

ISS projekt 2019

Vypracoval Adam Salih (xsalih01)

Projekt je psán v notebooku Jupyter.

Notebook obsahuje potřebné funkce, soubor `tools.py` pak obsahuje pomocné funkce pro formátovaný výpis, pomocnou datovou strukturu `AudioFile` pro zjednodušení práce s audio nahrávkami a dvě konstanty

- sentences
- queries

které obsahují pole nahrávek a hledaných výrazů.

Tento soubor načteme následujícím příkazem:

```
%run ./tools.py
```

1. Analýza vět

Nahrávky jsou použity pouze pro vyhodnocení tohoto ISS projektu.

```
def samplesAndLenght(audio):  
    samples = len(audio.data)  
    lenghtInSec = samples / audio.fs  
    return [audio.name, samples, lenghtInSec]
```

```
table(sentences, samplesAndLenght, ["Název", "Počet vzorků", "Čas  
[s]"])
```

Název	Počet vzorků	Čas [s]
si1734	64832	4.052
sx114	74389	4.6493125
sx204	92139	5.7586875

si1959	95552	5.972
sa2	70293	4.3933125
sa1	106475	6.6546875
sx384	62101	3.8813125
si1104	123883	7.7426875
sx24	63808	3.988
sx294	62443	3.9026875

2. Analýza vyhledávaných query

Pro nahrávky jsem si vybral slova `disapproves` uložené v souboru `q1.wav` a `occasionally` uložené v souboru `q2.wav`.

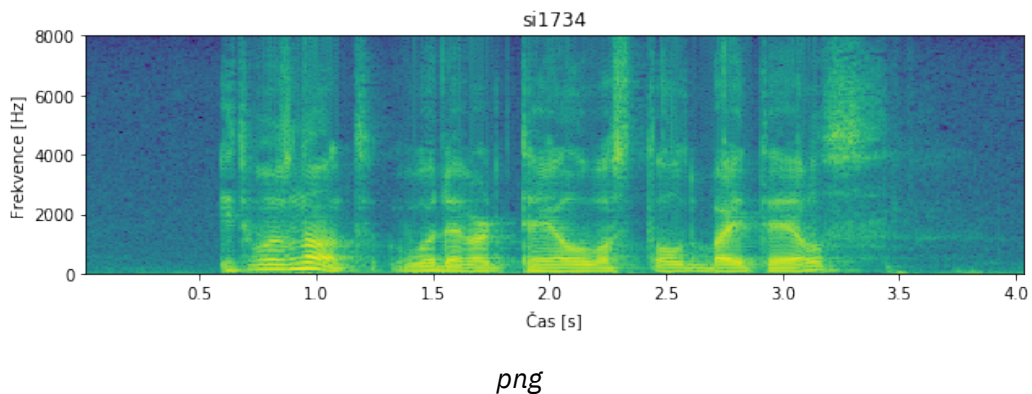
```
table(queries, samplesAndLenght, ["Název", "Počet vzorků", "Čas [s]"])
```

Název	Počet vzorků	Čas [s]
occasionally	13268	0.82925
disapproves	12685	0.7928125

3. Spectrogram

```
def spectrogramForAudio(audio):
    wlen = int(25e-3*audio.fs)
    wshift = int(10e-3*audio.fs)
    woverlap = wlen - wshift
    window = get_window("hamming", wlen)
    f, t, sgr = spectrogram(audio.data, audio.fs, window, wlen,
                             woverlap, 512)
    sgr_log = 10 * log10(sgr+1e-20)
    return f, t, sgr_log
```

```
plotSpectrogram(spectrogramForAudio(sentences[0]), sentences[0].name)
```



4. Parametry

Parametry jsem se rozhodl vytvořit pomocí lineární banky filtrů.

```
def generateFeatures(audio):
    fsgr, tsgr, sgr = audio.spectrogram
    final = []
    tmp = [0 for i in sgr[0]]
    f = [f for i, f in enumerate(fsgr) if i%16 == 0]
    f.pop(-1)
    for i, spec in enumerate(sgr):
        if i%16 == 0 and i != 0:
            final.append(tmp)
            tmp = spec
        else:
            for j, value in enumerate(spec):
                tmp[j] += value
    return f, tsgr, final
```

5. Výpočet podobnostního skóre

```
def scoreFor(audio, query, sampleNumber):
    q = transpose(query.features[2])
    f = transpose(audio.features[2])
    score = []
    if 0 > sampleNumber or sampleNumber > len(f) - len(q):
```

