

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

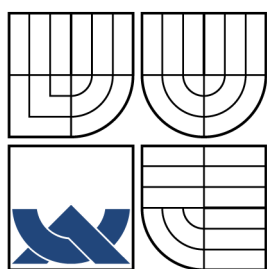
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

MODERNÍ ŘEČOVÉ PŘÍZNAKY POUŽÍVANÉ PŘI
DIAGNÓZE CHOROB

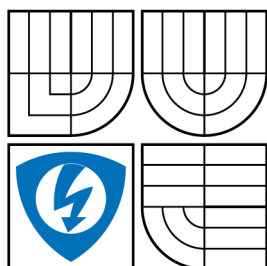
DIPLOMOVÁ PRÁCE
MASTER'S THESIS

AUTOR PRÁCE
AUTHOR

Bc. ONDŘEJ BÍLÝ



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV TELEKOMUNIKACÍ

FACULTY OF ELECTRICAL ENGINEERING AND
COMMUNICATION
DEPARTMENT OF TELECOMMUNICATIONS

MODERNÍ ŘEČOVÉ PŘÍZNAKY POUŽÍVANÉ PŘI DIAGNÓZE CHOROB

STATE OF THE ART SPEECH FEATURES USED FOR DIAGNOSTIC OF DISEASES

DIPLOMOVÁ PRÁCE
MASTER'S THESIS

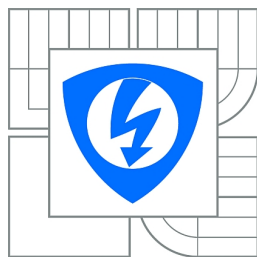
AUTOR PRÁCE
AUTHOR

Bc. ONDŘEJ BÍLÝ

VEDOUCÍ PRÁCE
SUPERVISOR

Ing. JIŘÍ MEKYSKA

BRNO 2010



**VYSOKÉ UČENÍ
TECHNICKÉ V BRNĚ**

**Fakulta elektrotechniky
a komunikačních technologií**

Ústav telekomunikací

Diplomová práce

magisterský navazující studijní obor
Telekomunikační a informační technika

Student: Bc. Ondřej Bílý

ID: 97986

Ročník: 2

Akademický rok: 2010/2011

NÁZEV TÉMATU:

Moderní řečové příznaky používané při diagnóze chorob

POKYNY PRO VYPRACOVÁNÍ:

Cílem diplomové práce je rozbor aktuální problematiky analýzy řečových signálů, především pro účely diagnózy Parkinsonovy choroby, a následný návrh řečových příznaků, pomocí nichž by bylo možné dostatečně odlišit skupinu zdravých řečníků a skupinu řečníků postižených Parkinsonovou chorobou. Mělo by se jednat o nejmodernější příznaky publikované v prestižních vědeckých časopisech. Funkce, určené k výpočtu příznaků, budou napsány v prostředí MATLAB. Výstupy této práce budou dále zpracovány v jiných závěrečných pracích věnujících se diagnóze Parkinsonovy nemoci.

DOPORUČENÁ LITERATURA:

- [1] PSUTKA, Josef, et al. Mluvíme s počítačem česky. Praha: Academia, 2006. 752 s. ISBN 80-200-1309-1.
- [2] HUANG, Xuedong; ACERO, Alex; HON, Hsiao-Wuen. Spoken Language Processing a Guide to Theory, Algorithm and System Development. Upper Saddle River: Prentice Hall PTR, 2001. 980 s. ISBN 0-13-022616-5.
- [3] F. QUATIERI, Thomas. Discrete-Time Speech Signal Processing: Principles and Practice. Upper Saddle River: Prentice Hall PTR, 2001. 816 s. ISBN 978-0132429429.

Termín zadání: 7.2.2011

Termín odevzdání: 26.5.2011

Vedoucí práce: Ing. Jiří Mekyska

prof. Ing. Kamil Vrba, CSc.

Předseda oborové rady

UPOZORNĚNÍ:

Autor diplomové práce nesmí při vytváření diplomové práce porušit autorská práva třetích osob, zejména nesmí zasahovat nedovoleným způsobem do cizích autorských práv osobnostních a musí si být plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení části druhé, hlavy VI. díl 4 Trestního zákoníku č.40/2009 Sb.

ABSTRAKT

Tato práce se zabývá diagnózou Parkinsonovy choroby na základě analýzy řečového signálu. Na začátku práce je popsána tvorba řečového signálu. Následuje popis analýzy řečového signálu, jeho předzpracování a následná extrakce příznaků. Dále je popsána Parkinsonova choroba a změna řečového signálu při tomto postižení. V následující části jsou popsány příznaky, které se používají pro diagnózu Parkinsonovy choroby (FCR, VSA, VOT atd.). Další část práce se zabývá metodami redukce a výběru příznaků pomocí učících se algoritmů (SVM, ANN, k -NN) a jejich následné ohodnocení. V poslední části diplomové práce je popsán vytvořený program pro počítání příznaků. Dále je popsán výběr příznaků a na konec jsou zhodnoceny všechny dosažené výsledky.

KLÍČOVÁ SLOVA

Řeč, prozodie, jitter, shimmer, Parkinsonova choroba, hypokinetická dysartrie, analýza řečového signálu, segmentace, parametrizace, segmentální příznak, formant, výběr příznaků, samohláska, souhláska, učící se algoritmy, matice záměn, klasifikace, přesnost, Matthewův korelační koeficient, algoritmy podpurných vektorů, umělé neuronové sítě, k nejbližších sousedů.

ABSTRACT

This work deals with the diagnosis of Parkinson's disease by analyzing the speech signal. At the beginning of this work there is described speech signal production. The following is a description of the speech signal analysis, its preparation and subsequent feature extraction. Next there is described Parkinson's disease and change of the speech signal by this disability. The following describes the symptoms, which are used for the diagnosis of Parkinson's disease (FCR, VSA, VOT, etc.). Another part of the work deals with the selection and reduction symptoms using the learning algorithms (SVM, ANN, k -NN) and their subsequent evaluation. In the last part of the thesis is described a program to count symptoms. Further is described selection and the end evaluated all the result.

KEYWORDS

Speech, prosody, jitter, shimmer, Parkinson's disease, hypokinetic dysarthria, speech signal analysis, segmentation, parameterization, segmentation symptom, formant, selection symptom, vowel, consonant, learning algorithms, confusion matrix, classification, accuracy, Matthews correlation coefficient, support vector machine, Artificial neural network, k nearest-neighbor.

BÍLÝ, Ondřej *Moderní řečové příznaky používané při diagnóze chorob*: diplomová práce. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací, 2010. 76 s. Vedoucí práce byl Ing. Jiří Mekyska

PROHLÁŠENÍ

Prohlašuji, že svou diplomovou práci na téma „Moderní řečové příznaky používané při diagnóze chorob“ jsem vypracoval samostatně pod vedením vedoucího diplomové práce a s použitím odborné literatury a dalších informačních zdrojů, které jsou všechny citovány v práci a uvedeny v seznamu literatury na konci práce.

Jako autor uvedené diplomové práce dále prohlašuji, že v souvislosti s vytvořením této diplomové práce jsem neporušil autorská práva třetích osob, zejména jsem nezasáhl nedovoleným způsobem do cizích autorských práv osobnostních a jsem si plně vědom následků porušení ustanovení § 11 a následujících autorského zákona č. 121/2000 Sb., včetně možných trestněprávních důsledků vyplývajících z ustanovení § 152 trestního zákona č. 140/1961 Sb.

Brno

.....

(podpis autora)

Poděkování

Velmi rád bych poděkoval vedoucímu diplomové práce panu Ing. Jiřímu Mekyskovi za metodickou a odbornou pomoc při jeho vytváření. Dále bych rád poděkoval celé svojí rodině za podporu po celou dobu studia a zejména svým rodičům.

