

ISS Projekt 2019 – protokol

Tomáš Ďuriš (xduris05)

1. úloha

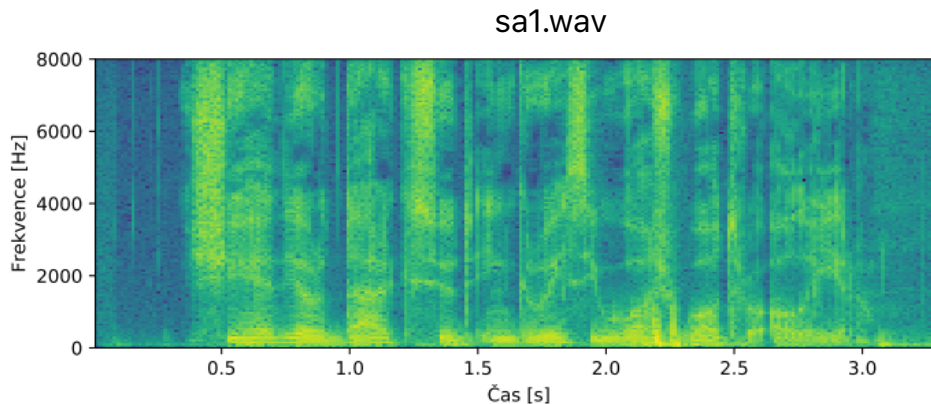
Názov vety	Dĺžka vo vzorkoch	Dĺžka v sekundách
sa1.wav	53227	3.33s
sa2.wav	62443	3.90s
si1013.wav	74731	4.67s
si1643.wav	50496	3.16s
si2273.wav	49472	3.09s
sx113.wav	53909	3.37s
sx203.wav	54933	3.43s
sx23.wav	55275	3.45s
sx293.wav	63808	3.99s
sx383.wav	56981	3.56s

Informácia o použití dát: c) súhlasím s využitím dát pre všetky varianty (a), b), tvorba voľne dostupnej databáze "Czenglish TIMIT"

2. úloha

Vybrané slová ("queries")	Dĺžka vo vzorkoch	Dĺžka v sekundách
musicians	9661	0.60s
beverages	11705	0.73s

3. úloha



Výpočet:

Pre výpočet som sa rozhodol použiť kód od Katky Žmolíkovej, keďže som projekt robil taktiež v Pythone, doplnil som potrebné parametre a vyššie uvedený spectrogram vykreslil. Prikladám kód:

```
s, fs = sf.read('sa9.wav') #načítanie súboru

#doplnenie parametrov do funkcie spectrogram
f, t, sgr = spectrogram(s, fs, nperseg=400, noverlap=240, nfft=511)

#výpočet logaritmickeho výkonového spektra
sgr_log = 10 * np.log10(sgr+1e-20)
```

4. úloha

Výpočet parametrov ("features") som sa rozhodol riešiť, podľa doporučenia v zadaní, pomocou lineárnej banky filtrov, ktorá produkuje pre každý rámec 16 koeficientov. Prikladám kód:

```
C = 0
# prejdeme maticu vety po stĺpcoch
for a in range(len(matrix_sentence[0])):
    # pre každý stlpec sčítame 16 prvkov
    for b in range(16):
        #sčítavame prvky (0-15,16-31,...,240-255)
        for c in range(c,c+16,1):
            #ukladáme na požadované indexy aby sme dosiahli maticu 16xXXX
            matrix_sentence[b,a] = matrix_sentence[b,a] + sgr_log[c,a]
        c = c + 1
    c = 0
```

