ASSIGNMENT
Text-based audio classification

PART A - Prepare and explore your dataset

To run the cells in this notebook you'll need to install the following Python dependencies:

- `numpy`
- `sklearn`

If you have not installed them you should be able to do so by running: `pip install numpy sklearn`.

Furthermore, you'll need to get **Freesound API** credentials so that you can retreive sounds' metadata for building your dataset. Please create a Freesound account (if you don't already have one) and go here: http://www.freesound.org/apiv2/apply/ (http://www.freesound.org/apiv2/apply/). Set `API_KEY` variable below to the given API key.

In [2]:

```
from __future__ import print_function
import freesound
import random
import utils
import json
from IPython.core.display import display, HTML
from collections import defaultdict

API_KEY='uo6MmiMPAJiWkAxjP0aMOmJoeS7H0IN7pshv2p7a'
c = freesound.FreesoundClient()
c.set_token(API_KEY,"token")
PAGE_SIZE = 50 # Page size for fs requests, no need to change that
```

1) Select your audio categories and get metadata from Freesound for each category

For the task in this session we provide a pre-selected number of audio categories with a number of Freesound sound IDs linked to them. The first step is to choose which classes to include in the dataset that we'll use later for classification and to get metadata for each sound in our dataset.

- Chose the categories you want to work with by editing the `DATASET_CLASSES` list below.
- You can limit the number of sounds chosen per class by setting the `N` parameter.
- The resulting dataset will be saved to a file named `DATASET_NAME.json`. This file will be loaded later in another notebook to carry out the classification task.

In [4]:

```
# Configure dataset parameters and audio categories

DATASET_NAME = 'instruments' # Dataset will be saved in a .json file with this name
N = None # Number of sounds per class (set 'None' to get all in dataset)
DATASET_CLASSES = [
    'Percussion',
    'Wind instrument, woodwind instrument',
    'Bowed string instrument',
    'Domestic sounds, home sounds',
 ]

# Get sound examples from Freesound
data_index =  json.load(open('data_index.json'))
dataset = defaultdict(list)

for klass in DATASET_CLASSES:
    sound_ids = data_index.get(klass, None)
    if sound_ids is None:
        print('Skipping class %s as no data is available for it' % klass)
        continue
    random.shuffle(sound_ids)  # Shuffle order of sound ids
    sound_ids = sound_ids[:N]  # Limit number of selected sound ids

    print('Getting sounds\' data for class %s...' % klass)
    for i in range(0, len(sound_ids), PAGE_SIZE):
        current_sound_ids = sound_ids[i:i + PAGE_SIZE]

        fields = "id,tags,description,username,analysis"
        # For the dataset classes :
        #'Percussion','Wind instrument, woodwind instrument','Bowed string instrument','Domestic sounds, hom
e sounds'
        # the audio features : pitch_salience, spectral_flux, hfc, average_loudness and dissonance will be h
elpful.
        descriptors = "lowlevel.pitch_salience,lowlevel.dynamic_complexity,lowlevel.spectral_flux,lowlevel.h
fc,lowlevel.average_loudness,lowlevel.dissonance"
        results_pager = c.text_search(
            filter='id:(%s)' % ' OR '.join([str(sid) for sid in current_sound_ids]),
            page_size=PAGE_SIZE,
            fields = fields,
            descriptors=descriptors
        )
        dataset[klass] += results_pager.results

        # TIP ON AUDIO FEATURES: you can get also audio features extracted in freesound by passing a 'descri
ptors'
        # parameter in the text_search function and including 'analysis' in the fields list
        # (see http://www.freesound.org/docs/api/resources_apiv2.html#response-sound-list):
        #
        # fields = "id,tags,description,username,analysis"
        # descriptors = "lowlevel.spectral_centroid,lowlevel.barkbands.mean"
        #
        # e.g.: results_page = c.text_search(query=target_query, ..., fields=fields, descriptors=descriptors
)
        # ...

# Show information and save dataset to file so we can work with it later on
utils.save_to_json('%s.json' % DATASET_NAME, dataset)
print('\nDataset created with %i classes:' % len(dataset))
for klass, sounds in dataset.items():
    print('\t%s: %i sounds' % (klass, len(sounds)))
print('Saved to %s' % '%s.json' % DATASET_NAME)
```

```
Getting sounds' data for class Percussion...
Getting sounds' data for class Wind instrument, woodwind instrument...
Getting sounds' data for class Bowed string instrument...
Getting sounds' data for class Domestic sounds, home sounds...

Dataset created with 4 classes:
        Wind instrument, woodwind instrument: 200 sounds
        Bowed string instrument: 200 sounds
        Percussion: 200 sounds
        Domestic sounds, home sounds: 200 sounds
Saved to instruments.json
```