V Brně:

.....
(podpis autora)

OBSAH

Úvod	11
1 Úvod do řečových signálů	12
1.1 Tvorba řeči	12
1.1.1 Dýchací ústrojí	12
1.1.2 Hlasové ústrojí	12
1.1.3 Artikulační ústrojí	14
1.2 Zobrazení řečového signálu	14
1.3 Formanty	15
1.4 Prozódie	15
2 Zpracování řečových signálů	16
2.1 Parametrizace	16
2.2 Předzpracování	16
2.2.1 Preemfáze	16
2.2.2 Ustřednění	17
2.2.3 Segmentace	17
3 Segmentální příznaky	20
3.1 LPC	20
3.2 MFCC	21
3.3 PLP	21
4 Parkinsonova choroba a příznaky k její diagnóze	23
4.1 Parkinsonova choroba	23
4.1.1 Řeč pacientů postižených Parkinsonovou chorobou	23
4.2 Suprasegmentální příznaky	24
4.3 Příznaky založené na formantech F_1, F_2	24
4.3.1 VSA	24
4.3.2 LnVSA	26
4.3.3 FCR	26
4.3.4 F_{2i}/F_{2u}	26
4.3.5 Shrnutí	27
4.4 VOT	28
4.4.1 VOT ration	28
4.4.2 VAI	28
4.5 Další příznaky používané k diagnóze Parkinsonovy choroby	28

5	Metody výběru příznaků a jejich optimalizace	30
5.1	Učící se algoritmy	30
5.1.1	Důvěryhodnost toho co bylo naučeno	30
5.1.2	Rozhodovací stromy	31
5.1.3	Metoda k nejbližších sousedů	31
5.1.4	Umělé neuronové sítě	32
5.1.5	Algoritmy podpůrných vektorů	33
5.1.6	Cross-validace	33
5.1.7	Vyhodnocení učících se algoritmů	34
5.2	Evoluční algoritmy	36
5.2.1	Genetické algoritmy	36
6	Systém pro diagnózu Parkinsonovy choroby	38
6.1	Nahrávky	38
6.2	Výpočet příznaků	39
6.2.1	Grafické prostředí	40
6.2.2	Extrakce požadované části signálu	41
6.2.3	Předzpracování signálu	41
6.2.4	Post-processing matice příznaků	41
6.3	Testování	41
6.3.1	Databáze nahrávek	42
6.3.2	Výběr příznaků	42
6.3.3	Testování algoritmů	44
6.3.4	Zhodnocení výsledků jednotlivých měření	53
6.4	Vyhodnocení příznaků	55
6.4.1	Realizace ohodnocení příznaků v RapidMineru	57
7	Závěr	58
	Literatura	60
	Seznam symbolů, veličin a zkratk	63
	Seznam příloh	66
A	Obsah přiloženého DVD	67
A.1	obsah adresáře DiplomovaPrace	67
A.1.1	obsah adresáře Matlab	67
A.1.2	obsah adresáře RapidMiner	67
B	Přehled cvičení	68

C	Příznaky	72
C.1	Všechny vypočítané příznaky	72
C.2	Příznaky použité pro klasifikaci	75

SEZNAM OBRÁZKŮ

1.1	Hlasový trakt člověka	13
1.2	Hrtan a hlasivky	14
2.1	Řečový signál před preemfází.	16
2.2	Řečový signál po preemfází.	17
2.3	Segmentace řečového signálu o délce N s nepřekrývajícími se segmenty.	18
2.4	Segmentace řečového signálu o délce N s překrývajícími se segmenty.	18
4.1	Grafické znázornění příznaku VSA.	25
5.1	Rozhodovací strom pro výběr dalšího studia	31
5.2	Klasifikace pomocí k -NN	32
5.3	Model neuronu	33
5.4	Oddělení dvou tříd pomocí SVM	34
6.1	Blokové schéma systému pro diagnózu Parkinsonovy choroby	38
6.2	Část vygenerované tabulky příznaků	40
6.3	Grafické prostředí pro výpočet příznaků	40
6.4	Blokové schéma klasifikátoru	42
6.5	Příznaky vhodné pro klasifikaci	43
6.6	Příznaky nevhodné pro klasifikaci	43
6.7	Rozdělení nahrávek pomocí příznaků VSA, lnVSA, FCR, VAI a F_{2u}/F_{2i}	45
6.8	Přesnost a MCC k -NN klasifikátoru	46
6.9	Přesnost a MCC SVM klasifikátoru	49
6.10	Přesnost a MCC ANN klasifikátoru	50
6.11	Průměrná přesnost a MCC všech klasifikátorů.	53
6.12	Přesnost klasifikátorů pro testovací množinu nahrávek	54
6.13	MCC klasifikátorů pro testovací množinu nahrávek	55
6.14	Přesnost klasifikátorů pro validační množinu nahrávek	56
6.15	MCC klasifikátorů pro validační množinu nahrávek	56

SEZNAM TABULEK

4.1	Srovnání příznaků VSA a FCR na pohlaví a věku řečníků. Uvedené hodnoty platí pro anglický jazyk [17].	26
4.2	Přehled formantů a jejich kmitočtů pro samohlásky /i/, /u/ a /a/ (PD – L – léčený pacient, PD – NL – neléčený pacient, ZJ – zdravý člověk) [17].	27
4.3	Parametry jednotlivých příznaků VSA, LnVSA, FCR, F_{2i}/F_{2u} (PD – L – léčený pacient, PD – NL – neléčený pacient, ZJ – zdravý člověk) [17].	27
5.1	Matice záměn	35
6.1	Počet nahrávek od fakultní nemocnice sv. Anny v Brně	39
6.2	Počet nahrávek celkem	39
6.3	Příklad všech příznaků pro klasifikaci	42
6.4	Matice záměn pro výběr příznaků	43
6.5	Matice záměn k -NN klasifikátoru (všechny příznaky)	46
6.6	Matice záměn k -NN klasifikátoru (vybrané příznaky)	47
6.7	Matice záměn k -NN klasifikátoru (Parkinsonovy příznaky)	47
6.8	Vyhodnocení kvality klasifikace k -NN klasifikátoru	48
6.9	Matice záměn SVM klasifikátoru (všechny příznaky)	48
6.10	Matice záměn SVM klasifikátoru (vybrané příznaky)	49
6.11	Matice záměn SVM klasifikátoru (Parkinsonovy příznaky)	49
6.12	Vyhodnocení kvality klasifikace SVM klasifikátoru	49
6.13	Matice záměn ANN klasifikátoru (všechny příznaky)	51
6.14	Matice záměn ANN klasifikátoru (vybrané příznaky)	51
6.15	Matice záměn ANN klasifikátoru (Parkinsonovy příznaky)	52
6.16	Vyhodnocení kvality klasifikace ANN klasifikátoru	52
6.17	Matice záměn pro ohodnocení příznaky	57
6.18	5 nejlepších a 5 nejhorších příznaků	57

ÚVOD

V dnešním moderním světě se každý člověk snaží ušetřit při své činnosti co možná nejvíce času a prostředků. Právě tyto dvě věci jsou hlavními hnacími motory pro vědecké týmy pracující v lékařském odvětví. Ti se snaží vyvíjet stále efektivnější způsoby vyšetřování pacientů na základě různých alternativních metod. Jedna z těchto metod je například rozpoznání nemoci na základě specifických příznaků v pacientově řeči.

Řeč je nejdůležitějším prostředkem k předávání informací mezi lidmi. Samotná tvorba řeči je jeden z nejsložitějších procesů, které lidské tělo produkuje. Nevýhoda všech složitých procesů je taková, že jsou ve většině případů složeny z celé řady dílčích prvků, které jsou na sobě závislé. Pak stačí, aby jen jeden z těchto prvků nepracoval správně a tím přestane pracovat správně i celý proces. Podobně tomu je i tvorby řeči, kde jednotlivé prvky představují samostatné orgány jako například mozek, plíce, hlasivky a jiné. Když je pak postižen jen jeden z těchto orgánů, tak je to na kvalitě řeči znatelné.

