



ÉCOLE NATIONALE SUPÉRIEURE D'INFORMATIQUE ET
D'ANALYSE DES SYSTÈMES

RAPPORT DU PROJET IA

**Outil de classification et detection du
genre musicale**

Réalisé par :

EL GUECHATI Hamza
EL HABTI Soufiane

Encadré par :

MME.BENBRAHIM HOUDA

Année universitaire 2020 - 2021

REMERCIEMENT

J'exprime ma profonde gratitude et respectueuses reconnaissances à Mme Houda Benbrahim pour sa bonne volonté d'accepter de m'encadrer, pour tout le temps qu'il m'a accordé et pour tous les conseils qu'il m'a prodigué durant la seance.

RÉSUMÉ

L'utilisation des techniques d'intelligence artificielle est aujourd'hui un élément clé pour les entreprises peu importe leur taille et le système d'information est devenu un élément central de l'activité des différents services et doit fonctionner pleinement et en permanence pour garantir l'efficacité de l'entreprise. Son rôle intervient à tous les niveaux : les réseaux, les terminaux utilisateurs, les serveurs d'applications ainsi que les données, avec comme objectif, la garantie du bon fonctionnement de l'entreprise. Pour pouvoir garantir une activité ainsi qu'une bonne notoriété de son entreprise, il est primordial de réduire au maximum les problèmes informatiques. C'est pour cela que les entreprises ont désormais recours à des outils intelligentes.

Le but est de mettre en place cette application c'est pour différencier entre les genres musicales en utilisant des techniques informatiques et des packages .

LISTE DES FIGURES

Figure 2.1 : 8 classes.....	13
Figure2.2 : neural network.....	13
Figure2.3 : matrice de confusion.....	14
Figure2.4 : Interface graphique.....	28
Figure2.5 : Representation graphique des genres dans un morceau.....	29

TABLE DES MATIÈRES

Introduction générale	7
1 Présentation et Cadre Général Du Projet	8
1 Introduction	9
2 Méthode d'apprentissage automatique	9
3 Les différentes approches	9
2 Travail Effectué	11
1 Introduction	12
2 extraction des attributs	12
2.1 Spectrogramme	12
2.1.1 MFCC - Coefficients Cepstraux Mel-Frequency	12
2.1.2 Centroïde spectral	13
3 Equilibrer la dataset	13
3.1 La forme de la dataset	13
3.2 Equilibrer la dataset	14
3.3 Preparation de données	15
4 MODEL DE CLASSIFICATION	15

4.1	Réseau de neurones artificiels	15
4.2	L'entraînement du model	15
4.3	Matrice de confusion	15
5	conclusion	16
6	Interface graphique et test	18
6.1	Interface d'utilisateur	18
6.2	Prediction	18
7	Conclusion	19

INTRODUCTION GÉNÉRALE

Actuellement, l'utilisation des systèmes d'Apprentissage automatique au sein des entreprises deviennent de plus en plus importants mais aussi complexes. Le besoin de maintenance et de gestion de ces systèmes est rapidement devenu une priorité. Grâce à cela l'analyse de données, et peut s'appliquer à des domaines tels que le e-commerce, la relation client, la sécurité ou encore la santé..

Ce rapport présente l'ensemble des étapes suivies pour étudier et mettre en place l'application. Il contient quelques chapitres qui décrivent la démarche suivie pour atteindre cet objectif .

CHAPITRE

1

PRÉSENTATION ET CADRE GÉNÉRAL DU PROJET

1 Introduction

Nous allons introduire dans ce premier chapitre l'étude faite concernant notre projet.

2 Méthode d'apprentissage automatique

Dans l'apprentissage automatique, les tâches sont généralement classées en grandes catégories. Ces catégories sont basées sur la façon dont l'apprentissage est reçu ou comment le feedback sur l'apprentissage est donné au système développé.

