



Projet ESIR2

FreeSound Audio Tagging

<https://www.kaggle.com/c/freesound-audio-tagging>

Plan

1. Objectifs et contraintes
2. Description des données
3. Extraction des “features”
4. Classifieurs
5. Résultats
6. Conclusion

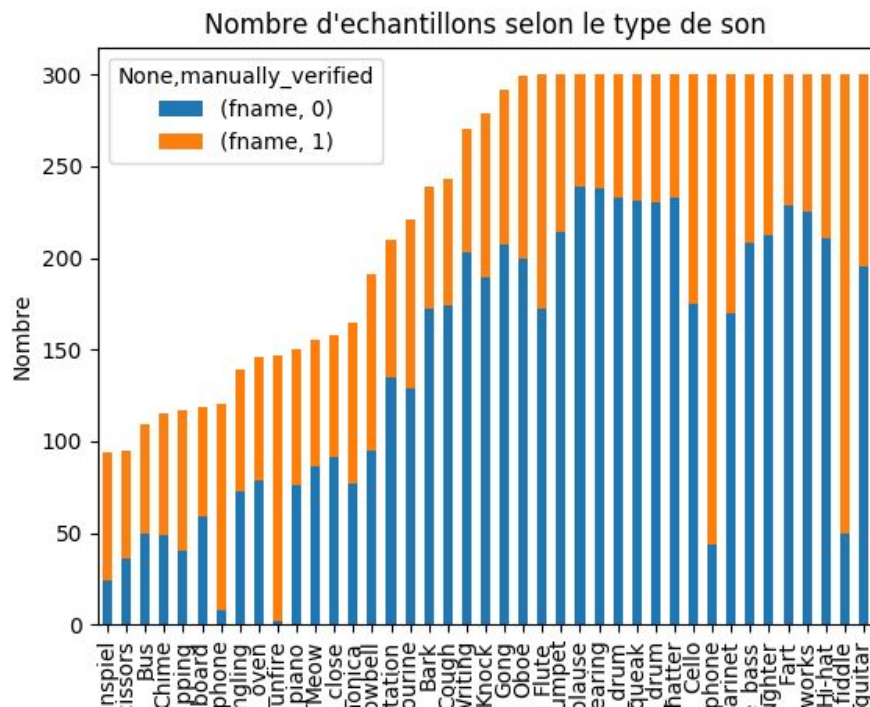


Objectifs

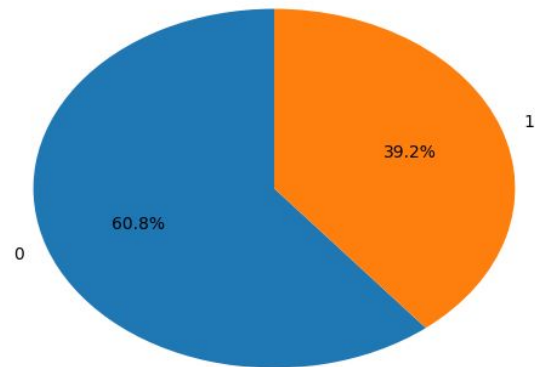
- Pas de système de reconnaissance audio.
- Coopération entre Freesound & Google Research's Machine Perception.
- Compétition Kaggle avec un dataset important.
- Dataset : train et test.



Description des données



Data manually verified 0:No 1:Yes



Number of data : 9473
Number of labels : 41
Missing Data Total
manually_verified 0
label 0
frame 0

fname,label,manually_verified
00044347.wav,Hi-hat,0
001ca53d.wav,Saxophone,1
002d256b.wav,Trumpet,0

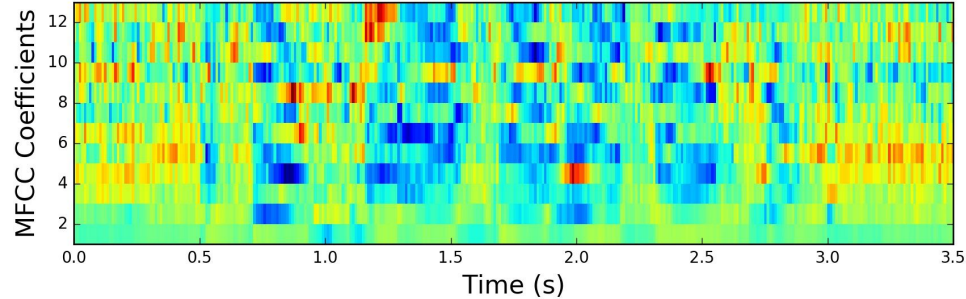
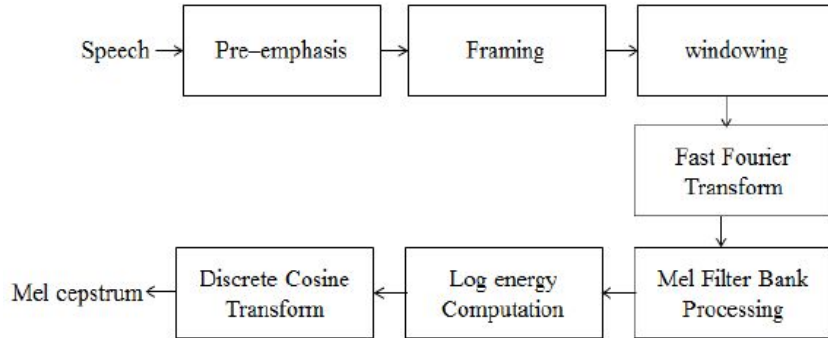