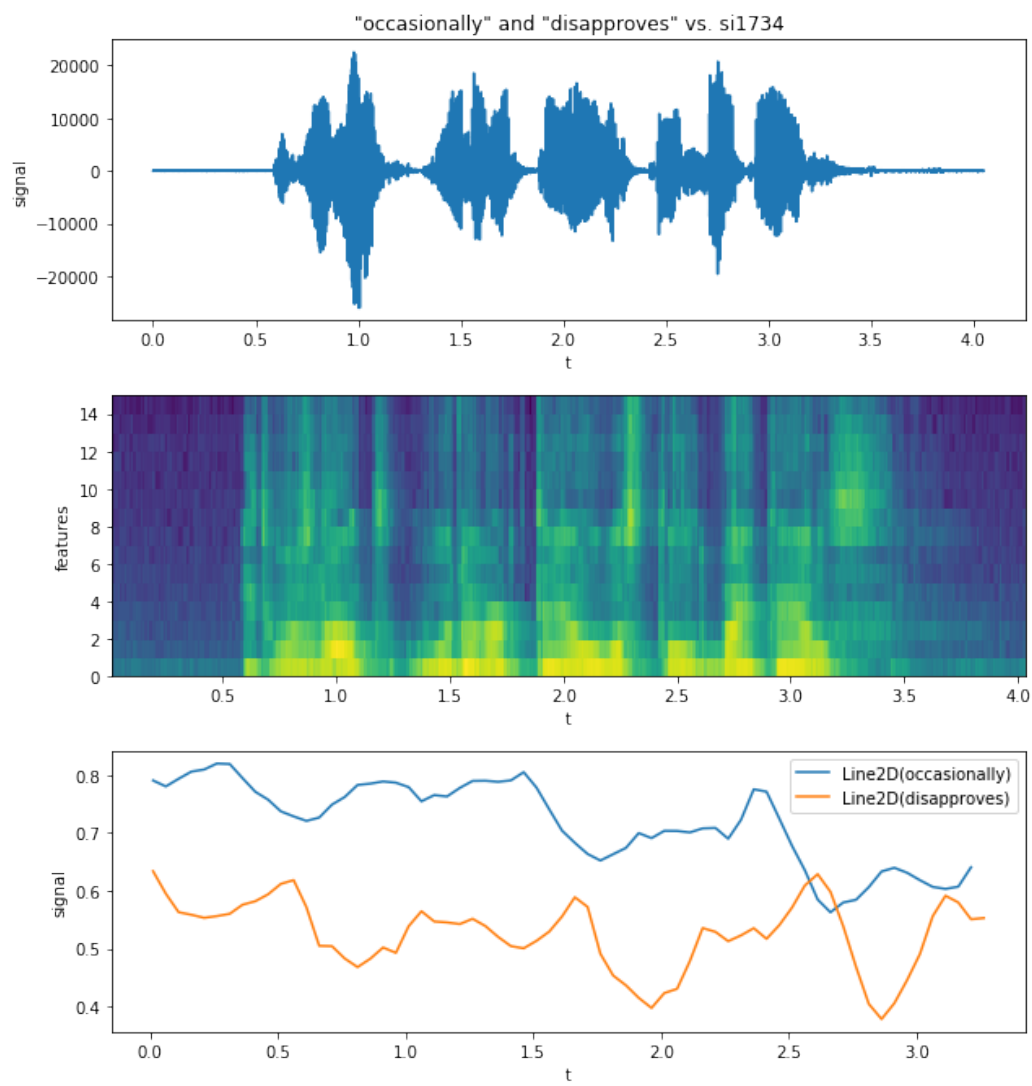
```
        return None
    delta = int(sampleNumber)
    fc = f[delta: delta + len(q)]
    p = 0.0
    for i in range(0, len(q)):
        p += pearsonr(q[i], fc[i])[0]
    return p / len(q)
```

6. grafický výstup

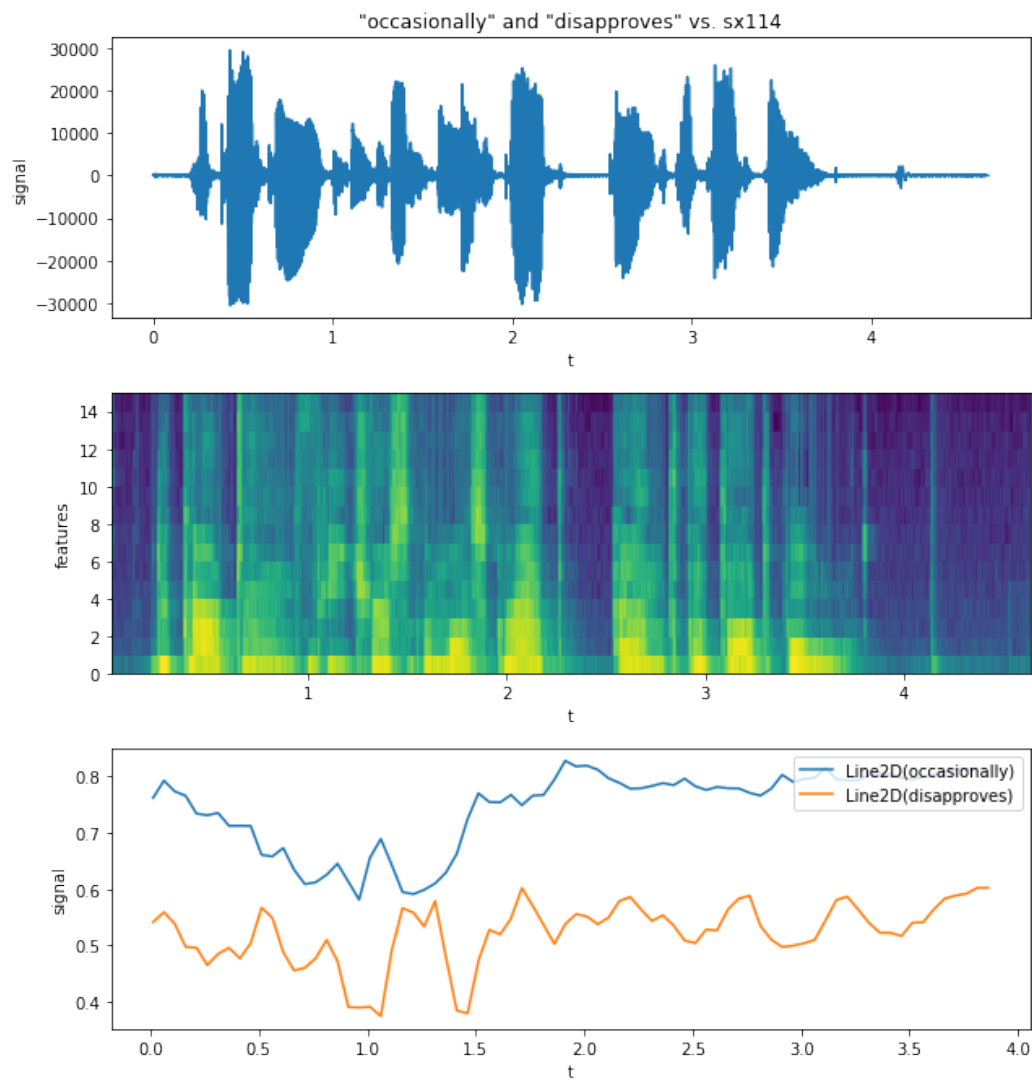
Slovo occasionally se objevuje v audionahrávce sx204