Deux des méthodes d'apprentissage automatique les plus largement adoptées sont l'apprentissage supervisé qui forme des algorithmes basés sur des données d'entrée et de sortie étiquetées par l'homme et l'apprentissage non supervisé qui ne fournit pas à l'algorithme des données étiquetées pour lui permettre de trouver une structure et de découvrir une logique dans données entrées. Explorons donc ces méthodes plus en détail.

L'apprentissage supervisée Dans l'apprentissage supervisé, l'ordinateur est fourni avec des exemples d'entrées qui sont étiquetés avec les sorties souhaitées. Le but de cette méthode est que l'algorithme puisse «apprendre» en comparant sa sortie réelle avec les sorties «enseignées» pour trouver des erreurs et modifier le modèle en conséquence. L'apprentissage supervisé utilise donc des modèles pour prédire les valeurs d'étiquettes sur des données non étiquetées supplémentaires.

L'apprentissage non supervisé Dans l'apprentissage non supervisé, les données sont non étiquetées, de sorte que l'algorithme d'apprentissage trouve tout seul des points communs parmi ses données d'entrée. Les données non étiquetées étant plus abondantes que les données étiquetées, les méthodes d'apprentissage automatique qui facilitent l'apprentissage non supervisé sont particulièrement utiles.

3 Les différentes approches

En tant que domaine informatique, l'apprentissage automatique est étroitement lié aux statistiques mathématiques; disposer d'une connaissance approfondie des statistiques est donc utile pour comprendre et exploiter les algorithmes d'apprentissage automatique.

Pour ceux qui n'ont peut-être pas étudié les statistiques, il peut être utile de définir d'abord la corrélation et la régression, car ce sont des techniques couramment utilisées pour étudier la relation entre les variables quantitatives. La corrélation est une mesure d'association entre deux variables qui ne sont pas désignées comme dépendantes ou indépendantes. La régression à un niveau de base

est utilisée pour examiner la relation entre une variable dépendante et une variable indépendante. Comme les statistiques de régression peuvent être utilisées pour anticiper la variable dépendante lorsque la variable indépendante est connue, la régression active les capacités de prédiction.

Les approches de l'apprentissage automatique sont continuellement développées. Pour nos besoins, nous allons passer en revue quelques-unes des approches populaires utilisées dans l'apprentissage automatique au moment de la rédaction de cet article.

Et pour notre projet on a choisi

L'apprentissage en profondeur - "Deep Learning" : Des tentatives profondes d'apprentissage pour imiter comment le cerveau humain peut traiter des stimuli lumineux et sonores dans la vision et l'ouïe sont à l'étude. Une architecture d'apprentissage en profondeur est inspirée par les réseaux neuronaux biologiques et se compose de plusieurs couches dans un réseau neuronal artificiel composé de matériel et de GPU.

L'apprentissage en profondeur utilise une cascade de couches d'unités de traitement non linéaires afin d'extraire ou de transformer les caractéristiques (ou représentations) des données. La sortie d'une couche sert d'entrée de la couche suivante. Dans l'apprentissage en profondeur, les algorithmes peuvent être supervisés et servir à classer les données, ou non supervisés et à effectuer une analyse de modèle. Parmi les algorithmes d'apprentissage machine actuellement utilisés et développés, l'apprentissage en profondeur absorbe le plus de données et a été capable de battre les humains dans certaines tâches cognitives. En raison de ces attributs, l'apprentissage en profondeur est devenu l'approche avec un potentiel significatif dans le monde de l'intelligence artificielle. La reconnaissance faciale par ordinateur et la reconnaissance vocale ont toutes deux permis de réaliser des progrès significatifs grâce à des approches d'apprentissage approfondies. IBM Watson est un exemple bien connu d'un système qui exploite l'apprentissage en profondeur.

CHAPITRE

2

TRAVAIL EFFECTUÉ

1 Introduction

Dans ce chapitre, nous allons présenter les étapes principal pour effectuer notre travail

2 extraction des attributs

en se utilisant la bibliothèque Librosa on peut visualiser les attribut qu'on peut extraire d'un morceau de musique

2.1 Spectrogramme

Un spectrogramme est une représentation visuelle du spectre des fréquences du son ou d'autres signaux qui varient avec le temps. C'est une représentation des fréquences changeant en fonction du temps pour des signaux musicaux donnés.