Extraction des features

Fichiers son en .wav 44100Hz inutilisable directement

Technique de reconnaissance vocale MFCC avec Librosa





Extraction des features

Liste des features utilisés :

- Zero crossing
- Spectral Contrast
- Spectral bandwidth
- Spectral centroid
- Spectral rolloff

Calcul de la moyenne, de la dérivée, min/max, skewness, kurtosis

Test de FFT plus généraliste





Classifieurs

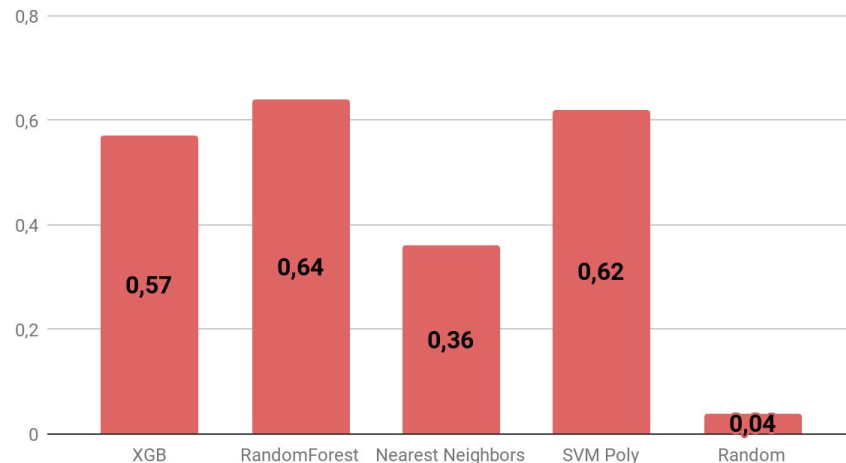
Classifieurs testés :

SVM, Gradient Boosting (XGB), Random forest, Nearest Neighbors et **essai** d'un réseau de neurones.

Classifieur retenu dans la soumission finale :

→ Random Forest

Taux de prédiction du label





Classifieurs

Trouver les meilleurs hyper-paramètres :

Utilisation de random search Cross-validation sur Random forest.

→ Légère amélioration des prédictions : score de **0,66**





Résultats

Rendu d'un fichier .csv de soumission à Kaggle avec 3 prédictions

Prédiction finale de **77,2%**

Affichage des erreurs



Freesound General-Purpose Audio Tagging Challenge | Kaggle - Mozilla Firefox

File Edit View History Bookmarks Tools Help

acidBaz sponsoring - Rec X bojou Victor - Recherche X FreeSound Audio Tagging X nicolas rochat ça va ? - R X Freesound General-Purpose X +

https://www.kaggle.com/c/freesound-audio-tagging/leaderboard

New Tab Bureau Development Music FreeSound

Overview Data Kernels Discussion Leaderboard Rules Team My Submissions Submit Predictions

148	▼ 2	ValentinLion		0.787	9	1d
149	▼ 16	NTU_b05902013_到底要選哪一...		0.785	6	10d
150	new	Dmitry Lunin		0.780	1	2d
151	▼ 17	Michael Kazachok	↔ CatBoost&MFCC	0.777	8	2mo
152	▼ 17	Nick Styles		0.777	1	1mo
153	▼ 17	Seiya Nakata		0.775	1	1mo
154	▼ 17	wajid zaman		0.773	1	2mo
155	new	Adrien Riou		0.772	1	2h
Your Best Entry ↑						
Your submission scored 0.772				Tweet this!		
156	▼ 18	Luc Forget		0.761	2	3d
157	new	Shafialsmail		0.760	4	1d
158	▼ 19	zhenkanglau		0.759	1	17d
159	▼ 19	finejuly		0.756	3	17d
160	new	arain		0.754	4	2d
161	▼ 20	sangwook		0.754	2	1mo
162	▼ 19	Valentyn Vaityshyn		0.753	4	20d
163	▼ 19	Mojtaba		0.753	8	1mo
164	▼ 19	SarthakYadav		0.747	2	22d
165	▼ 18	NTU_b04901009_正妹隊長		0.741	5	12d
166	new	lvi38		0.740	8	19h

Anna from Kaggle

Hi @adrienriou, Our revised Privacy Policy and Terms of Use are now in effect. We're making



Démonstration



Conclusion

- Score correct
- Extractions des features très longues
- Sons parfois difficiles à classifier par un humain

Axes de continuation :

- Mise en place d'un réseau de neurones
- Utilisation plus poussée des classifieurs
- Amélioration des features
- Utilisation du manually_verified