Slovo disapproves se objevuje v audionahrávce si1104

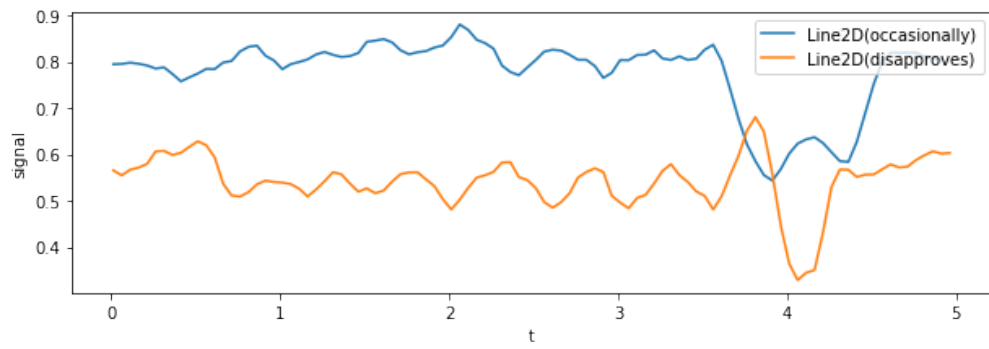
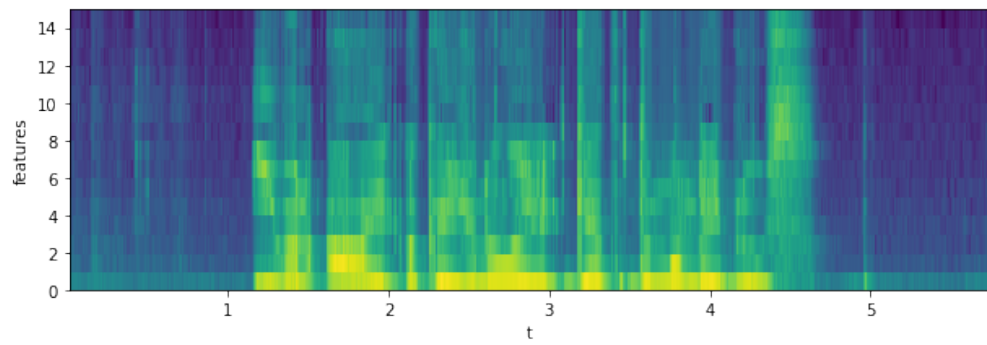
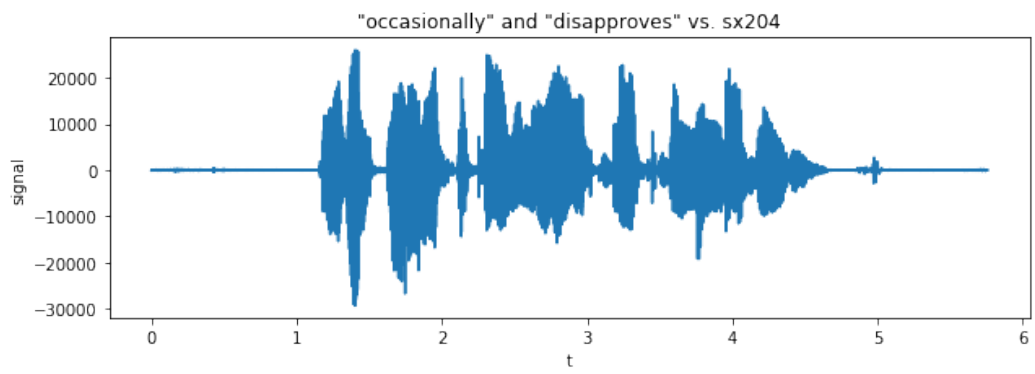
```
for sentence in sentences:
    plotStatistics(sentence)
```



