

Seminární práce II
z předmětu

Počítačové zpracování signálu (KI/PZS)

Klasifikace zvukových záznamů

Klasifikace zvukových záznamů

Zadání:

Ve zdrojové databázi najdete celkem 208 hlasových záznamů písmene a. Pomocí Vámi vybrané techniky v časové nebo frekvenční oblasti klasifikujte zvukové záznamy na dobré a patologické. V případě patologických poté klasifikujte jednotlivé poruchy. Jejich výčet najdete buď v hlavičkových souborech nebo v propisu databáze. Pro klasifikaci do jednotlivých skupin použijte veškeré techniky, které jste si v rámci kurzu osvojili včetně Fourierovy a keprstrální analýzy. Úspěšnost Vašeho postupu porovnejte s anotacemi, resp. rozřazením do skupin, které provedli experti, kteří data pořizovali.

1. Data

Data z jedné hlasové nahrávky se skládají ze souborů:

.dat – samotný záznam signálu

.hea – metadata o signálu

-info.txt – informace o pacientovi

Nejprve jsem se podívala na obsah souboru .hea pro první záznam signálu.

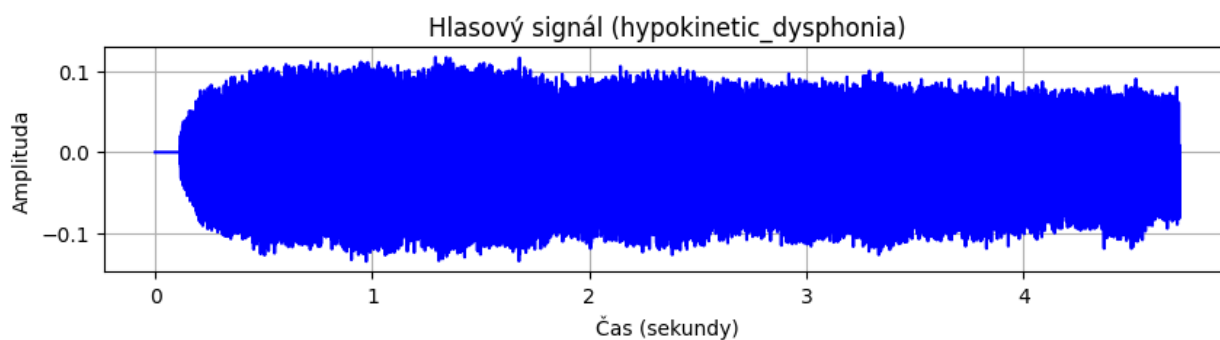
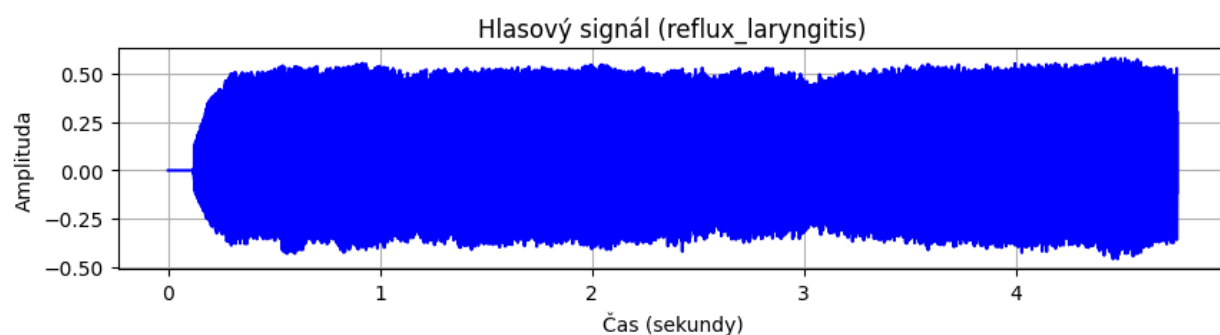
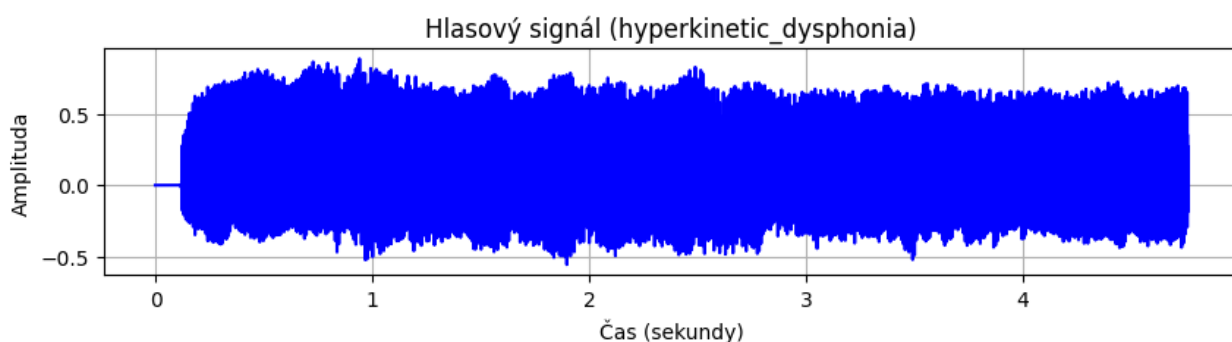
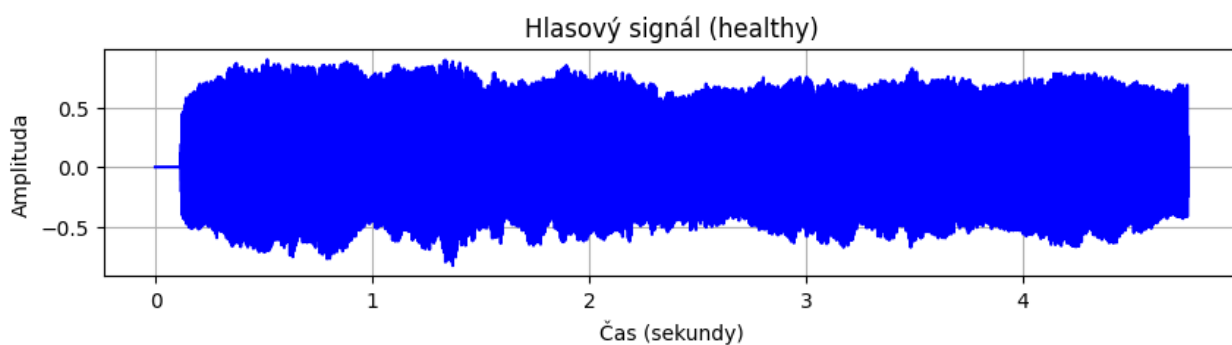
	Informace	Hodnota
0	Název souboru	voice001
1	Vzorkovací frekvence (Hz)	8000
2	Délka signálu (vzorky)	38080
3	Počet kanálů	1
4	Názvy kanálů	voice
5	Jednotky	NU