Důsledek Parkinsonovy choroby je postižení pohybového aparátu, který se mimo jiné podílí i u na tvorbě řeči. Udává se, že asi 90 % jedinců postižených Parkinsonovou chorobou trpí některou z poruch řeči. Nejčastěji to bývá porucha prozodie [22]. Z tohoto důvodu se zdá diagnóza Parkinsonovy choroby pomocí analýzy pacientovi řeči jako dobrá cesta, která by mohla ušetřit jak čas, tak i prostředky.

K diagnóze Parkinsonovy choroby se používají jak běžné příznaky pro zpracování řeči, tak i sofistikovanější příznaky k tomu vytvořené. Tyto příznaky jsou většinou počítány z několika běžných příznaků.

V této práci je použito přes 70 řečových příznaků, které jsou počítány nejen pro samohlásky, ale i pro celé slabiky, slova a věty. Manuálně kombinovat jednotlivé příznaky s celým poskytnutým inventářem by bylo časově nemožné. Ale v současnosti existují strojové algoritmy, které jsou k řešení takových problémů určeny. Nazývají se evoluční algoritmy, toto označení platí pro všechny matematické postupy, které jsou založeny na evolučním procesu v přírodě [2]. V práci je těchto algoritmů použito pro diagnózu Parkinsonovy choroby.

1 ÚVOD DO ŘEČOVÝCH SIGNÁLŮ

Řeč je nejjednodušší, ale také nejefektivnější prostředek sloužící k předávání informací mezi lidmi. Informace vytvořené pomocí řeči jsou přenášeny pomocí akustického signálu. Tímto způsobem vytvářené informace nenesou pouze samotnou myšlenku mluvčího, ale i specifické informace o jeho hlasovém traktu (intonace, barva hlasu atd.), o různých anomáliích (koktavost, vada výslovnosti, atd.), případně o jeho emocionálním stavu (radost, smutek, nervozita, atd.). K tvorbě řeči slouží Brocovo řečové centrum, které se nachází na levé straně čelního laloku lidského mozku. Z tohoto centra jsou posílány řídicí instrukce pro dané orgány [13].

1.1 Tvorba řeči

Samotná tvorba řeči je jeden z nejkomplicovanějších procesů lidského těla vůbec. Celý tento proces je koordinován mozkiem, který vytváří instrukce pro části těla podílejících se na tvorbě řeči. Tyto části těla se souhrnně nazývají hlasový trakt (obr. 1.1). Hlasový trakt je tvořen ze tří částí: dýchací, hlasové a artikulační ústrojí [13].

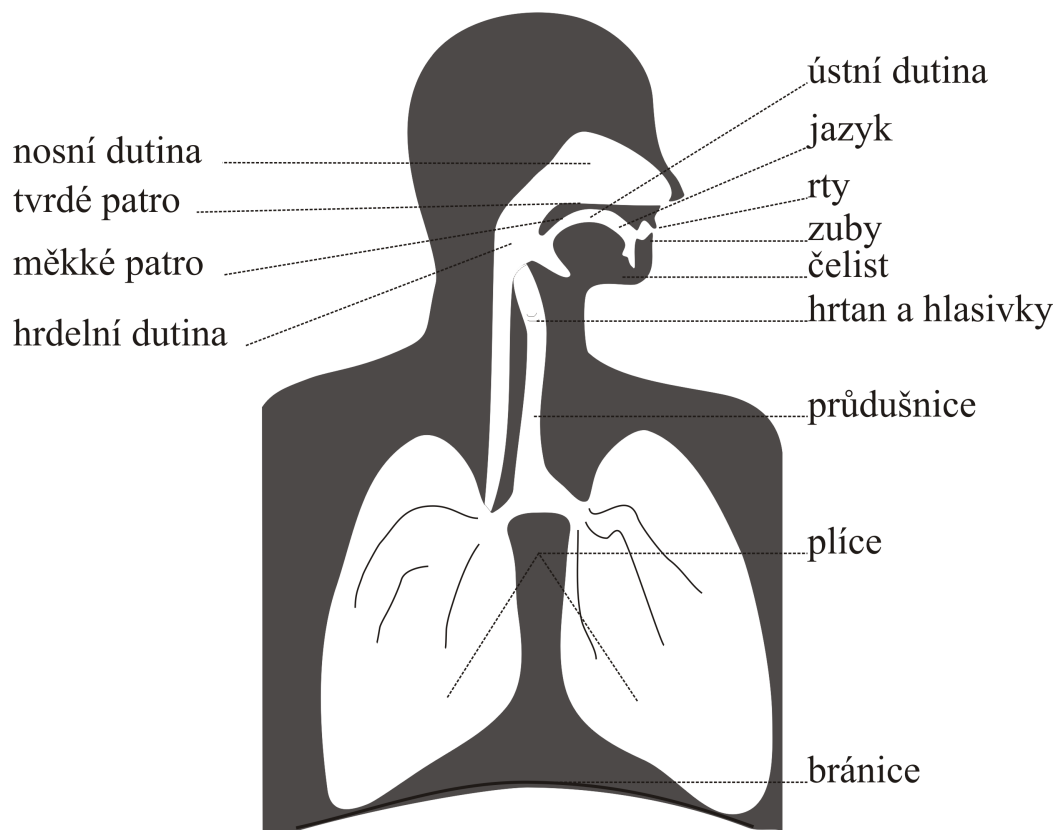
1.1.1 Dýchací ústrojí

Dýchací ústrojí je primárně určeno k zajištění základních životních funkcí. Při řeči se dýchání výrazně mění. U běžného dýchání je poměr vdechu a výdechu 2:3 a objem nadechnutého vzduchu je asi 0,5l. Při řeči se výrazně prodlužuje doba výdechu a to v poměru 1:7–1:12, dále pak výrazně narůstá objem nadechnutého vzduchu a to až na 1,5l. Při běžném dýchání člověk většinou dýchá jen nosem, ale když mluví, tak se nadechuje jak nosem tak i pusou. To proto, že je nutné v krátkém časovém úseku nabrat velký objem vzduchu do plic [11].

Člověk v drtivé většině případů mluví při vydechování vzduchu (vzduch dodává řeči potřebnou energii). Vydechovaný vzduch je přiváděn průdušnicí přes hrtan až k nadhrtanovým dutinám, kde se vytváří řečový signál. Existují i výjimky, kdy se řeč vytváří při vdechování vzduchu. V českém jazyce to jsou zejména citoslovečné výrazy např. „au“ [11].

1.1.2 Hlasové ústrojí

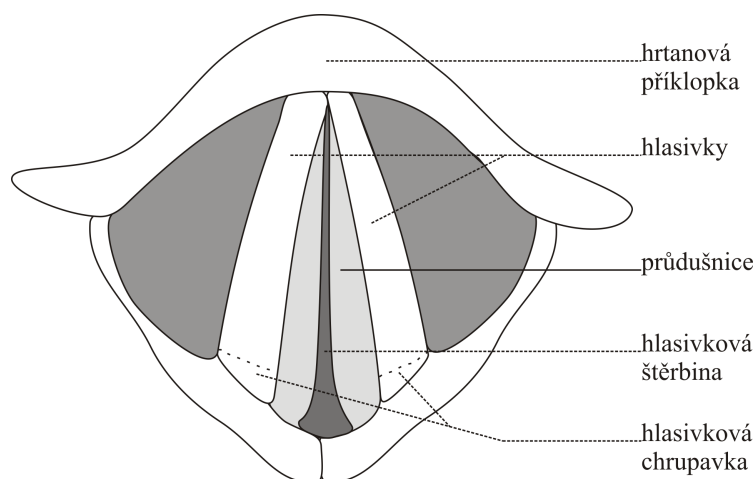
Hlasové ústrojí je uloženo v hrtanu. Je spojeno pomocí průdušnice k plicím. Základním stavebním kamenem hlasového ústrojí jsou hlasivky (obr. 1.2). Hlasivky jsou umístěny přímo za ohryzkem. Jsou tvořeny dvěma tenkými vazy, které se otvírají



Obr. 1.1: Hlasový trakt člověka

a zavírají při proudění vzduchu z plic. Jsou spojeny s dvěma chrupavkami, štítnou a hlasivkovou. Mezi hlasivkami se vytváří prostor, kterému se říká hlasivková štěrbina. Má trojúhelníkový tvar. Jestliže člověk nevydává žádný hlas, tak hlasivkovou štěrbinou proudí vzduch a nic významného se neděje. Při vytváření hlasu hlasivky cestu vzduchu uzavřou. Uzavřené hlasivky se pod tlakem vzduchu rozkmitají a prudce se začnou otvírat a zavírat. Takto upravený vzduch (uzavíráním a otvíráním hlasivek vznikne z homogenního, nehomogenní proudění vzduchu) tvoří základ lidského hlasu, také označován jako základní tón (F_0) [11, 13].