2) Explore the dataset (know your data!)

The cell below displays some data about your recently created dataset so you can have an idea of its contents. Feel free to experiment with the data and add any relevant plots that you might find useful for the future classificaiton task.

In [7]:

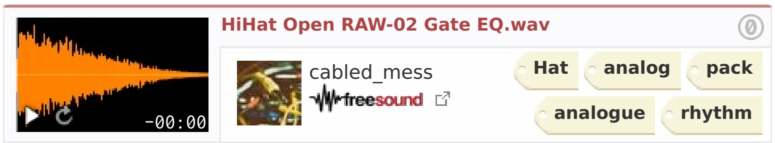
```
# Pick some sounds from each category and show players (Freesound embeds) to listen to them
for count, (class_name, sounds) in enumerate(dataset.items()):
    html = "<h3 style='color:#bbb;'>%i) %s</h3>" % (count + 1, class_name)
    html += "<h4>Example sounds:</h4><br>"
    html += utils.generate_html_with_sound_examples([sound['id'] for sound in sounds][:6])
    html += "<h4>Most commons tags tagcloud:</h4><br>"
    class_tags = utils.get_all_tags_from_class(class_name, dataset)
    html += utils.generate_html_tagcloud(class_tags, N=100, max_px=30, min_px=10, pow_scale=1.2)
    html += "<br><br>"
    display(HTML(html)) # <- This is pure jupyter notebook AWESOMENESS which renders the HTML in the output of the cell
```

1) Percussion

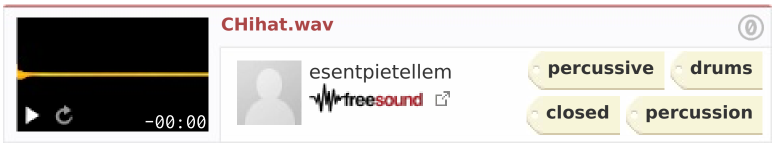
Example sounds:

Glockenspiel_35_c4_04
-00:03
cabled_mess
freesound
single-note
Glockenspiel

TOY_GLOCK-23.wav
-00:08
kilo one nine
freesound
incremental
glockenspiel

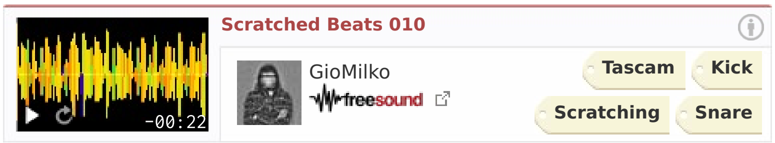


HiHat Open RAW-02 Gate EQ.wav
-00:00
cabled_mess
freesound
Hat analog pack
analogue rhythm



CHihat.wav
-00:00
esentpietellem
freesound
percussive drums
closed percussion

bicycle-bell_12.wav
-00:01
newagesoup
freesound
glockenspiel
hit ring metallic



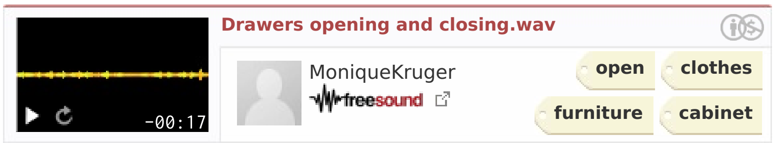
Scratched Beats 010
-00:22
GioMilko
freesound
Tascam Kick
Scratching Snare

