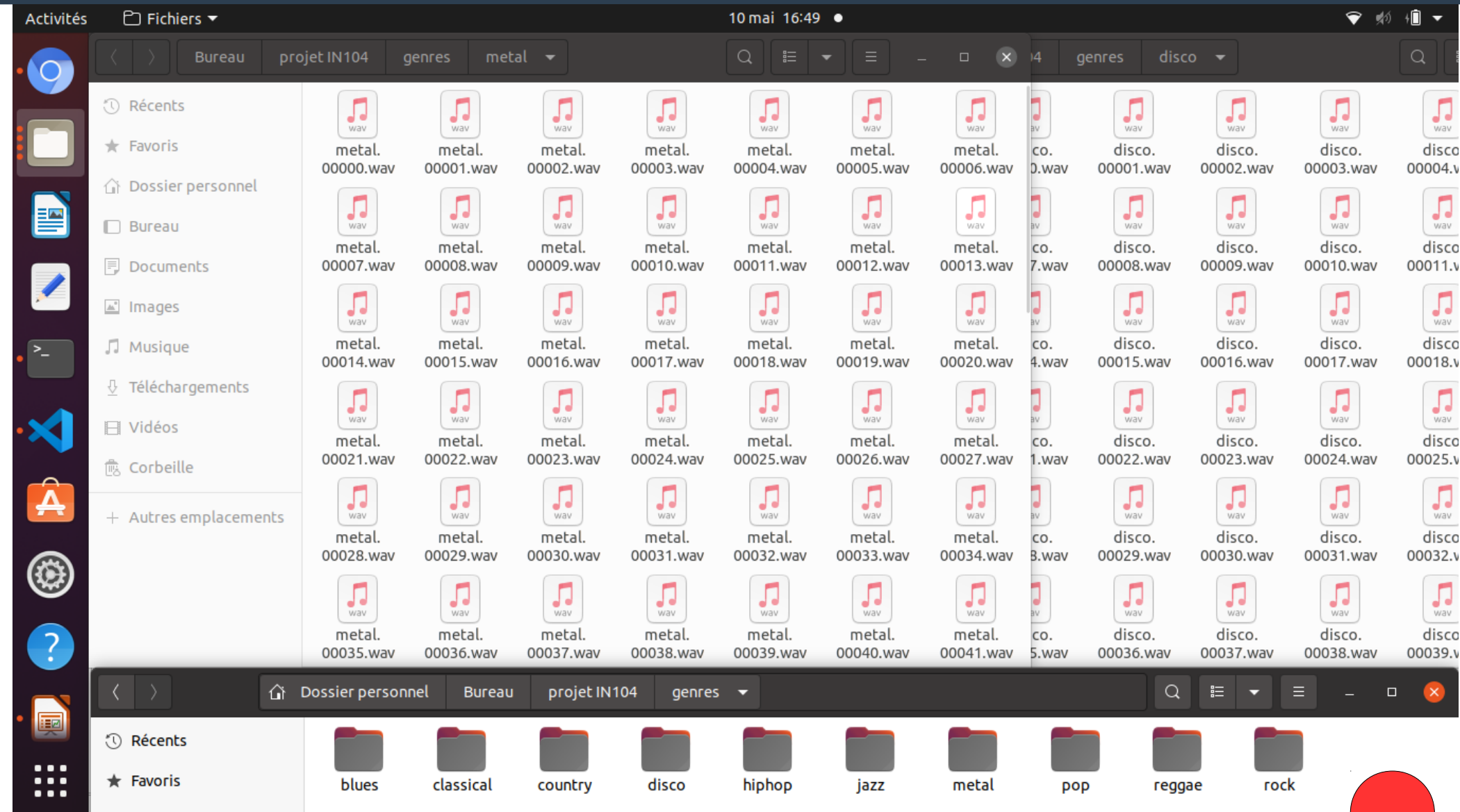


Bramas Timothé

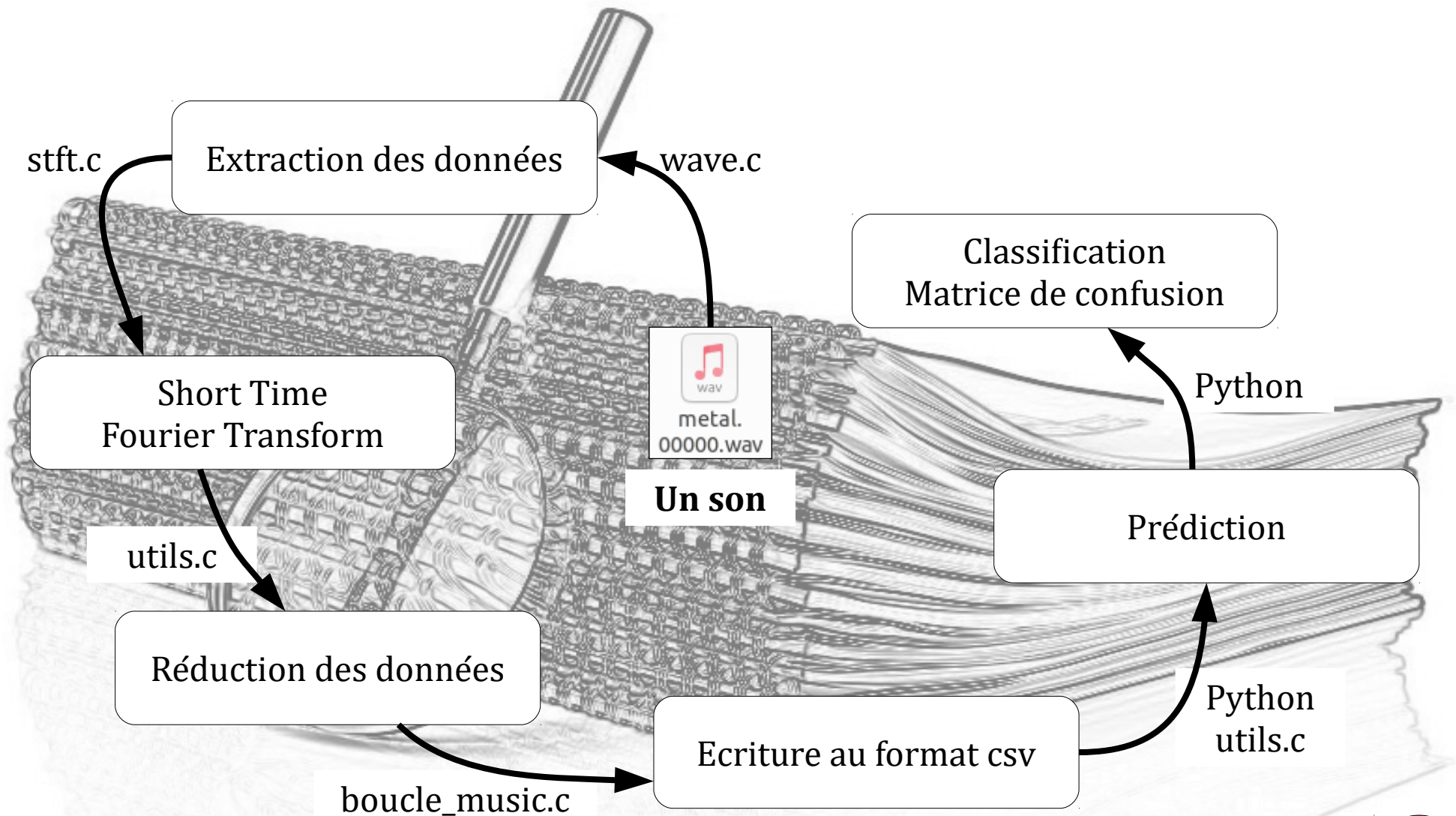
Boumaïza Sami



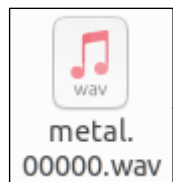
IN104 - Classifieur de sons



Sommaire



Short Time Fourier Transform



metal.
00000.wav

wave.c

Nombre d'échantillons (length)
Fréquence d'échantillonnage
Données d'amplitude

+

Hop size
Window size