O tom, jaký bude mít jedinec hlas, rozhoduje zejména velikost hlasivek (u žen v průměru 18 mm, u mužů v průměru 22 mm). Čím jsou hlasivky kratší, tím více kmitají a tím vzniká vyšší základní tón. Frekvence základního tónu se pohybuje v rozmezí od 50 Hz (hluboký hlas) až po 480 Hz (vysoký hlas). Frekvence základního tónu není konstantní, je závislá na zdravotním a emocionálním stavu řečníka. Kolísání základní periody se nazývá jitter. Kolísání amplitudy základního tónu se nazývá shimmer [13].



Obr. 1.2: Hrtan a hlasivky

1.1.3 Artikulační ústrojí

Artikulační ústrojí je poslední část hlasového traktu, která se podílí na tvorbě řeči. Nachází se v nadhrtanových dutinách a dělí se, na aktivní (pohybují se) a pasivní orgány (nepohybují se). Mezi aktivní artikulační orgány patří jazyk, dolní patro, rty, měkké patro a hrtan. Mezi pasivní artikulační orgány patří rty, zuby, dásně, střední patro a zadní patro [11].

Jazyk je nejvýznamnějším artikulačním orgánem. Skládá se ze tří částí: hrotu, hřbetu a kořene. Je jím možné pohybovat jako jedním celkem (při vokálech), nebo jen jeho jednotlivými částmi. Dokáže velice rychle měnit svůj tvar a svoje umístění v dutině ústní [11].

Z celého hlasového traktu právě artikulační ústrojí nejvíce ovlivňuje výsledný zvuk. Výsledný zvuk je ovlivněn v zásadě dvěma faktory: [13]

1. Vytvoření tónové struktury. Základní tón (F_0), který vzniká v hrtanovém ústrojí udává výšku tónu. V důsledku rezonancí v nadhrtanové dutině vznikají vyšší kmitočty ($F_1, F_2 \dots F_n$) také označované jako formanty. Složením vyšších frekvencí vznikají především samohlásky [13].
2. Vytvoření šumové složky. Tyto šumové složky vznikají, když se tříští proud vzduchu vycházející z plic o překážky, které mu brání v průchodu. Tyto překážky jsou vytvořeny artikulačními orgány. Šumovými složkami vznikají souhlásky [13].

1.2 Zobrazení řečového signálu

Řečový signál lze zobrazit v několika rovinách: v kmitočtové, časové, nebo kombinací předchozích. Touto kombinací vznikne spektrogram, kde na horizontální ose

je vynesena čas a na vertikální ose je vynesena kmitočet.

1.3 Formanty

Formanty jsou velice důležitou součástí řečového signálu. První tři formanty nesou informace o charakteru samohlásek a znělých souhlásek. Formanty jsou nejlépe vidět ve spektrální obálce řečového signálu. Pro nalezení formantů v kmitočtové rovině se nejčastěji používá metoda LPC. Snaha najít kmitočet formantů nese v zásadě dva problémy: [13]

1. Výskyt nepravých vrcholů ve spektrální obálce. Maxima ve spektrální obálce jsou téměř výhradně způsobena formanty. Ale může nastat výjimka, kdy je maximum způsobeno předimenzováním řádu prediktoru u metody LPC [13].
2. Splývání formantů může nastat, když dva, či více formantů mají sobě blízké kmitočty [13].

1.4 Prozódie

Prozódie, nebo-li suprasegmentální jevy, jsou definovány jako souhrn intonace, intenzity a tempa řeči. Při špatných akustických podmínkách mohou být tyto vlastnosti jediné, které posluchač uslyší [21].

Intonace je nejvíce rozpoznatelná prozodická vlastnost. Značí změnu kmitočtu základního tónu (F_0). Z lingvistického hlediska to znamená změnu výšky hlasu řečníka. Tento jev je nejvíce vnímán při mluvení k odlišení tázací od oznamovací věty, ukončení větného celku, nebo při proslovu [21].

Změna hlasitosti řeči je označována jako intenzita. Nejčastěji vyjadřuje emocionální stav řečníka (nervózní člověk má kolísavou hlasitost projevu, rozložený člověk má vyšší hlasitost projevu apod.) [21].

Tempo je další významná prozodická vlastnost. Vyjadřuje rychlost řeči, tzn. počet slov, které řečník vysloví za určitý časový úsek. Tempo řeči je závislé na mnoha okolnostech: stylu mluvení, emocionálním stavu, umísťování pauz, způsobu artikulace, nebo na rytmu řeči [21].

2 ZPRACOVÁNÍ ŘEČOVÝCH SIGNÁLŮ

2.1 Parametrizace

Parametrizací řečového signálu rozumíme vyjádření tohoto signálu s omezeným počtem hodnot. Používají se dvě metody: [4]

1. **Neparametrická** metoda vycházející pouze z poznatků zpracování běžných signálů.
2. **Parametrická** metoda vycházející z poznatků tvorby řeči.

Hodnoty vypočtené těmito metodami nazýváme příznaky. Příznaky se dělí na:

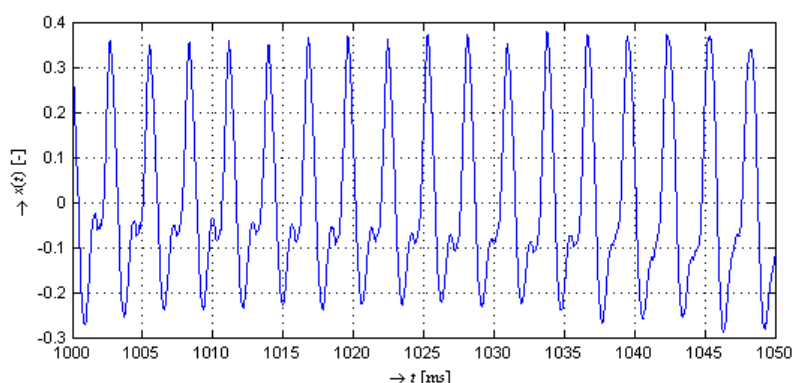
1. **Skalární** – jen jedno číslo na řečový úsek.
2. **Vektorové** – více čísel řazených do vektorů, případně do matic [4].

2.2 Předzpracování

Před parametrizací řečového signálu je vhodné signál upravit do vhodné podoby pro výpočet příznaků [4].