Most commons tags tagcloud:

kick processed machine Glockenspiel reverb kick-drum 909 beat realistic toy single single-note drum-machine electronic
tone hi-hat distortion techno hardstyle china plastic wav hardcore tamborine vsco-2 pack xylophone snare gong sample
CompMusic compmusic chinese-traditional qmul beijing-opera Tabla mallets peking-opera heavy metal glock musical-instrument
tones mallet Tabla-stroke bassdrum icassp2014-dataset idiophone tambourine echo dance hard rhythm bd tribal dirty
sampler xiaoluo-instrument Percussion 1-shot edited recording tony acoustic 808 analog from cymbal hihat drum
kit house timpani percussion-loop samples percussive note loop bass vibraphone one-shot real metallophone instrument
multihit velocity drums tabla notes hit drum-loop incremental kickdrum multisample chinese
percussion glockenspiel roll percs Indian-percussion

2) Domestic sounds, home sounds

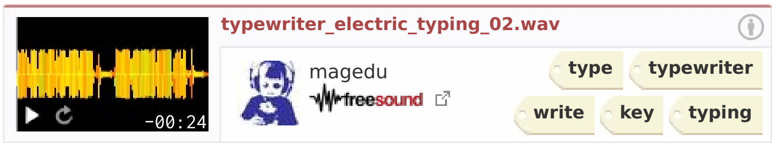
Example sounds:


Drawers opening and closing.wav

-00:17

MoniqueKruger

open clothes furniture cabinet


typewriter_electric_typing_02.wav

-00:24

magedu

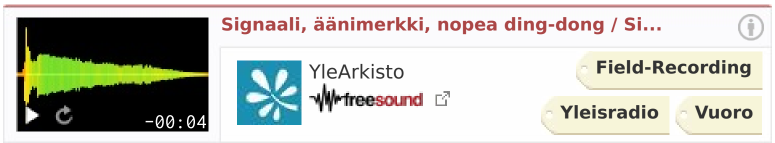
type typewriter write key typing

Kitchen Drawer Close.wav

-00:02

Glitchedtones

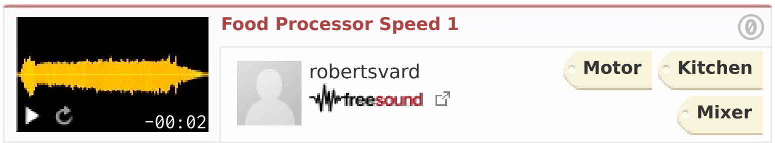
sliding kitchen close closing


Signaali, äänimerkki, nopea ding-dong / Si...