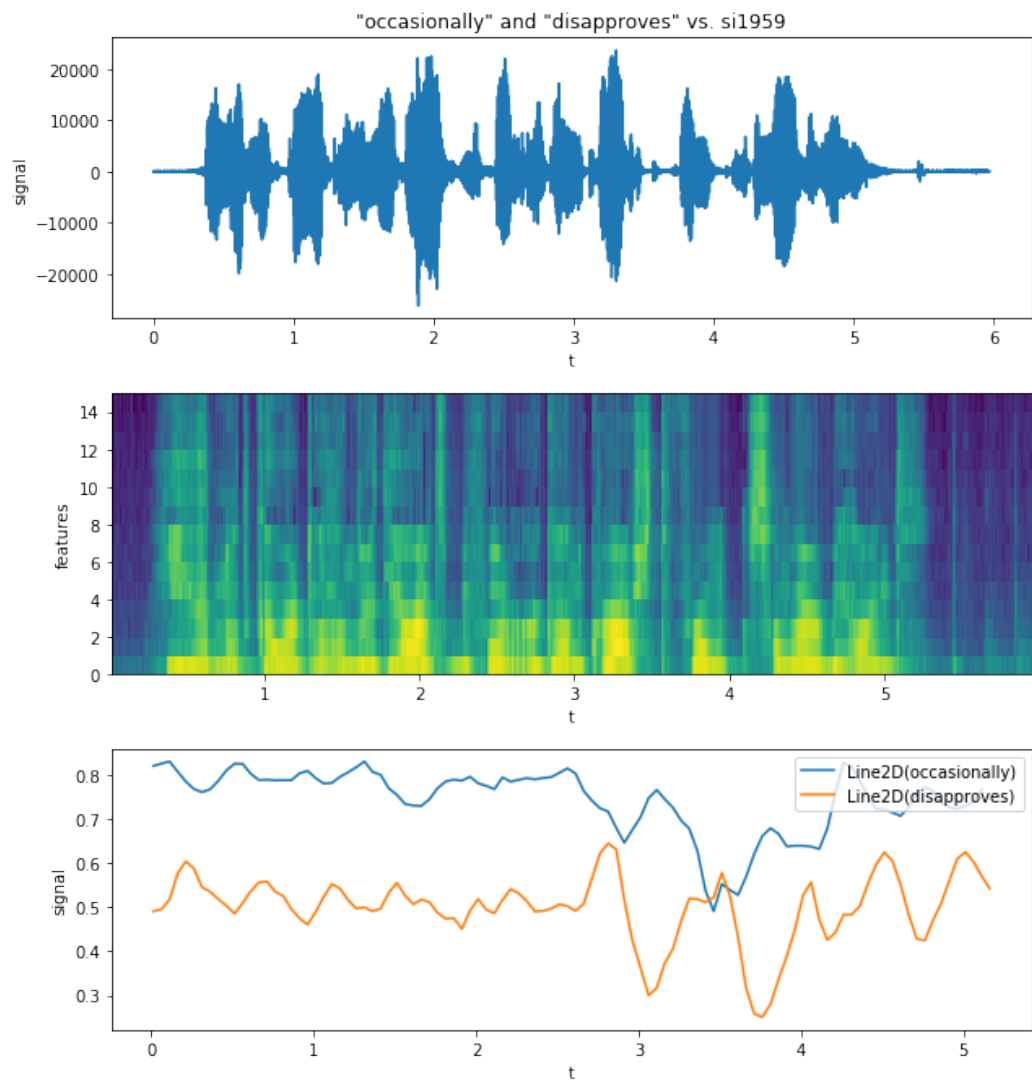
png



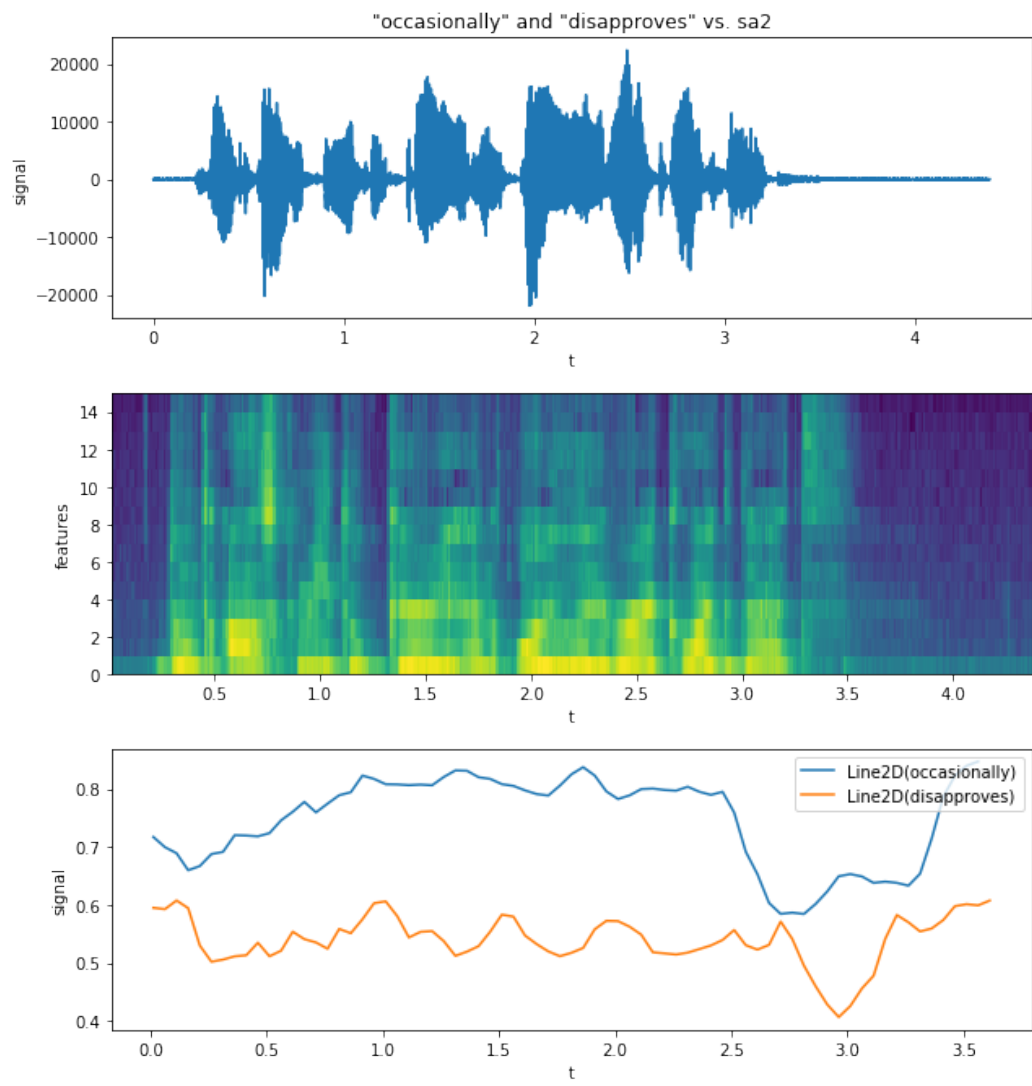
png