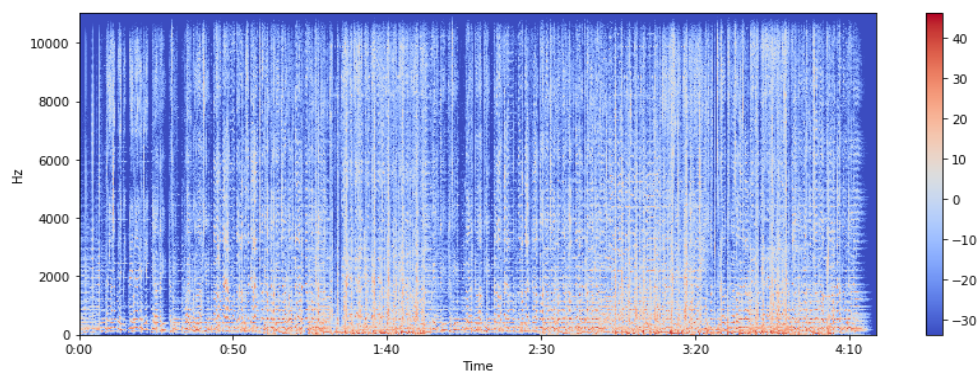


FIGURE 2.1: Spectrogramme

2.1.1 MFCC - Coefficients Cepstraux Mel-Frequency

Cette fonction est l'une des méthodes les plus importantes pour extraire une fonction d'un signal audio et est utilisée principalement lorsque vous travaillez sur des signaux audio. Les coefficients cepstraux de fréquence de mélange (MFCC) d'un signal sont un petit ensemble de caractéristiques (généralement environ 10-20) qui décrivent de manière concise la forme globale d'une enveloppe spectrale.

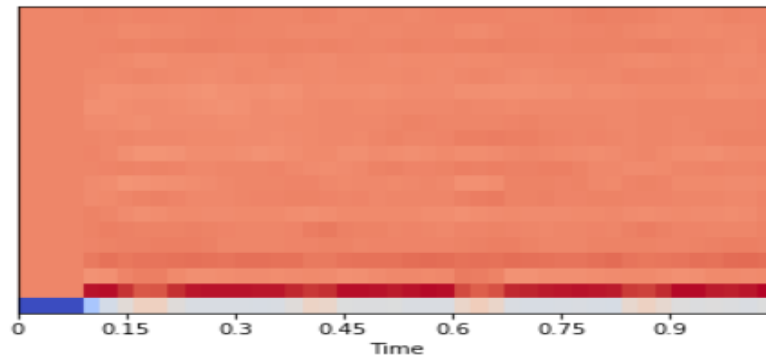


FIGURE 2.2: MFCC

2.1.2 Centroïde spectral

Il indique où se situe le «centre de masse» d'un son et est calculé comme la moyenne pondérée des fréquences présentes dans le son. Si les fréquences en musique sont les mêmes partout, le centroïde spectral serait autour d'un centre et s'il y a des fréquences élevées à la fin du son, le centroïde serait vers sa fin.

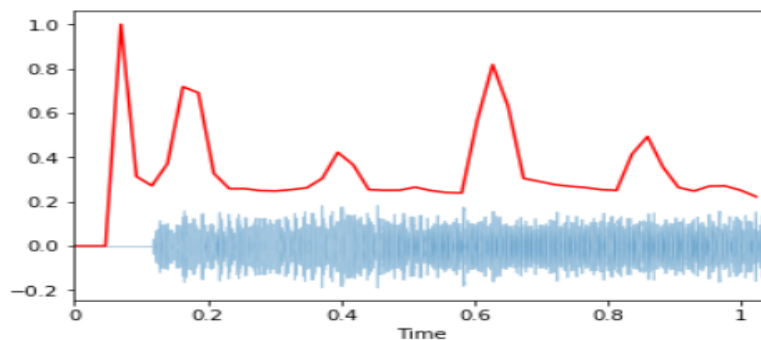


FIGURE 2.3: centroïde spectral

3 Equilibrer la dataset

3.1 La forme de la dataset

Nous avons maintenant 104343 observations et 16 genres 52,47 % des observations ont plus d'un genre En moyenne, une chanson a 1,70 genres L'équilibre du genre est [1] :