-00:04

YleArkisto

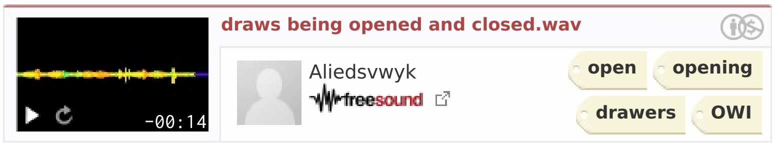
Field-Recording Yleisradio Vuoro


Food Processor Speed 1

-00:02

robertsvard

Motor Kitchen Mixer


draws being opened and closed.wav

-00:14

Aliedsvwyk

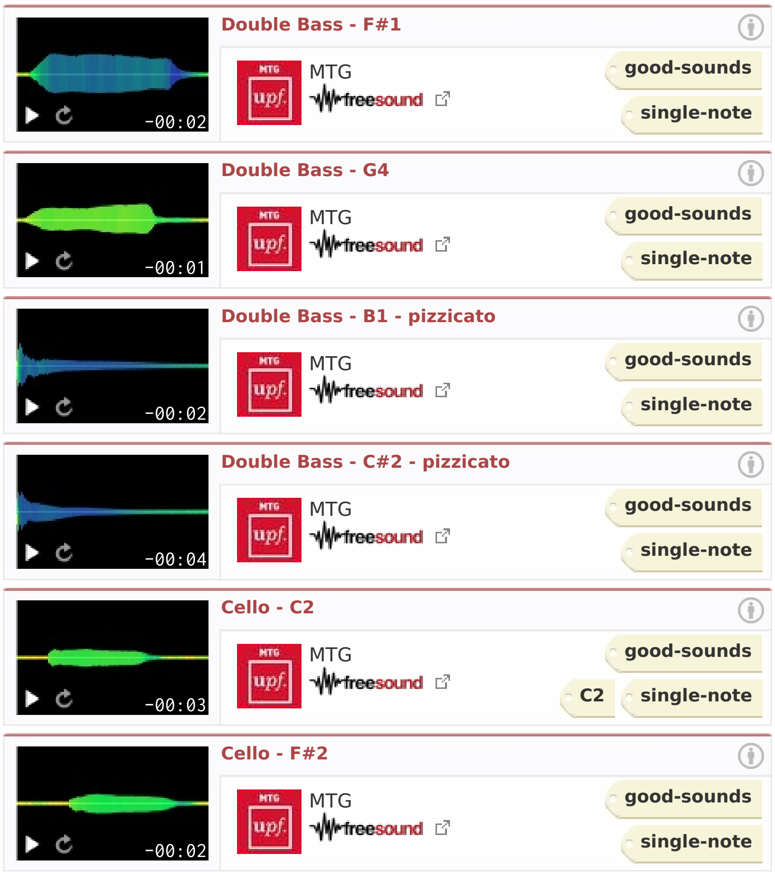
open opening drawers OWI

Most commons tags tagcloud:

door-bell Sell steel closing dry scrape office indiegamedev Money tearing bong drawer type old ring
write Video-Game typewriter foley dryer percussion glass pencil jangling games Coins cutting ding hair house
duct-tape kitchen sound ping short letter wooden typing scissors tape wood Gold ding-dong tap bag
pen campus-upf keyboard sticky rip indiedev zipper Game close ripping money metal squeak
mechanical coins drop Purchase key field-recording Currency electric microwave Realistic gamedeveloping Coin tear
videogame paper hit writing squeaky sfx keys coin jingle household gamedev door gaming
opening scribble gate creaking bell slide doorbell dong knock fabric open videogames OWI chime velcro
cupboard

3) Bowed string instrument

Example sounds:



Double Bass - F#1

MTG

good-sounds single-note

-00:02

Double Bass - G4

MTG

good-sounds single-note

-00:01

Double Bass - B1 - pizzicato

MTG

good-sounds single-note

-00:02

Double Bass - C#2 - pizzicato

MTG

good-sounds single-note

-00:04

Cello - C2

MTG

good-sounds C2 single-note

-00:03

Cello - F#2

MTG

good-sounds single-note

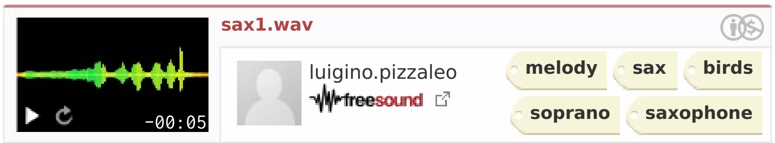
-00:02

Most commons tags tagcloud:

C3 non-vibrato c5 chordophone Gsharp3 midi-note-60 Asharp3 mezzoforte zoom-h2n slow tenuto phone acoustic C4 strings Double Gsharp2 bad Fsharp2 Dsharp4 violin dirty Gsharp4 bass Csharp2 midi-note-86 col-legno contrabass vsco-2 C2 loops g2 double zoom g-3 message descending classic glitchy double-bass vibrato midi-note-72 scale bowed d6 esharp3 orchestra plucked midi-note-67 pizzicato Stereo bowing tight-pizzicato a3 B3 midi-note-83 g-sharp-1 midi-velocity-31 b5 good-sounds G2 midi-note-57 b3 Acoustic ensemble g-sharp-3 multisample fsharp4 midi-note-69 note violin-section viola Fsharp3 Upright orchestral glitch midi-velocity-95 pluck b-4 cello recording A3 c4 Plucked cello-section cheap single-note Instrument neumann-U87 G3 Bass string-instrument midi-note-53 solo-violin Standup F4 D3 a4 midi-velocity-63 midi-note-59