2.2.1 Preemfáze

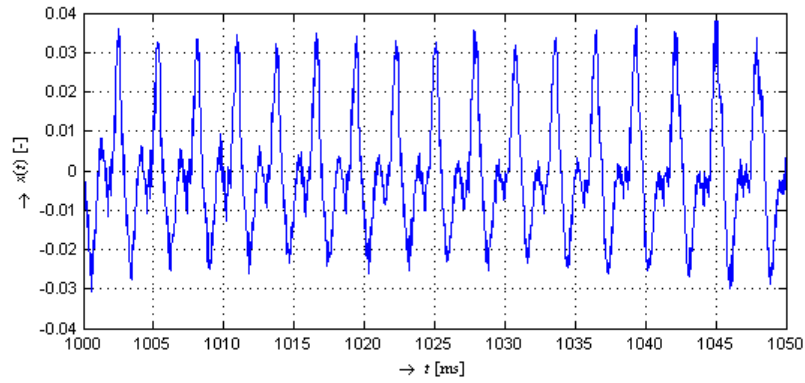
Energie řečového signálu klesá k vyšším kmitočtům. Preemfáze se stará o vyrovnávání kmitočtové charakteristiky. Na obr. 2.1 lze vidět řečový signál před průchodem filtrem. Na obr. 2.2 je vidět signál po průchodu filtrem, tento signál je zřetelně kostrbatější a má více ostrých hran, to dokládá větší podíl vyšších kmitočtů [4].



Obr. 2.1: Řečový signál před preemfází.

Ke zvýraznění vyšších kmitočtů lze použít filtr prvního řádu s přenosovou funkcí: [4]

$$H(z) = 1 - kz^{-1}, \quad (2.1)$$



Obr. 2.2: Řečový signál po preemfází.

kde $k \in [0.9, 1]$. Prakticky tento filtr odečítá od sebe dva sousední vzorky: [4]

$$s'[n] = s[n] - ks[n-1]. \quad (2.2)$$

2.2.2 Ustřednění

Ustřednění slouží k odstranění stejnosměrné složky, která nenese žádné užitečné informace a mohla by být při dalším zpracování signálu rušivá. Stejnosměrná složka se odstraní odečtením střední hodnoty signálu: [4]

$$s'[n] = s[n] - \mu_s, \quad (2.3)$$

Střední hodnotu lze počítat dvěma způsoby:

1. **Off-line** výpočet – po ukončení signálu se signál zprůměruje: [4]

$$\mu_s = \frac{1}{N} \sum_{n=0}^{N-1} s[n]. \quad (2.4)$$

2. **On-line** výpočet – celý signál není k dispozici, pak se střední hodnota vypočítá rekurzivně:

$$\mu_s = \gamma\mu_s[n-1] + (1-\gamma)s[n], \quad (2.5)$$

kde $\gamma \rightarrow 1$.

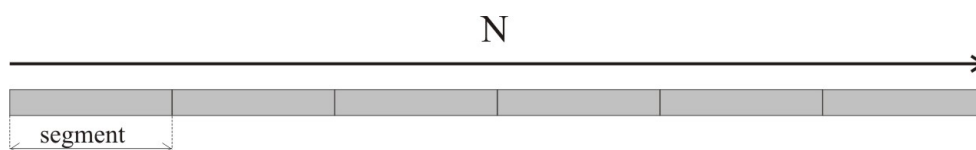
2.2.3 Segmentace

Řečový signál není náhodný stacionární proces¹. To znamená, že by se jeho vlastnosti získávali velice obtížně. Signál se proto rozdělí na kratší úseky (segmenty), které lze považovat za stacionární a lze z nich odhadnout jejich vlastnosti [4].

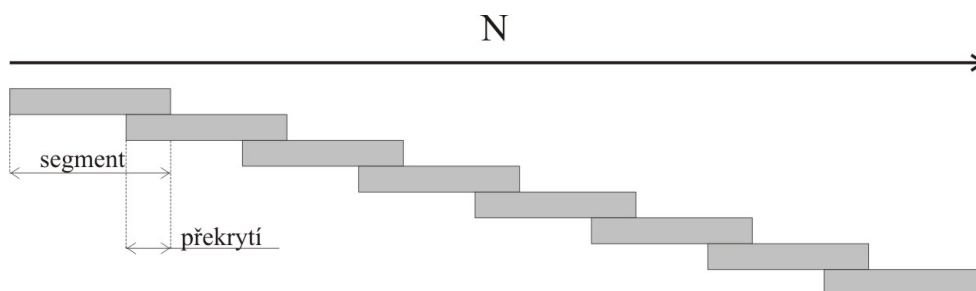
¹Stacionární proces má parametry, které jsou neproměnné v čase.

Délka segmentů by měla vyhovovat několika požadavkům. Měla by být co nejkratší, aby je bylo možné požadovat za stacionární, ale na druhou stranu musí být co nejdelší, aby z nich bylo možné co nejlépe odhadnout požadované vlastnosti. Proto je zde kompromis, který je založen na setrvačnosti hlasového traktu, typická délka segmentu bývá 20–25 ms [4].

Jednotlivé segmenty se mohou a nemusí překrývat. Segmenty, které se nepřekrývají (obr. 2.3), zajistí rychlý časový posun, malé nároky na procesor a paměť, ale zároveň hodnoty parametrů se můžou segment od segmentu velice lišit. U segmentů, které se překrývají (obr. 2.4), je časový posun pomalý, jsou zde kladeny velké nároky na procesor i paměť, ale parametry segmentů jsou vyhlazené [4].



Obr. 2.3: Segmentace řečového signálu o délce N s nepřekrývajícími se segmenty.



Obr. 2.4: Segmentace řečového signálu o délce N s překrývajícími se segmenty.

Okénkové funkce

Pro tzv. „vykrojení“ signálu se nejčastěji používají dva druhy okénkových funkcí: [4]

1. **Pravoúhlé okno** – signál zůstává nezměněn:

$$w[n] = \begin{cases} 1 & \text{pro } 0 \leq n \leq L - 1, \\ 0 & \text{jinde.} \end{cases} \quad (2.6)$$

2. **Hammingovo okno** – signál je na okrajích utlumen:

$$w[n] = \begin{cases} 0,54 - 0,46 \cos\left(\frac{2\pi n}{L-1}\right) & \text{pro } 0 \leq n \leq L - 1, \\ 0 & \text{jinde,} \end{cases} \quad (2.7)$$

kde L značí délku segmentu [4].

Při segmentaci pomocí okna, dochází ke změně původního spektra signálu. Násobením okna v časové oblasti odpovídá konvoluce spekter: [4]

$$X(f) = S(f) * W(f). \quad (2.8)$$

3 SEGMENTÁLNÍ PŘÍZNAKY

Jedná se o příznaky, ze kterých lze určit parametry řečového signálu. Z těchto parametrů lze vyvodit vlastnosti hlasu [21].

3.1 LPC

Na základě získaných koeficientů a_i u lineární predikce LPC (Linear Predictive Coding) lze modelovat hlasový trakt. Řečový signál po odstranění impulzní odezvy hlasového traktu odpovídá hlasivkovému buzení. Výhodou lineární predikce je jednoduchý výpočet LPC koeficientů pomocí soustavy lineárních rovnic. Model hlasového traktu odpovídá IIR filtru: [21, 4]

1. Přenosová funkce IIR filtru:

$$H(z) = \frac{1}{A(z)}, \quad (3.1)$$

kde polynom $A(z) = 1 + a_1 z^{-1} + a_2 z^{-2} + \dots + a_P z^{-P}$ má řád P .

2. Predikce signálu:

$$\tilde{s}[n] = - \sum_{i=1}^P a_i s[n-i]. \quad (3.2)$$

3. Chyba predikce:

$$e[n] = s[n] - \tilde{s}[n]. \quad (3.3)$$

4. Výpočet LPC koeficientů pomocí soustavy lineárních rovnic:

$$\Gamma_P \cdot a_p = \gamma_p, \quad (3.4)$$

$$\Gamma_P = \begin{bmatrix} \gamma_{xx}[0] & \gamma_{xx}^*[-1] & \gamma_{xx}^*[-2] & \cdots & \gamma_{xx}^*[-p+1] \\ \gamma_{xx}[1] & \gamma_{xx}[0] & \gamma_{xx}^*[-1] & \cdots & \gamma_{xx}^*[-p+2] \\ \gamma_{xx}[2] & \gamma_{xx}[1] & \gamma_{xx}[0] & \cdots & \gamma_{xx}^*[-p+3] \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{xx}[p-1] & \gamma_{xx}[p-2] & \gamma_{xx}[p-3] & \cdots & \gamma_{xx}[0] \end{bmatrix}, \quad (3.5)$$

$$a_p = [a_p[1], a_p[2], a_p[3], \dots, a_p[p]]^T, \quad (3.6)$$

$$\gamma_p = [-\gamma_{xx}[1], -\gamma_{xx}[2], -\gamma_{xx}[3], \dots, -\gamma_{xx}[p]]^T, \quad (3.7)$$

kde $\gamma_{xx}[n]$ značí autokorelační koeficienty a Γ_P vytváří Toeplitzovu matici [21].