Magnitude (tableau vide)
Taille de magnitude

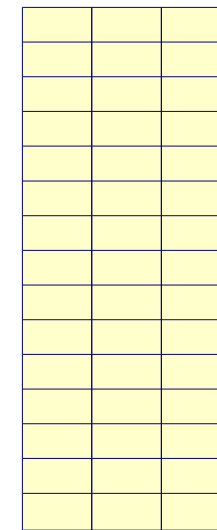
stft.c

Magnitude

HopSize = 512
WindowSize = 512
n_elements = 664345

$$\begin{aligned} n_elements &= \\ \text{length}/(\text{WindowSize}/2) * (\text{WindowSize}/2 + 1) \\ &= n\text{Row} * n\text{Col} \end{aligned}$$

freq



nRow=2585
nCol = 257

time



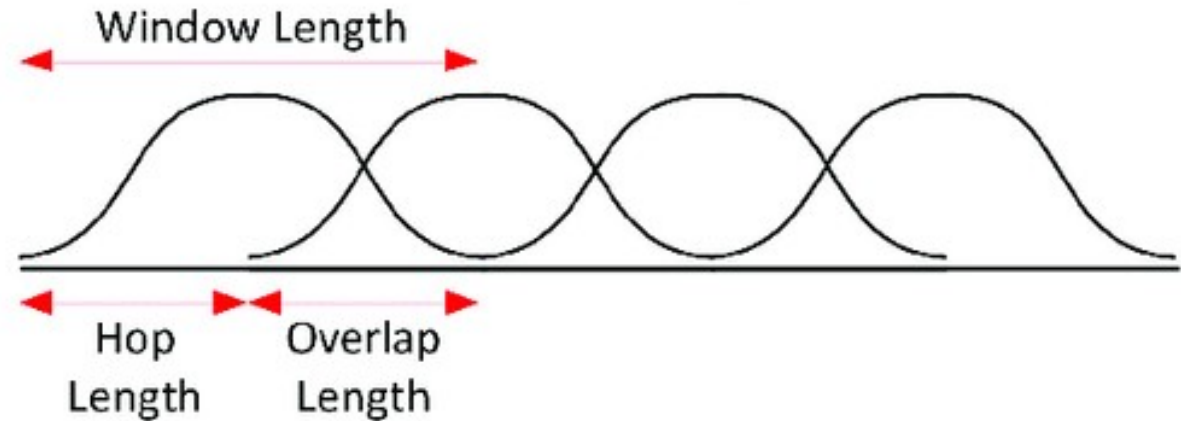
Short Time Fourier Transform

Nombre d'échantillons
 Fréquence d'échantillonnage
 Données d'amplitude
 +
 Hop size
 Window size
 Magnitude (tableau vide)
 Taille de magnitude

Signal



Windows

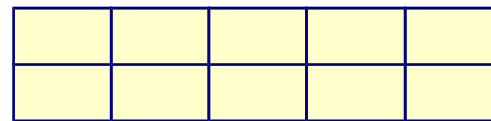


n_elements

=

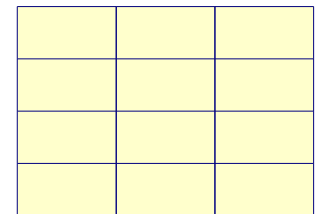
$\text{length}/(\text{WindowSize}/2) * (\text{WindowSize}/2 + 1)$