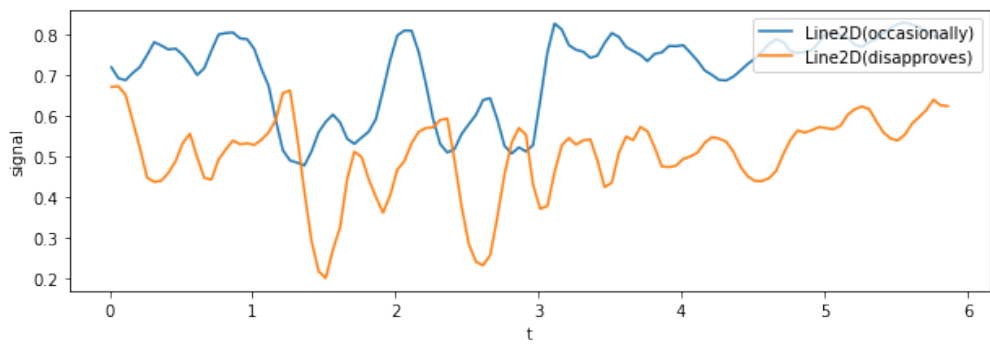
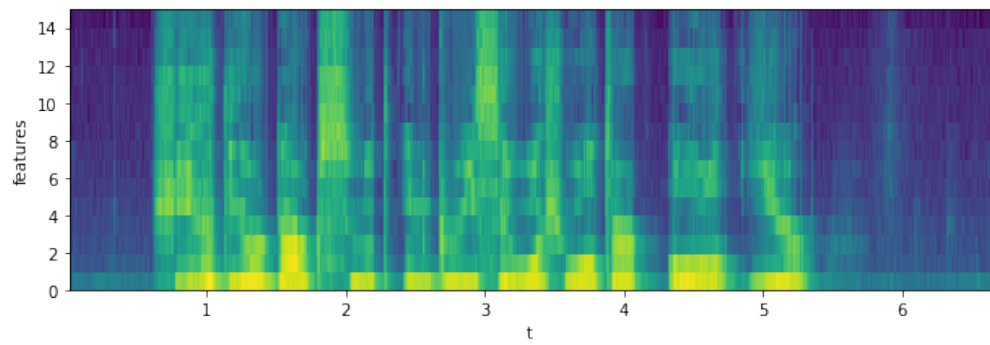
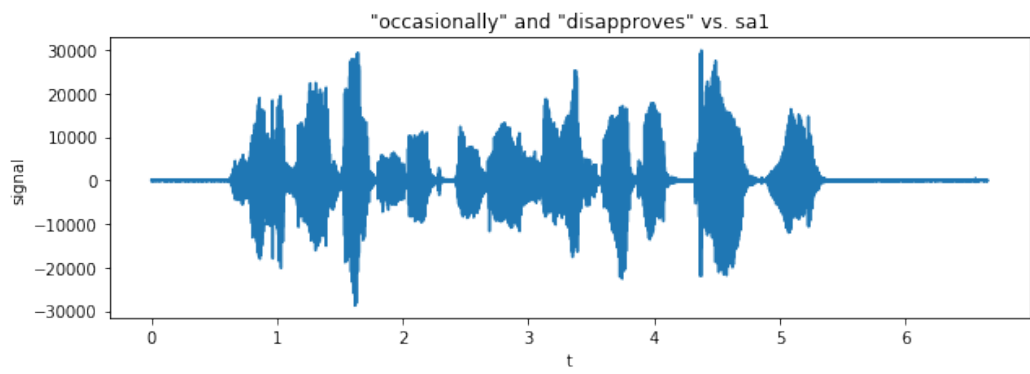
png