Poté jsem se podívala na obsah souboru -info.txt. Z celého popisu pacienta mě zajímala jeho diagnóza.

Vytvořila jsem seznam všech diagnóz, které se v souborech vyskytují bez dalších informací o každé z nich, které byly zapsány v závorkách.

Zjistila jsem, že ve 208 souborech se vyskytují pouze 4 kategorie diagnóz:

1. Healthy
2. Hyperkinetic dysphonia
3. Reflux laryngitis
4. Hypokinetic dysphonia



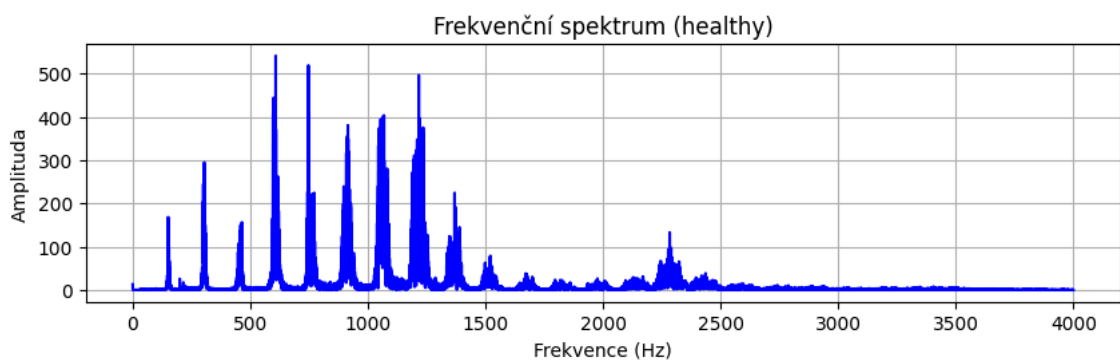
2. Frekvenční spektrum

Abych identifikovala vlastnosti hlasu a pochopila, čím se liší zdravé hlasy od patologických, bylo nutné určit klíčové frekvenční charakteristiky.