freq

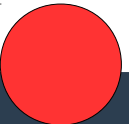


time

freq



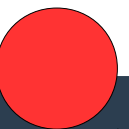
time



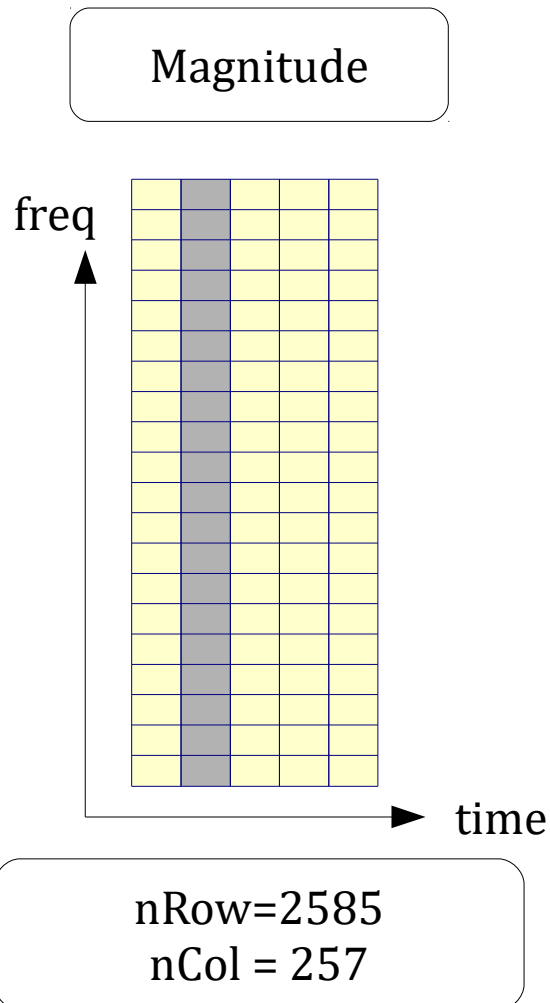
STFT - Résultat

Test unitaire pour blues.00000.wav, valeurs de magnitude :

96.167805
327.900165
93.793411
186.160593
180.747299
114.125950
140.482962
148.101793
116.486766
303.920547



Réduction des données



$\mu 1$	$\mu 2$	$\mu 3$	$\mu 4$...
$\sigma 1$	$\sigma 2$	$\sigma 3$	$\sigma 4$...

514 colonnes
2 lignes

Test unitaire
pour blues.00000.wav :

Moyenne 0 : 300.816232
Ecart-type 0 : 297.775569
Moyenne 1 : 735.819754
Ecart-type 1 : 595.734107
Moyenne 2 : 772.668620
Ecart-type 2 : 605.819411
Moyenne 3 : 872.611176
Ecart-type 3 : 659.907046
Moyenne 4 : 878.853137
Ecart-type 4 : 604.878590



Ecriture au format csv

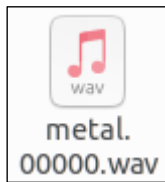
Projet

genres

blues

classical

...



IN104

Makefile

readme.md

src

fichier c et h

build

exec

Tableau
défini
constant

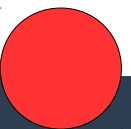
"../..genres/blues/blues.00000.wav"
"../..genres/classical/classical.00000.wav"
...

boucle_music.c

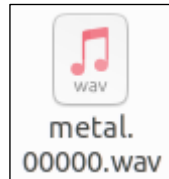
Boucle sur les genres pour
changer les deux derniers chiffres

Test unitaire : bonne chaîne de
caractères et ne renvoie pas le
pointeur nul

Directement incorporé dans
le main



Prédiction



80 % entraînement
20 % test

Classifieur Python :

- score (précision du modèle)
- Matrice W et vecteur b
- Matrice de confusion

Prédiction :

$\text{argmax de } W * @ + b$

$W : 10 * 514$

$b : 10 * 1$

@ = moy et écart-type

L'indice du maximum correspond au genre !