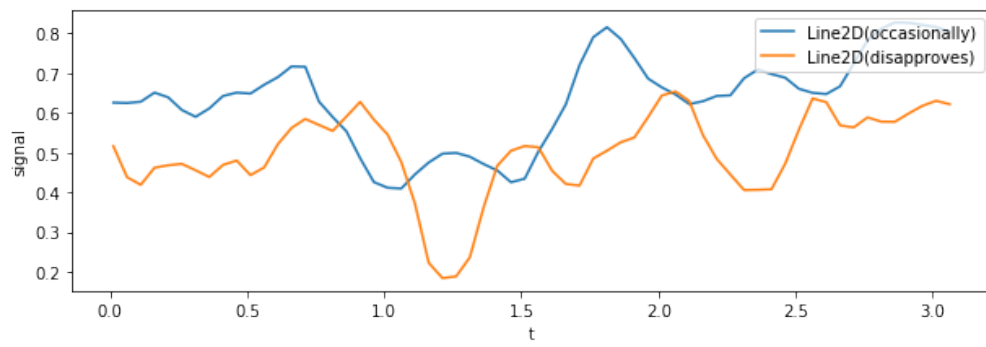
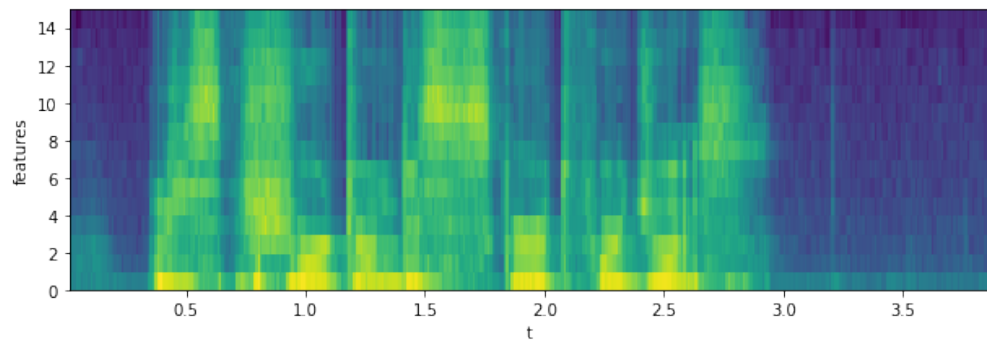
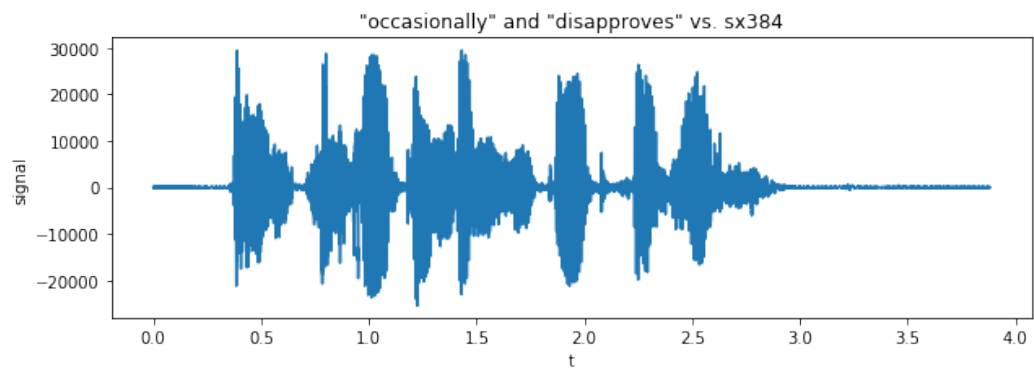
png