K tomu jsem použila Fourierovu analýzu, která umožňuje rozložit hlasový signál na jeho frekvenční složky.

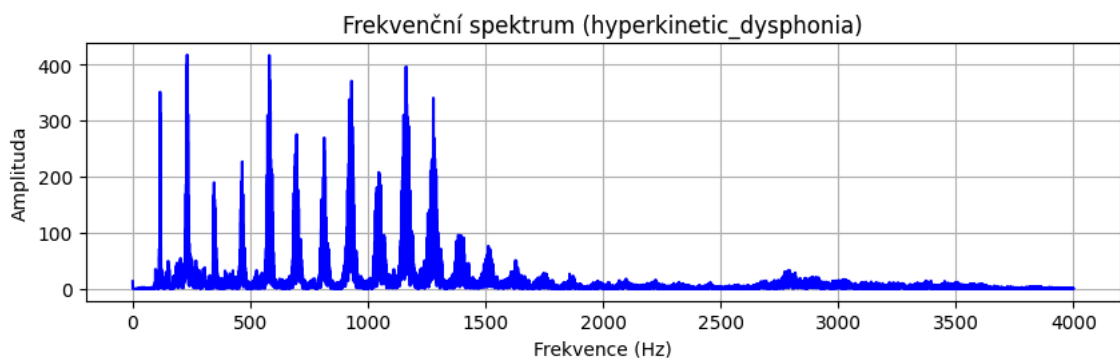
1. Healthy (Zdravý hlas)

Spektrum obsahuje jasně definované harmonické složky s rovnoměrným poklesem amplitudy, což svědčí o stabilní hlasové funkci.



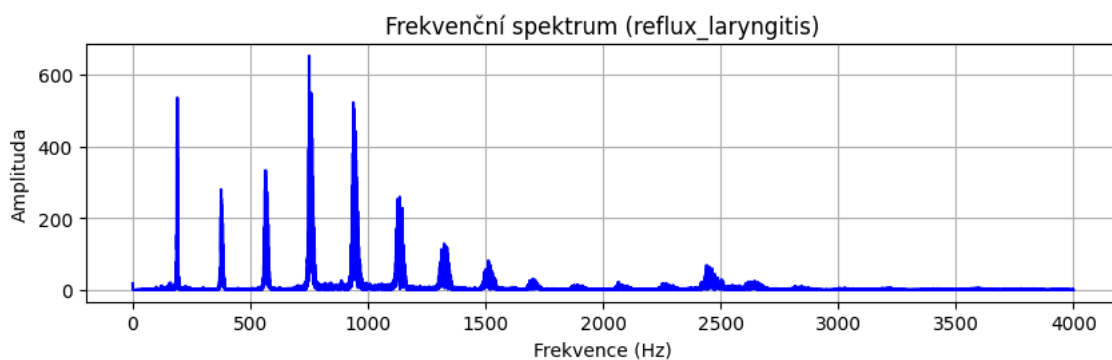
2. Hyperkinetic dysphonia (Hyperkinetická dysfonie)

Ve spektru se vyskytuje zvýšená úroveň šumu a dodatečné frekvenční píky, které naznačují nadměrné napětí hlasivek a přidání nekontrolované oscilace.



3. Reflux laryngitis (Refluxní laryngitida)

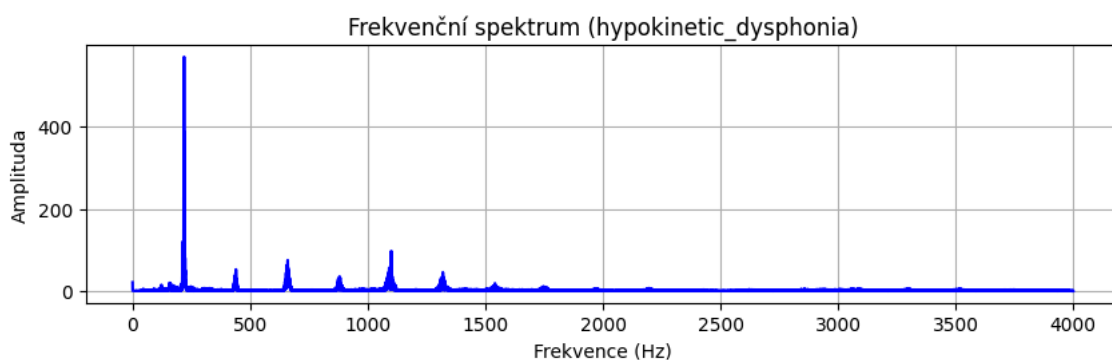
Spektrum ukazuje sníženou amplitudu ve vyšších frekvencích, což může být spojeno s zánětlivými procesy ovlivňujícími vibrační schopnost hlasivek



4. Hypokinetic dysphonia (Hypokinetická dysfonie)

Velmi slabé vysoké frekvence.