3.2 MFCC

Koeficienty MFCC (Mel-frequency Cepstral Coefficients), byly prvními koeficienty, které brali v potaz nelineární a maskovací vlastnosti lidského sluchu. Nelinearita sluchu v kmitočtové oblasti se pak odstraní převodem kmitočtové osy na mel osu, která je pro sluch lineární¹. Převod kmitočtové stupnice v Hz na kmitočtovou osu v melech lze vyjádřit vztahem: [21]

$$f_m = 2595,0375 \log \left(1 + \frac{f}{700} \right) \quad [\text{mel}], \quad (3.8)$$

pro zpětný převod se používá vztah:

$$f = 700(10^{\frac{f_m}{2595,0375}}) \quad [\text{Hz}]. \quad (3.9)$$

Koeficienty se odvodí v několika krocích:

1. Řečový signál je segmentován pomocí vhodného typu okna a následně je podroben diskrétní Fourierově transformaci za pomoci rychlé Fourierovi transformace [21].
2. Nelineární a maskovací vlastnosti sluchu jsou vyjádřeny bankou číslicových filtrů s trojúhelníkovou modulovou kmitočtovou charakteristikou. Nejdříve se převede kmitočtové pásmo v Hz na pásmo v jednotkách mel pomocí vztahu 3.8. Pak se do tohoto pásma rovnoměrně rozmístí trojúhelníkové kmitočtové filtry a převede ze zpět na pásmo v Hz podle vztahu 3.9. Tím je docíleno, že kmitočtové pásmo v Hz je rozloženo podle nelineárních vlastností sluchu [21].
3. Určí se výkonové spektrum řečového signálu. Toto spektrum je následně násobeno filtry z banky [21].
4. Vlastní výpočet koeficientů MFCC se provede zpětnou diskrétní Fourierovou transformací ze spektra násobeného řečového signálu filtry z banky. Pro tyto účely se používá diskrétní kosinova transformace [21].

3.3 PLP

Výpočet percepčních lineárních koeficientů vychází z lineární predikční analýzy. Výpočet vlastních koeficientů předchází předzpracování řečového signálu. Nejdříve se zpracují maskovací vlivy, křivky stejné hlasitosti a nelineární vlastnosti sluchu. A až po této proceduře proběhne výpočet PLP koeficientů. Ten je rozdělen do několika kroků: [21]

¹Mel stupnice vyjadřuje dva zvuky, které jsou od sebe vzdáleny stejný počet melů, jsou stejně od sebe vzdáleny i percepčně.

1. Řečový signál je zpracován stejným způsobem jak u MFCC (segmentace, diskrétní Fourierova transformace). Pak je stanoven odhad krátkodobé výkonové spektrální hustoty segmentu [21].
2. V tomto kroku se do modelu přiřadí maskovací vlastnosti sluchu a provede se transformace kmitočtové osy z $f[\text{Hz}]$ na $f_b[\text{bark}]$, podle vztahu: [21]

$$f_b = 13 \arctg(7,6 \cdot 10^{-4} f) + 3,5 \arctg \left[\left(\frac{f}{7,5 \cdot 10^3} \right)^2 \right] \quad [\text{bark}]. \quad (3.10)$$

Následně se podobně jako u MFCC vytvoří banka filtrů, která je nyní lichoběžníková a lineárně se rozloží v barkově pásmu. Následuje převod zpět do pásma v Hz, kde vznikne nelineární rozložení, které odpovídá vlastnostem sluchu.

3. Před násobením krátkodobé spektrální hustoty bankou filtrů se upraví moduly jednotlivých barkových filtrů podle křivek hlasitosti [21].
4. Použije se zpětná Fourierova transformace pro získání časových vzorků. Takto získané vzorky, ale už obsahují vlastnosti lidského sluchu [21].
5. Nakonec se provede vlastní výpočet PLP koeficientů, které vzniknou řešením soustavy rovnic 3.5 [21].

4 PARKINSONOVA CHOROBA A PŘÍZNAKY K JEJÍ DIAGNÓZE

4.1 Parkinsonova choroba

Parkinsonova choroba je neurologické degenerativní onemocnění, které postihuje centrální nervovou soustavu. Byla popsána v roce 1817 anglickým lékařem Jamesem Parkinsonem. V populaci je výskyt velice častý u lidí na 50 let, ve výjimečných případech postihuje i mladší lidi. V současné době se jedná o nevyléčitelnou chorobu, kterou lze současnými léky jen zpomalit [22].

Příčina této nemoci je způsobena ztrátou nervových buněk, které produkují sloučeninu zvanou dopamin. Tato sloučenina slouží ke komunikaci buněk v části mozku zvané bazální ganglia [22].

Nemoc se projevuje zhoršenou ovladatelností celého pohybového aparátu. Pacient trpí rigiditou (rezistencí svalů vůči pohybu), akinézií (neschopnost uvést svaly do pohybu), bradykinézií (pomalé pohyby svalstva), třesem a postižením řečového aparátu [22].

4.1.1 Řeč pacientů postižených Parkinsonovou chorobou

70 % pacientů postižených Parkinsonovou chorobou trpí hypokinetickou dysartrií. Ta způsobuje problém s fonací, intonací i artikulací. Hypokinetická dysartrie postihuje všechny tři části hlasového traktu: [22]

1. V důsledku ochabování svalstva dýchacího ústrojí se zmenšuje objem vzduchu, který je člověk schopen vdechnout. To má za následek zhoršení kvality hlasu, zkrácení délky promluvy a zmenšenou sílu hlasu. Takto postižení pacienti pracují s dechem neefektivně. Mluví v krátkých větách, zrychlují tempo řeči, obtížně koordinují mluvení s dýcháním [22].
2. Abnormálně zvýšená aktivita hrtanových svalů uložených v hlasovém ústrojí způsobuje, že síla hlasu mnohem rychleji klesá a výška hlasu nabývá menšího rozsahu. Posluchač tyto vady vnímá jako monotónní řeč, kde je přítomna hypofonie (snížená síla hlasu, hlas je tichý, nevýrazný) a dysfonie (hlas je drsný, zastřený, chraptivý). Někdy může být hlas vnímán jako labilní, napjatý a neznelý [22].
3. Artikulační ústrojí, na kterém je nejvíce závislá konečná kvalita hlasu, rozlišuje tři poruchy: [22]
 - V důsledku snížení rychlosti a rozsahu pohybu artikulačních orgánů je narušena schopnost výslovnosti souhlásek. Z logopedického hlediska se

rozlišují souhlásky na ražené–závěrové (/p/, /b/, /m/, /t/, /d/, /n/, /t̥/, /d̥/, /ŋ/, /k/, /g/), poloražené–polozávěrové (/c/, /č/) a třené–užínové (/f/, /v/, /s/, /š/, /z/, /ž/, /l/, /r/, /ř/, /j/, /h/, /ch/). Hlávky ražené a poloražené jsou běžně tvořeny nekontinuálně (tzn. třístěním vydechovaného proudu vzduchu o překážky), u pacientů jsou tvořeny kontinuálně. Třené hlávky jsou běžně tvořeny napjatě, u pacientů jsou tvořeny nenapjatě [22].