5. úloha

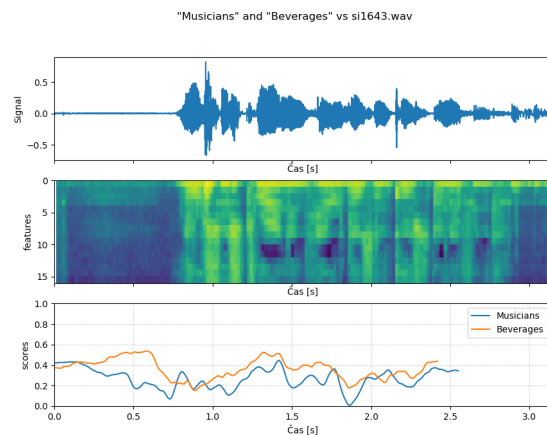
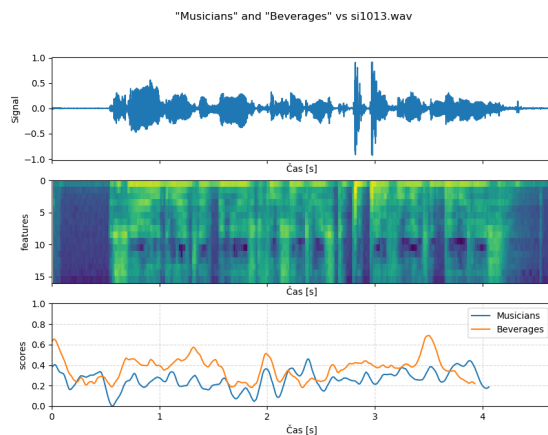
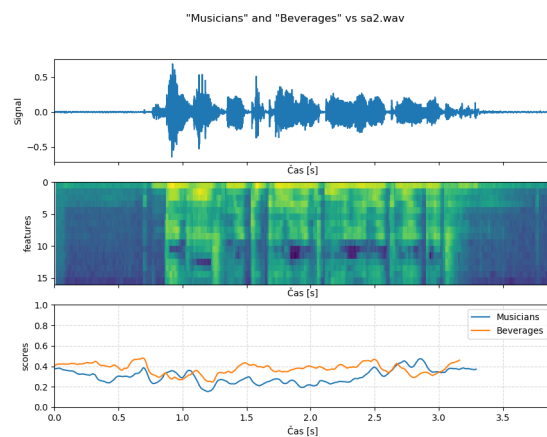
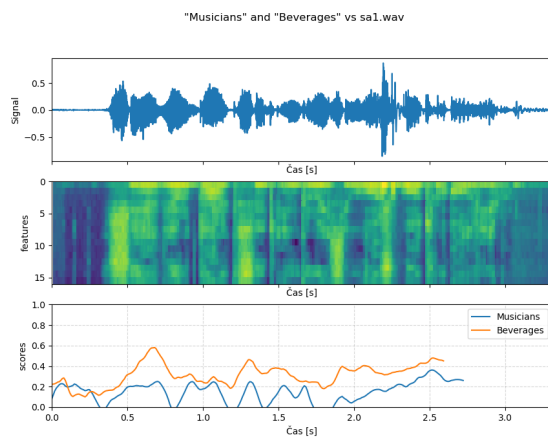
Pre výpočet skóre kľúčové slova som sa rozhodol taktiež použiť metódu zo zadania a nechal som maticu parametrov (matrix_query1) prechádzať okolo matice parametrov vety (matrix_sentence) a *pp* som si zvolil 1. Matice som transponoval na veľkosť 16xXXX a 16xYYY. Prikladám kód:

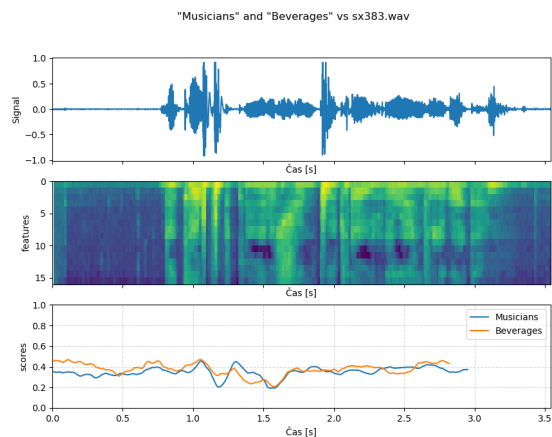
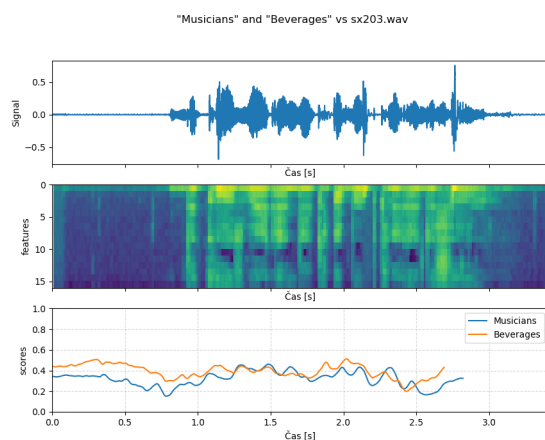
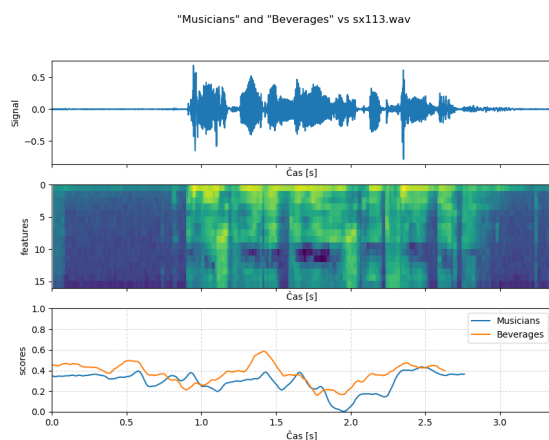
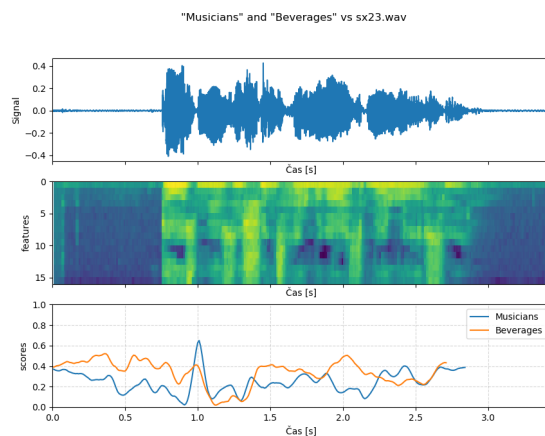
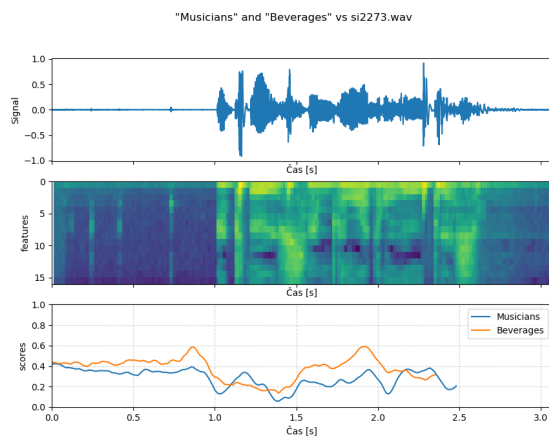
```

for d in range(len_final): #vykonávame pre dĺžku (matice vety – matice slova)
    sum = 0
    for i in range(len(matrix_query1)): #vykonávame pre dĺžku matice slova
        finalmatrix, secondvalue = pearsonr(matrix_query1[i], matrix_sentence[d + i])
        sum = sum + finalmatrix
    scores.append(sum / np.size(matrix_query1,0))
#prípajáme posčítavané prvé parametre z funkcie pearsonr

```

6. úloha





7. úloha

Rozhodol som sa určiť prah na základe bližšieho preskúmania obrázkov, a porovnania očakávaného výskytu. Moje výsledné prahy:

Vybrané slová ("queries")	Prah
musicians (q1)	0.6
beverages (q2)	0.66

8. úloha

Veta	Musicians (q1)	Beverages (q2)
sa1.wav	nie	nie
sa2.wav	nie	nie
si1013.wav	nie	áno (55520 až 66880)
si1643.wav	nie	nie
si2273.wav	nie	nie
sx113.wav	nie	nie
sx203.wav	nie	nie
sx23.wav	áno (16000 až 25280)	nie
sx293.wav	nie	áno (8480 až 19840)
sx383.wav	nie	nie

Pri nahrávke sx293.wav sa naozaj slovo musicians na danom úseku vyskytovalo, takže zhoda je správna

Pri nahrávke si1013.wav sa naozaj slovo beverages na danom úseku vyskytovalo, takže zhoda je správna

Pri nahrávke sx293.wav sa slovo beverages na danom úseku nevyskytovalo, takže zhoda nie je správna.

9. úloha

Záver:

Myslím, že tento projekt mi celkom vyšiel a výsledky splňajú zhruba to čo som očakával, sú tam menšie nepresnosti hlavne pri odhaľovaní "výskytu queries", ale podarilo sa mi to určiť správne, až na jednu zhodu. Myslím si, že toto môže byť spôsobené hlavne rušivými vonkajšími vplyvmi a nahrávaním na mobilný mikrofón, avšak otestoval som aj vystrihnutie slova priamo z vety a následné použitie kde my vyšlo dokonca skóre 1. Z tohto pokusu by som usúdil, že algoritmus je správny a v teoretickej rovine funguje.

Projekt sa mi veľmi páčil a bavilo ma ho robiť, hodnotil by som ho ako jeden z najzaujímavejších na tejto škole a podarilo sa mi pochopiť ako v praxi fungujú viaceré veci, s ktorými som sa keď tak stretol vždy len v teoretickej rovine. Jediná vec čo ma mrzí bola tá, že projekt sa prekrýval s projektami, ktoré bolo nutné robiť, aby bol dosiahnutý zápočet a teda nebol čas poskúšať kvalitnejšie nahrávky, lepšie prostredie,...