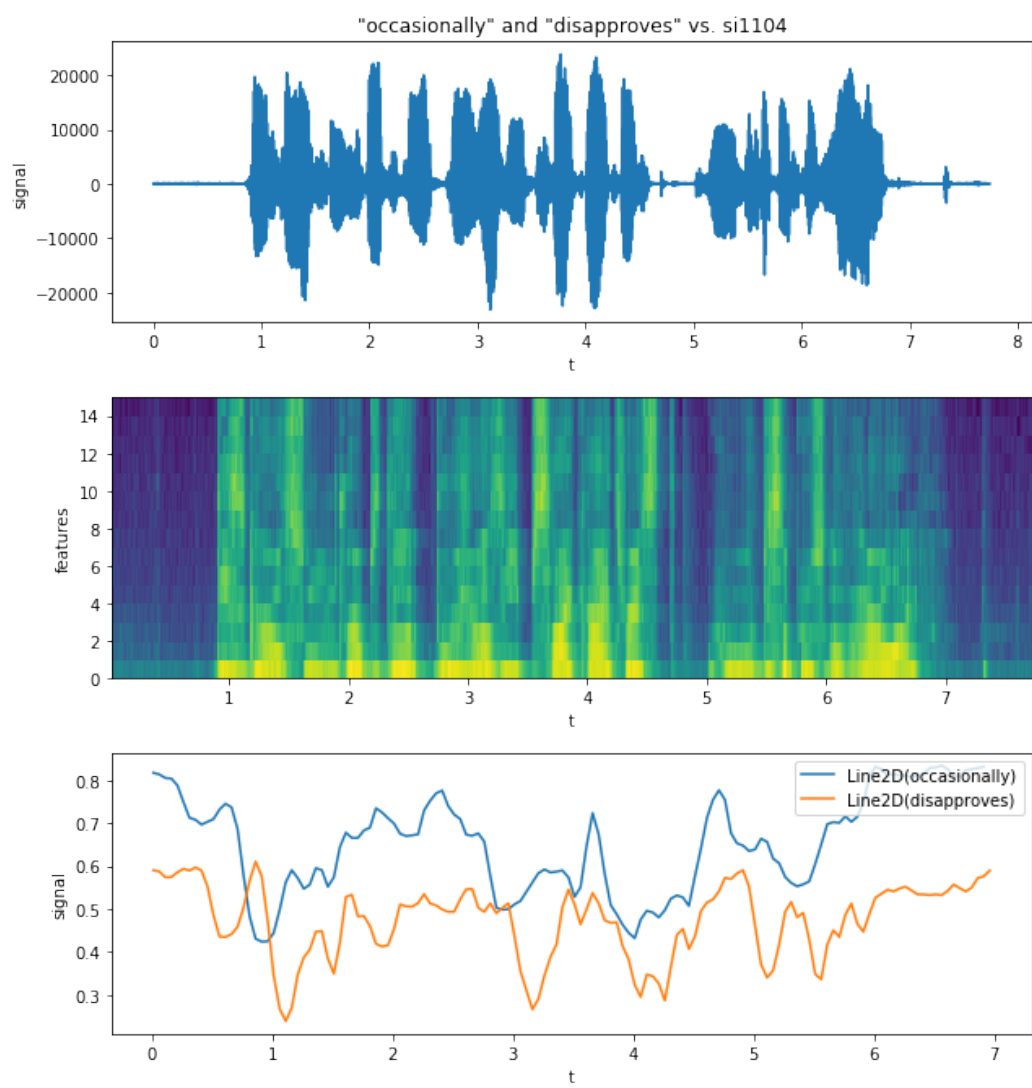
png



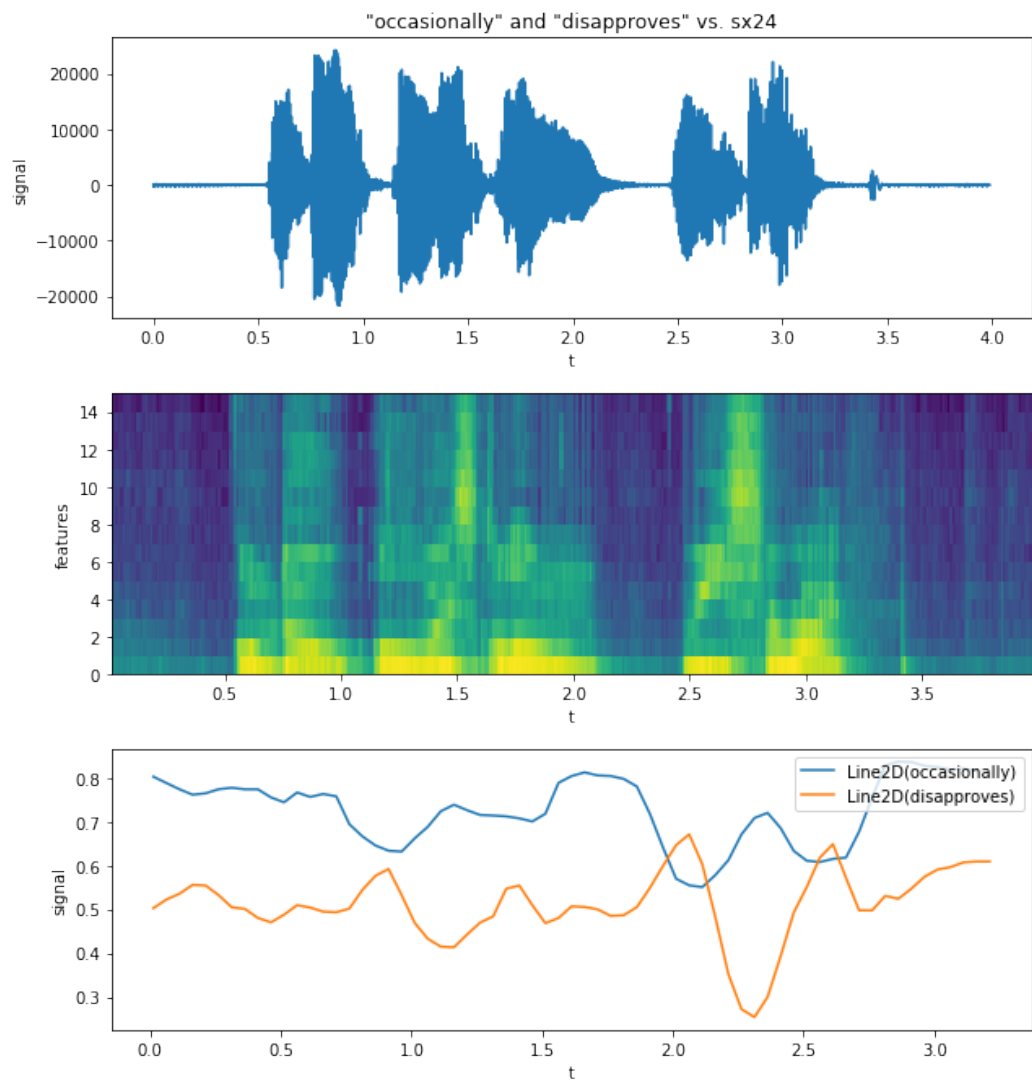
png



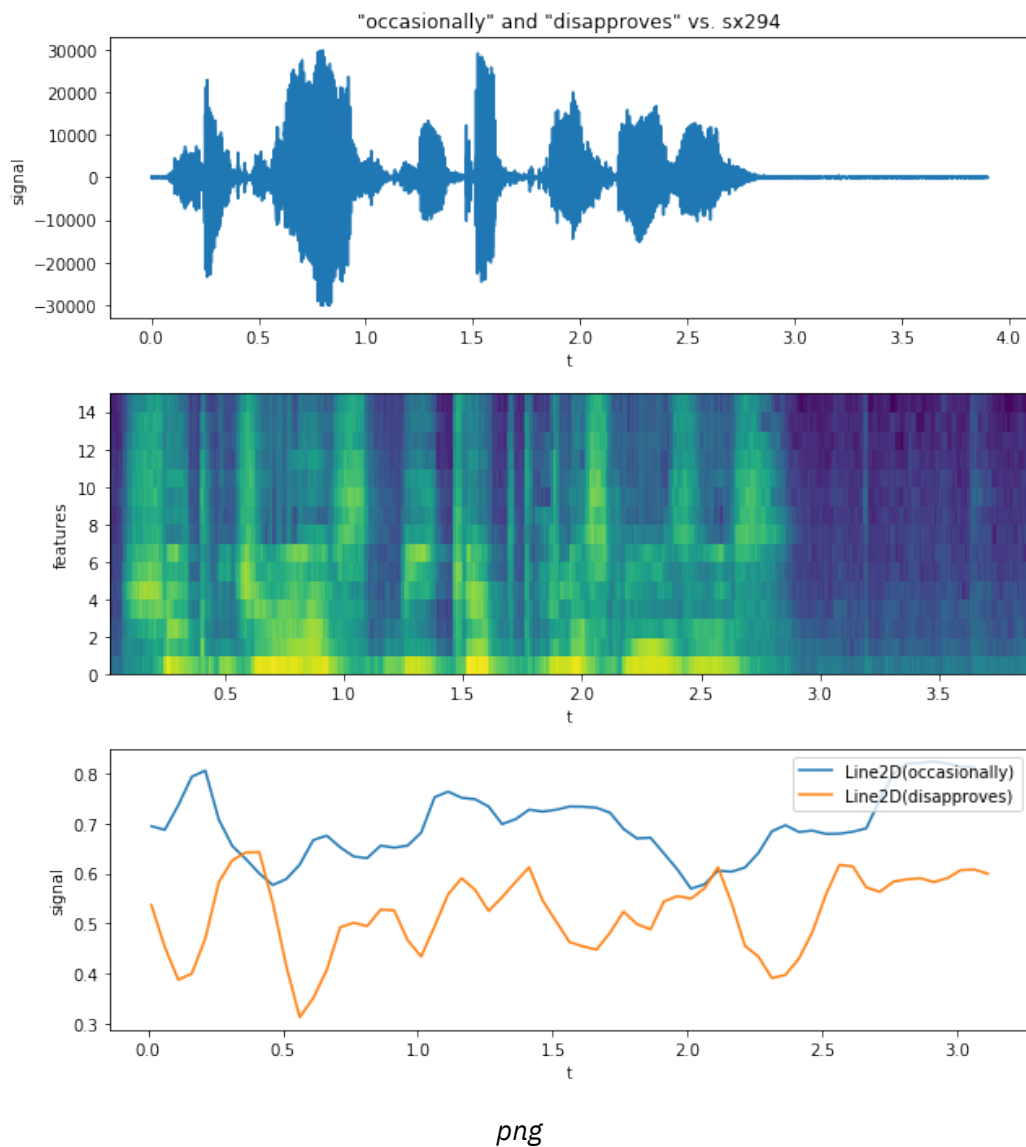
png



png



png



7. Threshold

Hodnotu prahů jsem si určil tak, že jsem si vybral první 'pěkné' číslo, které je o něco málo menší, než maxiální hodnota skóre ve větách, které dané slovo obsahovaly.

```
for query in queries:
    if query.name == "occasionally":
        query.treshhold = 0.85
    elif query.name == "disapproves":
        query.treshhold = 0.61
```

8. Detekce query

```
def containsQuery(audio):
    contains = []
    for query in queries:
        score = [audio.scoreFor(query, i) for i in
range(len(audio.features[1])) if i%5 == 0]
        score = [nonNone for nonNone in score if nonNone != None]
        if max(score) > query.treshold:
            contains.append(True)
        else:
            contains.append(False)
    return [audio.name] + contains
```

```
table(sentences, containsQuery, ["název"] + [query.name for query in
queries])
```

název	occasionally	disapproves
si1734	False	True
sx114	False	False
sx204	True	True
si1959	False	True
sa2	False	False
sa1	False	True
sx384	False	True
si1104	False	True
sx24	False	True
sx294	False	True

9. Závěr

První query se vyhledala správně, u druhé query byl práh nastaven tak nízko, že se falešně pozitivně vyhodnotil nález i u vět, které slovo `disapproves` neobsahovaly. Domnívám se, že je vinnou délka nahrávky. Je možné, že jsem vyslovení slova

neumíslně protáhl a tím jsem posunul nalézané vzorky o offset, který skóre hledaného výrazu efektivně snížili na napoznatelnou hodnotu.

Toto by se dalo vyřešit transformací hledané query, které by délku smrštilo, nebo natáhlo a tím zvýšilo možnost hledaný výraz najít.