- U poruch prozodie jsou postiženy všechny její vlastnosti. Poruchy intonace se projevuje monotónností projevu. Narušení tempa způsobuje zpomalení, nebo zrychlení projevu. Výrazně je ovlivněna poloha pauz ve větách [22].
- V plynulosti řeči se rozlišují dvě formy poruch. Mohou existovat jak samostatně, tak i současně (palilalie a hezitace). Palilalie způsobuje zrychlené opakování slabik nebo opakování celých slov. Hezitace má za následek neplánované pauzy v projevu [22].

Příznaky rozebrané v kapitole 3 (LPC, MFCC, PLP) nejsou pro diagnózu Parkinsonovy choroby příliš vhodné. Pro tyto účely se používají příznaky, které jsou založeny na prozodických vlastnostech a kvalitě řečového signálu.

4.2 Suprasegmentální příznaky

Suprasegmentální příznaky byly podrobně rozebrány v kapitole 1.4.

Významnou roli v odhalení choroby hraje intonace, která vychází z kmitočtu základního tónu F_0 . Střední hodnota F_0 (při vyslovování samohlásek) je u pacientů významně vyšší než u zdravých lidí ve stejném věku. Dále pak narůstá hodnota jitteru i shimmeru. Tyto jevy se dějí v důsledku ochabnutí nadhrtanových svalů, které pak nedokáží udržet na dostatečně dlouhou dobu proud vzduchu potřebný k rozkmitání hlasivek [7, 19].

4.3 Příznaky založené na formantech F_1, F_2

4.3.1 VSA

Příznak VSA (Vowel Space Area) je používán jako akustická míra dysartrie v řeči. Princip je založen na výpočtu plochy, kterou vytvoří formanty F_1, F_2 samohlásek /a/, /e/, /i/, /o/, /u/. Obsah plochy vypočtené pomocí VSA je vidět na obr. 4.1. Tato plocha se rozdělí na trojúhelníky a výpočtem se dílí plochy, které se pak sečtou. Výpočet trojúhelníkové plochy pro samohlávky /a/, /i/ a /u/ pak probíhá

podle vztahu: [17]

$$VSA = \left| \left(\frac{(F_{1i} \cdot (F_{2a} - F_{2u}) + F_{1a} \cdot (F_{2u} - F_{1u}) + F_{1u} \cdot (F_{2i} - F_{2a}))}{2} \right) \right|, \quad (4.1)$$

nebo podle vztahu:

$$VSA = \sqrt{(S \cdot (S - ED_{iu}) \cdot (S - ED_{ia}) \cdot (S - ED_{au})), \quad (4.2)$$

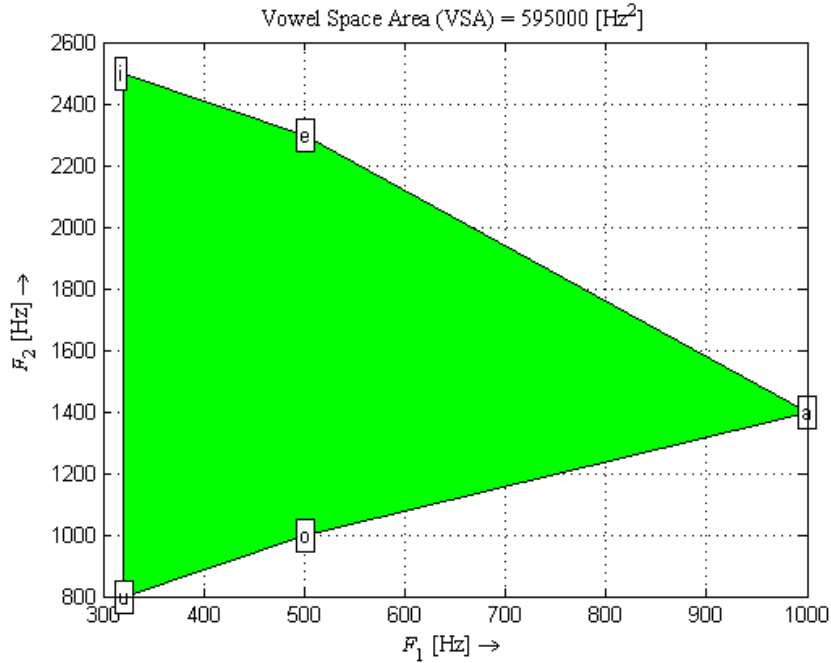
$$ED_{iu} = \sqrt{(F_{1i} - F_{1u})^2 + (F_{2i} - F_{2u})^2}, \quad (4.3)$$

$$ED_{ia} = \sqrt{(F_{1i} - F_{1a})^2 + (F_{2i} - F_{2a})^2}, \quad (4.4)$$

$$ED_{au} = \sqrt{(F_{1a} - F_{1u})^2 + (F_{2a} - F_{2u})^2}, \quad (4.5)$$

$$S = \frac{(ED_{iu} + ED_{ia} + ED_{au})}{2}, \quad (4.6)$$

kde jednotlivé indexy u proměnné ED_{xx} značí Euklidovu vzdálenost mezi jednotlivými samohláskami. Pacienti postižení Parkinsonovou chorobou mají tuto plochu tvořenou formanty jednotlivých samohlásek výrazně menší, než tomu je u zdravých lidí [17].



Obr. 4.1: Grafické znázornění příznaku VSA.

	F_{1i}	F_{2i}	F_{1a}	F_{2a}	F_{1u}	F_{2u}	VSA	FCR
	[Hz]	[Hz]	[Hz]	[Hz]	[Hz]	[Hz]	[Hz ²]	[–]
Muž	342	2322	768	1333	378	997	264423	0,99
Žena	437	2761	936	1551	459	1105	399862	0,96
Dítě	452	3081	1002	1688	494	1345	448147	0,97
Střední hodnota	410	2721	902	1524	444	1149	370811	0,97
Směrodatná odchylka	60	381	121	179	60	178	95245	0,01
Koeficient kolísání	15 %	14 %	13 %	12 %	13 %	16 %	26 %	1 %

Tab. 4.1: Srovnání příznaků VSA a FCR na pohlaví a věku řečníků. Uvedené hodnoty platí pro anglický jazyk [17].

4.3.2 LnVSA

Logaritmická verze VSA (LnVSA) vychází z původní verze. U LnVSA jsou nejdříve jednotlivé formanty zlogaritmovány a až potom je počítána plocha VSA. Výhoda oproti VSA je, že výsledek je méně ovlivněn kolísáním hlasu. LnVSA lze vyjádřit vztahem: [17]

$$\ln VSA = \sqrt{(\ln S \cdot (\ln S - \ln ED_{iu}) \cdot (\ln S - \ln ED_{ia}) \cdot (\ln S - \ln ED_{au})).} \quad (4.7)$$

4.3.3 FCR

Nevýhodou příznaku VSA je, že se jeho hodnoty výrazně mění s kolísáním hlasu, se změnou věku a pohlaví řečníka. Všechny tyto nevýhody odstraňuje příznak FCR (Formant Centralization Ratio), který se na rozdíl od VSA počítá jako poměr formantů samohlásek. Příklad výpočtu poměrného počítání formantů: [17]

- Poměr formantů u dospělého muže s $F_{1i} = 300\text{Hz}$, $F_{2i} = 2400\text{Hz}$ je $\frac{F_{2i}}{F_{1i}} = 8$.
- Poměr formantů u dospělé ženy s $F_{1i} = 350\text{Hz}$, $F_{2i} = 2800\text{Hz}$ je $\frac{F_{2i}}{F_{1i}} = 8$.
- Poměr formantů u dítěte s $F_{1i} = 400\text{Hz}$, $F_{2i} = 3200\text{Hz}$ je $\frac{F_{2i}}{F_{1i}} = 8$.

Jedná se o modelový příklad, kde kolísání vychází přesně 0 %. Ve skutečnosti se hodnota kolísání příliš neliší od modelového příkladu a pohybuje v jednotkách procent. Z tab. 4.1 je zřejmé, že FCR je podstatně odolnější vůči pohlaví a věku řečníka než VSA [17].