To odpovídá slabému hlasu, který je typický pro hypokinetickou dysfonii.



3. Analýza hlavních frekvenčních peaků

Dalším krokem je analýza hlavních frekvenčních peaků ve spektru, abych zjistila: základní frekvenci (F_0) – základní frekvence kmitání hlasivek, rozptyl frekvenčních peaků – míra chaotičnosti rozložení frekvencí a pravidelnost harmonických složek – doplňkové frekvence ovlivňující tón hlasu.

Zdravé hlasy vykazují nejvyšší pravidelnost harmonik a vyváženou F_0 , což potvrzuje jejich stabilitu.

```

healthy
Základní frekvence ( $F_0$ ): 149.37 Hz
Pravidelnost harmonických: 50.37
Variabilita základní frekvence: 315.43
  
```

Hyperkinetická dysfonie se vyznačuje sníženou F_0 a vysokou variabilitou, což může naznačovat nestabilní vibrace hlasivek.

```
hyperkinetic_dysphonia  
Základní frekvence ( $F_0$ ): 114.92 Hz  
Pravidelnost harmonických: 14.85  
Variabilita základní frekvence: 340.12
```

Refluxní laryngitida vykazuje zvýšenou F_0 a nízkou pravidelnost harmonik, což naznačuje zánětlivé procesy ovlivňující hlas.

```
reflux_laryngitis  
Základní frekvence ( $F_0$ ): 186.13 Hz  
Pravidelnost harmonických: 30.27  
Variabilita základní frekvence: 259.56
```

Hypokinetická dysfonie má nejvyšší F_0 , nízkou pravidelnost harmonik a téměř nulovou variabilitu, což může být silným indikátorem patologie.

```
hypokinetic_dysphonia  
Základní frekvence ( $F_0$ ): 215.68 Hz  
Pravidelnost harmonických: 0.22  
Variabilita základní frekvence: 2.03
```

4. Kepstrální analýza

Po analýze frekvenčního spektra a hlavních frekvenčních peaků přecházím ke kepstrální analýze, abych vyhodnotila periodičnost frekvenčních peaků.

Nejprve vezmu spektrum signálu, lagaritmizuji ho a poté pomocí zpětné Fourierovy transformace získám kepstr - periodičnost frekvenčních vrcholů. Poté vypočítám kepstrální vrchol, který ukazuje přítomnost patologií nebo šumu (čím nižší hodnota, tím více šumu), a střední hodnotu kepstrálního vrcholu, která ukazuje, jak rovnoměrné jsou frekvenční oscilace.

```

healthy
{'Kepstrální vrchol (CPP)': 0.4698635190583461, 'Průměrná hodnota kepsra (CEPS-Mean)': 6.758879894261984e-05}

hyperkinetic_dysphonia
{'Kepstrální vrchol (CPP)': 0.46849538432126037, 'Průměrná hodnota kepsra (CEPS-Mean)': 6.875924078725585e-05}

reflux_laryngitis
{'Kepstrální vrchol (CPP)': 0.4458947045596854, 'Průměrná hodnota kepsra (CEPS-Mean)': 7.595083791840906e-05}

hypokinetic_dysphonia
{'Kepstrální vrchol (CPP)': 0.2023185086793501, 'Průměrná hodnota kepsra (CEPS-Mean)': 8.087600510342135e-05}

```

5. Vyhodnocení stavu hlasu

Po extrakci spektrálních a kepstrálních charakteristik přecházím k závěrečnému hodnocení stavu hlasu.

Funkce `is_voice_healthy()` analyzuje klíčové parametry a určí, zda je hlas zdravý.

Pro klasifikaci hlasu jako zdravého nebo patologického nastavím prahové hodnoty klíčových parametrů vzhledem k dříve zjištěným hodnotám:

```

# prahové hodnoty pro zdravý hlas
cpp_threshold = 0.4
ceps_mean_threshold = 0.00007
entropy_threshold = 6.5
centroid_min, centroid_max = 1100, 1800
spread_threshold = 800
skewness_threshold = 1.5

```

Hlas byl považován za zdravý, pokud splňoval alespoň 5 z 6 kritérií.

Poté jsem ověřila každý záznam a vypočítala konečnou úspěšnost

	Correct	Wrong	Total	Success Rate (%)
0	140	68	208	67.307692