

Seminární práce z předmětu KI/PZS

Počítačové zpracování signálu II.

JAN HRANIČKA

Osobní číslo: F24132

Obsah

1	Zadání	2
2	Použitá data	2
3	Postup řešení	2
3.1	Předzpracování signálu	2
3.2	Extrakce příznaků	2
3.3	Klasifikace	3
4	Výsledky	3
4.1	Binární klasifikace	3
4.2	Vícetřídní klasifikace patologií	3
4.3	Grafické výstupy	4
5	Závěr	4

1 Zadání

Cílem seminární práce je zpracování hlasových signálů z databáze VOICED Database dostupné na portálu PhysioNet. Databáze obsahuje celkem 208 hlasových záznamů samohlásky *a*, z nichž část odpovídá zdravým subjektům a část subjektům s patologickými poruchami hlasu.

Úkolem práce je:

- načíst a analyzovat hlasové signály a jejich anotace,
- extrahovat vhodné příznaky v časové, frekvenční a kepstrální oblasti,
- klasifikovat signály na zdravé a patologické,
- u patologických signálů se pokusit o klasifikaci jednotlivých poruch,
- vyhodnotit úspěšnost klasifikace porovnáním s anotacemi expertů.

Součástí práce jsou také grafické výstupy demonstруjící zpracování signálu a tabulky s výsledky klasifikace.

2 Použitá data

Zdrojem dat je databáze VOICED Database z portálu PhysioNet. Každý subjekt je reprezentován:

- hlavičkovým souborem obsahujícím základní informace a diagnózu,
- textovým souborem,
- datovým souborem se samotným hlasovým signálem.

Pro načítání dat byl využit balík `wfdb`. Pro samotné zpracování signálu nebyly použity žádné předpřipravené nástroje z tohoto balíku v souladu se zadáním práce.

3 Postup řešení

3.1 Předzpracování signálu

Hlasové signály byly normalizovány a rozděleny do rámců pevné délky. Pro odstranění stejnosměrné složky byl od signálu odečten jeho střední hodnota.

3.2 Extrakce příznaků

Z každého rámce byly extrahovány následující příznaky:

- časové příznaky (energie, rozptyl),
- frekvenční příznaky získané pomocí Fourierovy transformace,
- kepstrální příznaky popisující spektrální obálku signálu.

Výsledné příznaky byly agregovány do příznakových vektorů reprezentujících jednotlivé hlasové záznamy.

3.3 Klasifikace

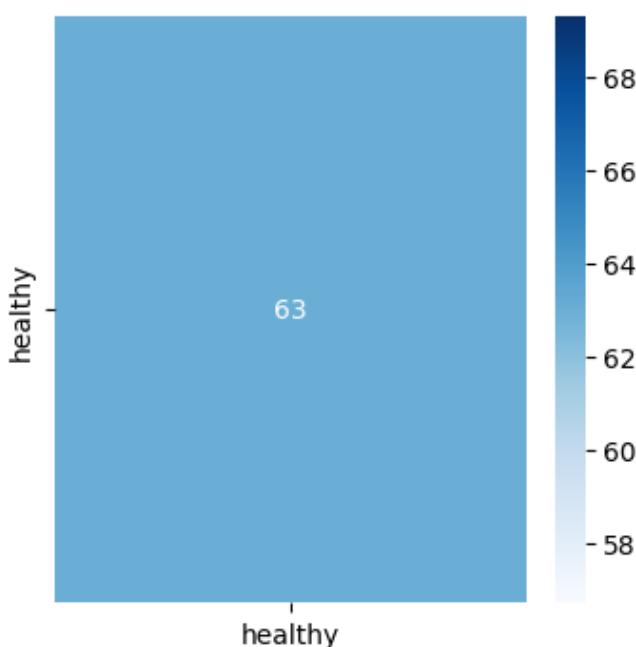
Pro klasifikaci byl zvolen algoritmus *Random Forest*. Nejprve byla provedena binární klasifikace na zdravé a patologické záznamy. Následně byla testována vícetřídní klasifikace patologických poruch.

Data byla rozdělena na trénovací a testovací množinu v poměru 80:20.

4 Výsledky

4.1 Binární klasifikace

Výsledky binární klasifikace zdravých a patologických hlasů jsou shrnutý v následující tabulce.



Obrázek 1: Míra vnitřní podobnosti mezi zdravými hlasovými záznamy vyjádřená pomocí korelační matice.