4) Wind instrument, woodwind instrument

Example sounds:

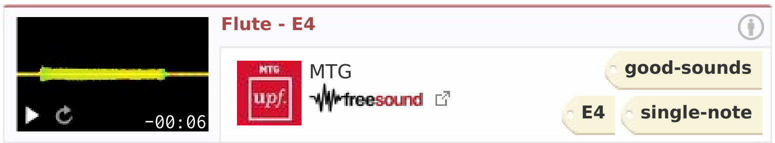

sax1.wav

luigino.pizzaleo

freesound

-00:05

melody sax birds soprano saxophone

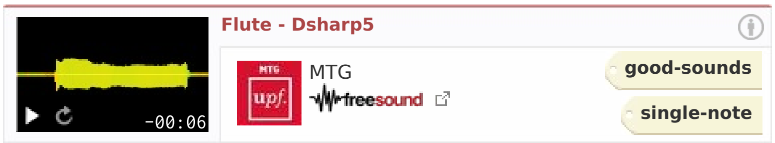

Flute - E4

MTG

freesound

-00:06

good-sounds E4 single-note

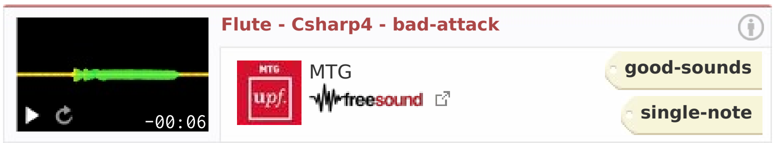

Flute - Dsharp5

MTG

freesound

-00:06

good-sounds single-note

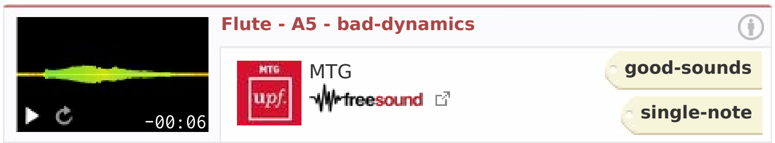

Flute - Csharp4 - bad-attack

MTG

freesound

-00:06

good-sounds single-note

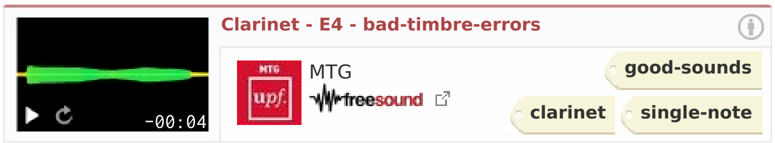

Flute - A5 - bad-dynamics

MTG

freesound

-00:06

good-sounds single-note


Clarinet - E4 - bad-timbre-errors

MTG

freesound

-00:04

good-sounds clarinet single-note

Most commons tags tagcloud:

medieval staccato Instruments midi-velocity-62 native-american Enatural bassoon multiphonics midi-note-52 soprano-saxophone neumann-U87 D6 woodwind Gmajor aerophone c-sharp-3 birds saxophone E4 Asharp3 midi-note-72 Csharp5 good-sounds E3 vibrato clarinet c4 D4 d-2 F4 midi-note-88 Fsharp5 flauto-dolce vibrato-sustain midi-velocity-95 multisample C4 non-vibrato Dsharp5 alto-sax minor single-note melody music midi-note-89 native-american-flute scale soprano e6 pad jazz sax oboe caroli soprano-sax B5 Asharp4 Gsharp4 midi-note-45 sustain B4 flute Fsharp4 midi-note-76 e-3 midi-velocity-105 vsco-2 mezzoforte midi-note-70 Asharp5 C6 vst D3 midi-note-58 african swing G3 d-sharp-2 tenuto Dnatural c5 B3 Acoustic midi-note-60 e5 sampled-instruments Fmajor schoeps-mk4 asharp4 woodwinds double-reed hip-hop G6 a2 Asharpmajor loop D5 midi-velocity-31 E5 recorder