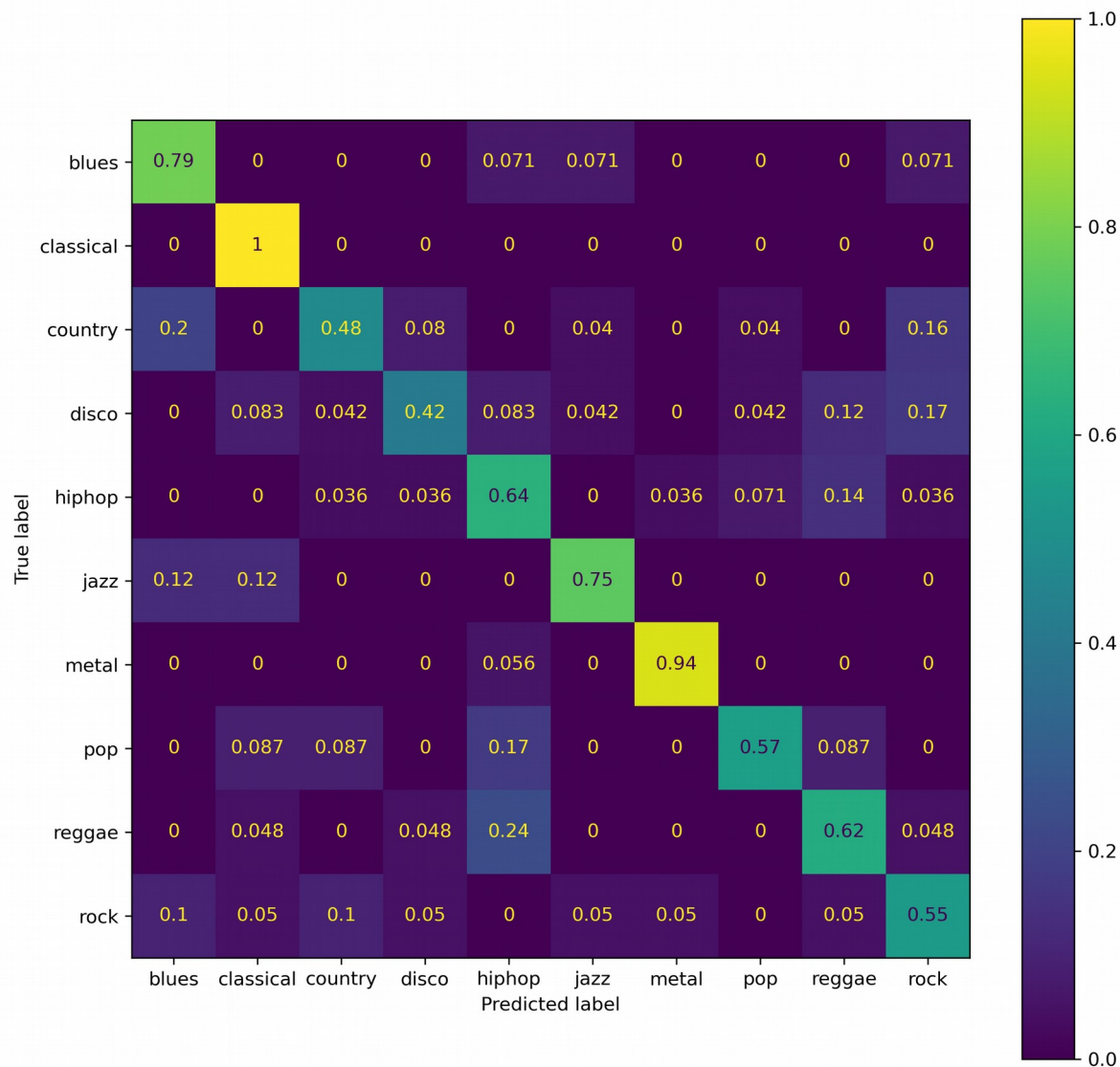
Score 0,49 – 0,56 sans normalisation

Score 0,64 – 0,66 avec normalisation Batch

Modifier les paramètres n'est pas significatif



Matrice de confusion



Score 0.66