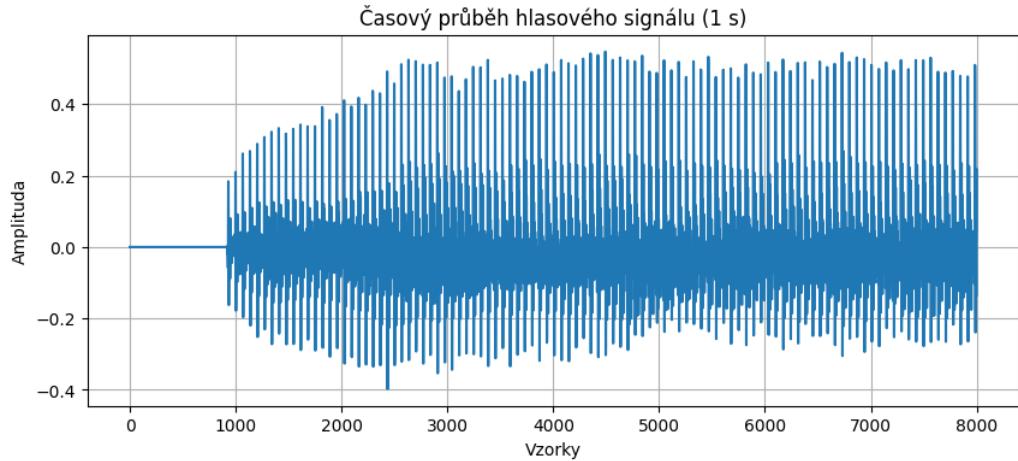
Na obrázku 1 je znázorněna korelační matice vypočtená pouze pro třídu zdravých hlasových záznamů. Výsledná hodnota reprezentuje míru vzájemné podobnosti mezi jednotlivými zdravými signály a potvrzuje relativní homogenitu této třídy.

Tento výsledek neposkytuje informaci o separovatelnosti zdravých a patologických hlasů, ale slouží jako doplnková analýza vnitřní variability dat. Pro hodnocení rozdílů mezi třídami jsou dále využity klasifikační metriky a vizualizace v prostoru příznaků.

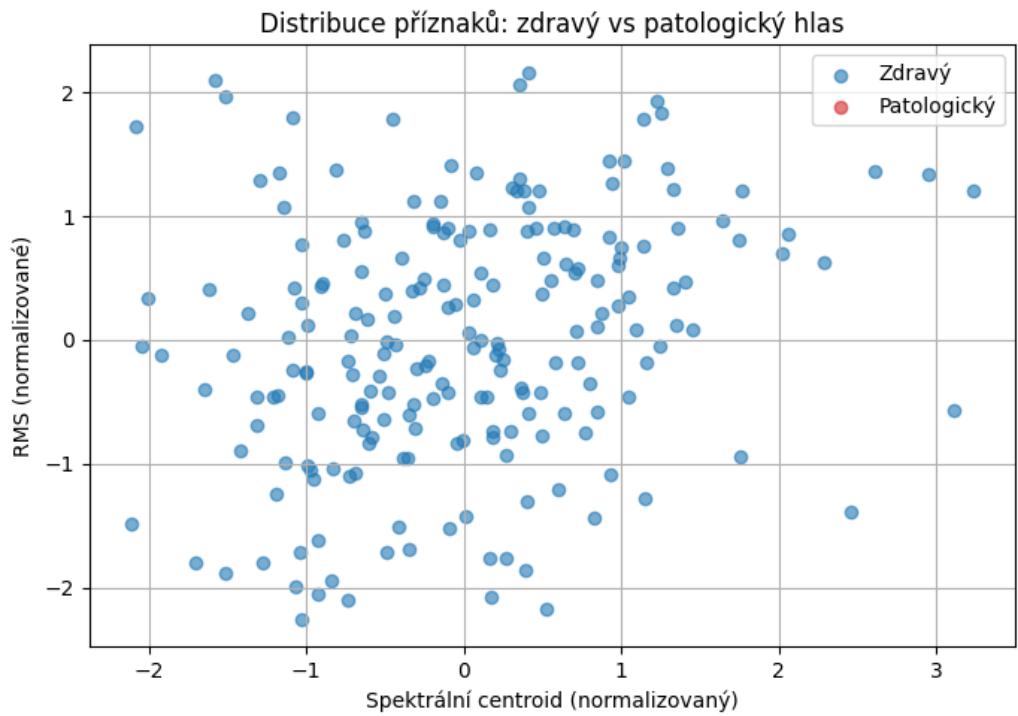
4.2 Vícetřídní klasifikace patologií

Vícetřídní klasifikace jednotlivých poruch vykazovala výrazně nižší úspěšnost. Hlavním důvodem je nerovnoměrné zastoupení tříd a nejednoznačnost anotací v databázi, což vedlo k seskupování většiny záznamů do obecné třídy patologických hlasů.

4.3 Grafické výstupy



Obrázek 2: Ukázka hlasového signálu v časové oblasti.



Obrázek 3: Distribuce příznaků hlasových záznamů v prostoru normalizovaného spektrálního centroidu a RMS hodnoty pro zdravé a patologické hlasy.

5 Závěr

Cílem práce bylo navrhnout a ověřit postup pro automatickou klasifikaci hlasových signálů na zdravé a patologické. Binární klasifikace dosáhla uspokojivých výsledků, které potvrzují, že zvolené příznaky vhodně popisují rozdíly mezi zdravým a patologickým hlasem.

Vícetřídní klasifikace jednotlivých poruch se ukázala jako problematická, především kvůli omezenému počtu vzorků v jednotlivých třídách a nejednoznačným anotacím. Pro

budoucí zlepšení by bylo vhodné využít větší databázi nebo pokročilejší metody extrakce příznaků.