Blues — 1752

Classical — 4106

Country — 1987

Easy Listening — 730
 Electronic — 34413
 Experimental — 38154
 Folk — 12706
 Hip-Hop — 8389
 Instrumental — 14938
 International — 5271
 Jazz — 4126
 Old-Time / Historic — 868
 Pop — 13845
 Rock — 32923
 Soul-RnB — 1499
 Spoken — 1876

3.2 Equilibrer la dataset

Ce que nous avons vu précédemment, c'est que l'ensemble de données est très déséquilibré. Malheureusement, il n'existe aucun moyen d'équilibrer un BoW pour le moment. Nous allons donc utiliser une astuce pour le faire. L'idée est de dupliquer des fonctionnalités audio ayant plusieurs genres, puis de ne conserver qu'un seul genre pour chacun. Ensuite, nous pouvons équilibrer l'ensemble de données en utilisant une solution classique de suréchantillonnage .

Et puisque notre dataset contient 16 genres nous avons pris juste 8 genres pour les predire

track_id	Electronic	Experimental	Folk	Hip-Hop	Instrumental	International	Pop	Rock
0	2	0	0	0	1	0	0	0
1	3	0	0	0	1	0	0	0
2	5	0	0	0	1	0	0	0
3	10	0	0	0	0	0	0	1
4	20	0	0	1	0	0	0	1

FIGURE 2.4: 8 classes

3.3 Preparation de données

Nous avons correctement le même nombre de classes et le même nombre d'observations. Avec 610464 observations, nous avons appliqué l'ACP sur la totalité de 518 composants.

pour préparer on a utilisé : `StandardScaler()` [4] qui normalisera les entités afin que chaque colonne/entité/variable ait $\text{mean} = 0$ et $\text{standard deviation} = 1$.

4 MODEL DE CLASSIFICATION

Le concept de supervision a été normalisé par l'ISO (International Organization for Standardization). Voici les différentes fonctions qui ont été définies par l'ISO :

4.1 Réseau de neurones artificiels

nous avons utilisé un modèle de l'architecture générée par la bibliothèque Keras [3] suivante : Avec 3 couches cachées de 256, 128, 64 et un output de 8 variables qui vont représenter les genres

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 256)	89344
dense_2 (Dense)	(None, 128)	32896
dropout_1 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 64)	8256
dropout_2 (Dropout)	(None, 64)	0
dense_4 (Dense)	(None, 8)	520

FIGURE 2.5: neural network

4.2 L'entraînement du modèle

pour entraîner notre modèle on a utilisé un nombre d'époques égal à 40 pour avoir la meilleure précision et éviter l'overfitting, alors on a atteint 0.57 précision de validation après on a augmenté le nombre d'époques à 200

4.3 Matrice de confusion

En apprentissage automatique supervisé, la matrice de confusion est une matrice qui mesure la qualité d'un système de classification. Chaque ligne correspond à une classe réelle, chaque colonne

correspond à une classe estimée. La cellule ligne L, colonne C contient le nombre d'éléments de la classe réelle L qui ont été estimés comme appartenant à la classe C1.

Un des intérêts de la matrice de confusion est qu'elle montre rapidement si un système de classification parvient à classer correctement. Notre matrice de confusion le model on a constater deux resultat Pour 40 epochs :

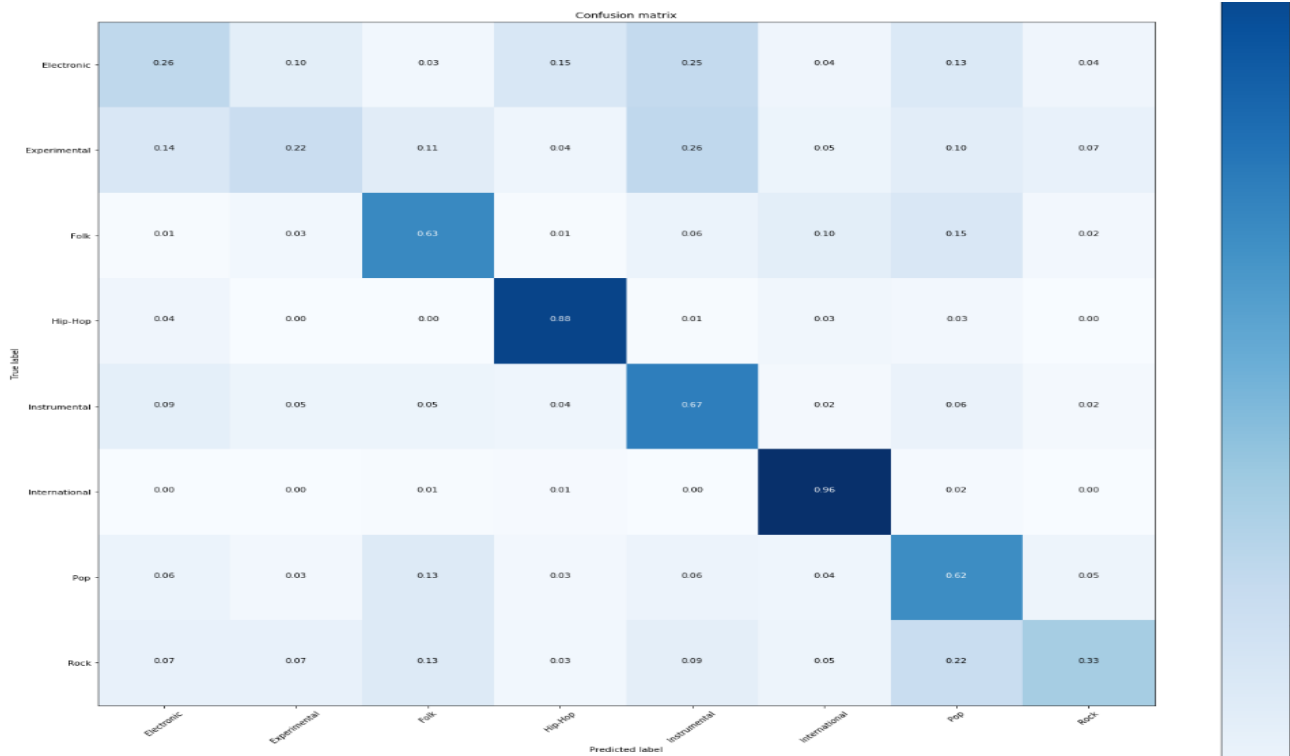


FIGURE 2.6: matrice de confusion

Pour 200 epochs :

5 conclusion

on peut constater que l'augmentation de nombre d'epochs n'a pas fait un grand changement .

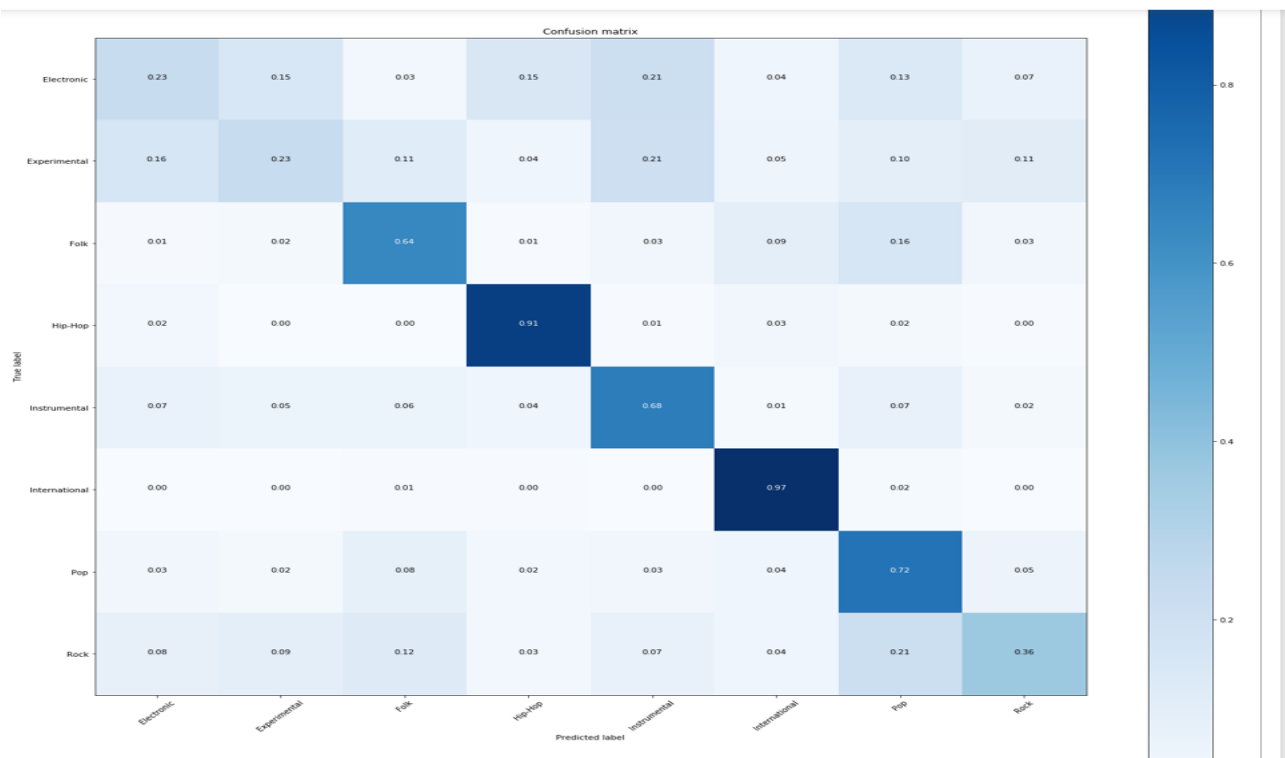


FIGURE 2.7: matrice de confusion

6 Interface graphique et test

6.1 Interface d'utilisateur

Nous avons mis en place une interface très simple et facile à utiliser à l'aide du tkinter [5] package fournis par le langage de programmation Python.

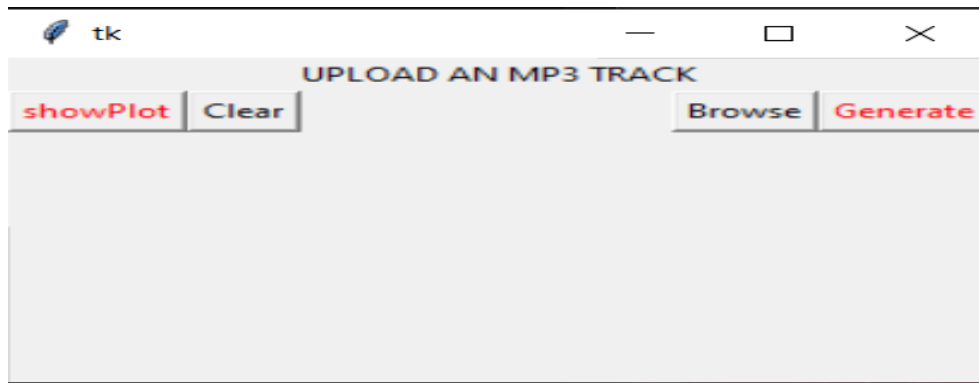


FIGURE 2.8: Interface graphique

6.2 Prediction

apres qu'on choisit un morceau de musique on clique sur generate pour predire et sur showPlot pour affihcer la prediction sous frome de de cette representation graphique :

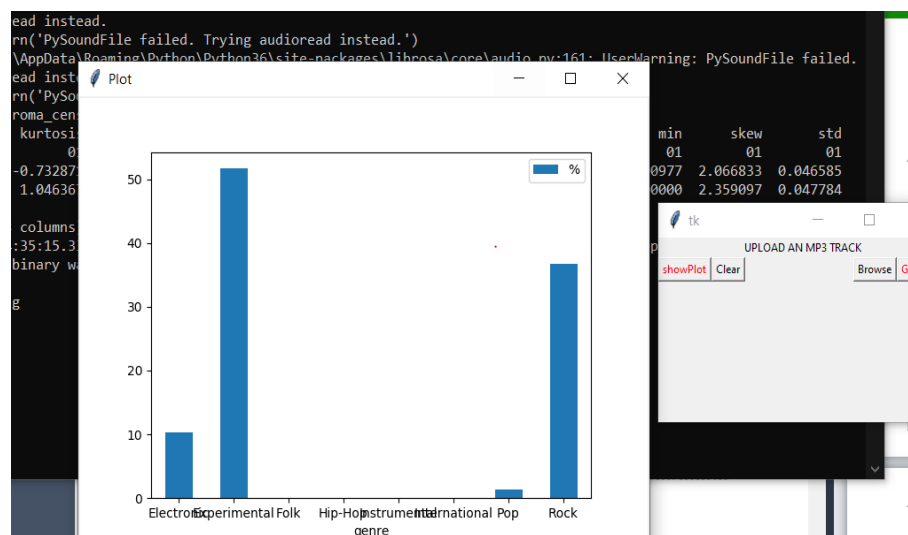


FIGURE 2.9: Representation graphique des genres dans un morceau

7 Conclusion

par une accuracy de 0.57 ce model est un peu bon pour classifier, cependant on n'as utilisé d'autre model pour comparer les resultats .

CONCLUSION GÉNÉRALE

machine learning est un outil très puissant qui permet d'effectuer de multiples actions comme classifier des données, faire apprendre à un programme à partir d'expérimentations ou encore de créer un programme évolutionnaire qui s'améliore sans cesse. Ainsi, même avec un échantillon peu fourni (le machine learning nécessite habituellement des échantillons avec 50 spécimens) et des données influencées par la subjectivité de celui qui les mesure, le machine learning reste relativement précis malgré quelques lacunes.

Néanmoins, le machine learning n'a pas que des qualités, il doit être constamment adapté au problème qu'il tente de résoudre. En effet, le programmeur doit tout d'abord se procurer l'échantillon le plus représentatif possible. Ensuite il faudra qu'il choisisse la fonction la plus fidèle à l'échantillon, ce qui n'est pas nécessaire dans notre cas car les classes sont suffisamment distinctes pour que changer la forme de la fonction n'influe pas les résultats obtenus. Enfin, le machine learning doit être utilisé comme un outil car tous les problèmes ne nécessitent pas un programme complexe en machine learning. Dans notre cas, une simple clé de détermination basique couplée à l'expérience d'un botaniste sont bien plus efficaces. Mais après se pose la question où faire la limite entre les tâches à apprendre à un programme et celles à apprendre soi-même ?

WEBLIOGRAPHIE

Dataset et bibliothèques

/[1] <https://librosa.github.io/librosa/>

/[2] <https://arxiv.org/pdf/1612.01840.pdf> , FMA : A DATASET FOR MUSIC ANALYSIS

/[3] <https://keras.io/>

/[4] <https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html>

/[5] <https://wiki.python.org/moin/TkInter>

/[6] http://www.astro.caltech.edu/~george/aybi199/Donalek_classification.pdf