Výpočet příznaku FCR pro samohlásky /a/, /i/ a /u/ lze vyjádřit vztahem: [17]

$$FCR = \frac{F_{2u} + F_{2a} + F_{1i} + F_{1u}}{F_{2i} + F_{1a}}. \quad (4.8)$$

4.3.4 F_{2i}/F_{2u}

Stejně jako příznak FCR je příznak F_{2i}/F_{2u} počítán poměrově. Dokáže dobře rozlišit zdravého člověka od nemocného případně i jeho zlepšení. Ale na rozdíl od FCR, které

		F_{1i}	F_{1u}	F_{1a}	F_{2i}	F_{2u}	F_{2a}
		[Hz]	[Hz]	[Hz]	[Hz]	[Hz]	[Hz]
PD – L	Střední hodnota	330	361	756	2417	1364	1326
	Směrodatná odchylka	67	60	114	303	193	182
	Koeficient kolísání	20,3 %	16,6 %	15,1 %	12,5 %	14,1 %	13,7 %
PD – NL	Střední hodnota	338	363	786	2480	1323	1335
	Střední odchylka	30	45	103	335	216	135
	Koeficient kolísání	8,8 %	12,4 %	13,0 %	13,5 %	16,3 %	10,1 %
ZJ	Střední hodnota	318	384	788	2565	1189	1307
	Směrodatná odchylka	41	35	89	222	163	102
	Koeficient kolísání	13,0 %	9,2 %	11,3 %	8,7 %	13,7 %	7,8 %

Tab. 4.2: Přehled formantů a jejich kmitočtů pro samohlásky /i/, /u/ a /a/ (PD – L – léčený pacient, PD – NL – neléčený pacient, ZJ – zdravý člověk) [17].

		FCR	VSA	LnVSA	F_{2i}/F_{2u}
		[-]	[Hz ²]	[LnHz]	[-]
PD – L	Střední hodnota	1,07	217551	0,21	1,79
	Střední odchylka	0,08	99982	0,08	0,24
	Koeficient kolísání	7,5 %	46,0 %	37,7 %	13,5 %
PD – NL	Střední hodnota	1,03	233508	0,24	1,90
	Střední odchylka	0,09	83369	0,07	0,24
	Koeficient kolísání	8,3 %	35,7 %	28,7 %	12,9 %
ZJ	Střední hodnota	0,96	280420	0,28	2,18
	Střední odchylka	0,07	77579	0,07	0,27
	Koeficient kolísání	7,60 %	27,7 %	24,0 %	12,5 %

Tab. 4.3: Parametry jednotlivých příznaků VSA, LnVSA, FCR, F_{2i}/F_{2u} (PD – L – léčený pacient, PD – NL – neléčený pacient, ZJ – zdravý člověk) [17].

je počítáno ze tří samohlásek a ze dvou formantů, je příznak F_{2i}/F_{2u} počítán jen ze dvou samohlásek a z jednoho formantu. Tota omezující vlastnost dělá příznak FCR efektivnější k odhalování artikulačních abnormalit [17].

4.3.5 Shrnutí

V tab. 4.3, která vychází z hodnot tab. 4.2 (v této tabulce jsou průměrné hodnoty formantů od 38 anglicky mluvících pacientů a 14 zdravých lidí, z toho bylo 19 pacientů podrobeno 3 denní intenzivní léčbě hlasu), jsou uvedeny důležité hodnoty pro všechny čtyři zmíněné příznaky. Lze z ní odvodit, že všechny 4 příznaky (VSA, LnVSA, FCR, F_{2i}/F_{2u}) mají dobrou rozlišovací schopnost rozeznání zdravého od nemocného člověka dokonce dokáží rozlišit i případné zlepšení stavu pacienta při léčbě. Výhoda FCR je, že není ovlivněn kolísáním hlasu [17].

4.4 VOT

Příznak VOT (Voice Onset Time) je definován jako interval mezi zákmitem plosiv¹ a začátkem samohlásky. Hodnota VOT je pro různý druh souhlásek i samohlásek různá. Pro bilabiální² souhlásky má VOT hodnotu nejmenší, alveolární³ ji mají středně velkou a velární⁴ mají hodnotu VOT největší. U samohlásek je větší hodnota VOT pro vyšší samohlásky (/i/, /u/). Dalšími faktory, které hodnotu VOT mohou ovlivnit, jsou: věk řečníka, tempo, znělost a hlasitost proslovu [5].

Pro diagnózu Parkinsonovy choroby příznakem VOT se používají nejčastěji samohlásky /u/ a /i/ a neznělé souhlásky (/p/, /t/, /k/). A to proto, že mají u pacientů s Parkinsonovou chorobou největší vliv na hodnotu VOT [5].

4.4.1 VOT ration

VOT ration vychází s příznaku VOT jako funkce tempa řeči. VOT ration není závislé na tempu řeči. Výpočet pak probíhá podle vztahu:

$$\text{VOT ration} = \left[\frac{\text{VOT}}{\text{doba trvání slova}} \right]. \quad (4.9)$$

Hlavní problémem příznaku VOT je to, že provedené výzkumy se různí v tom, jestli hodnota VOT má být pro pacienty postižené Parkinsonovou chorobou větší nebo menší [5].

4.4.2 VAI

Příznak VAI (Vowel Articulation Index) obdobně jako předchozí příznaky se počítá z kmitočtů F_1 a F_2 pro tři samohlásky /a/, /i/ a /u/. Tento parametr je nezávislý na pohlaví řečníka. Příznak VAI je vyjádřen vztahem: [20]

$$\text{VAI} = \frac{F_{2i} + F_{1a}}{F_{1u} + F_{1i} + F_{2u} + F_{2a}}. \quad (4.10)$$

4.5 Další příznaky používané k diagnóze Parkinsonovy choroby

- Articulation rate – počet slabik vyslovených za dobu jedné vteřiny po odstranění pauz větších než 50 ms [6].

¹Plosivní souhláska vzniká při uvolnění závěru (okluze) v hlasovém traktu.

²Pro bilabiální souhlásky je uzávěr tvořen rty (/b/, /p/).

³Pro alveolární souhlásky je uzávěr tvořen jazykem a dásňovým obloukem (/d/, /t/).

⁴Pro velární souhlásky je uzávěr tvořen zadní částí jazyka a uvulou (/g/, /k/).

- Pause ration – procentuální vyjádření počtu pauz ve vzorku.
- Net Speech Rate (NSR) – počet vyslovených slabik za dobu jedné vteřiny v celé řeči po odstranění pauz větších než 10ms [6].
- Total Speech Rate (TSR) – počet vyslovených slabik za vteřinu [20].
- Variation Range (F_0 VR) – rozdíl mezi maximální a minimální hodnotou F_0 [19].
- Relative Variation Range ($\text{rel}F_0\text{VR}$) – $\frac{F_0\text{VR}}{\text{mean}F_0} \cdot 100$ [%], relativní kolísání F_0 [19].
- Relative Standart Deviation ($\text{rel}F_0\text{SD}$) – $\frac{F_0\text{SD}}{\text{mean}F_0} \cdot 100$ [%], relativní směrodatná odchylka F_0 [19].
- Interpause-Speech Duration (ISD) – doba trvání mezi dvěma pauzami [3].
- Speech Index of Rhythmicity (SPIR) – počet mezislovních pauz za minutu [3].

5 METODY VÝBĚRU PŘÍZNAKŮ A JEJICH OPTIMALIZACE

Akusticky podobné zvuky lze vyjádřit příznakovými vektory, které při geometrickém znázornění leží blízko sebe a tvoří shluky. Tyto shluky tvoří při zpracování akustického signálu jednotlivé třídy zvuků. Metody výběru jsou založeny právě na těchto geometrických vlastnostech, kde se metody výběru příznaků snaží nalézt takovou transformaci, která změní dimenzi geometrického prostoru, kde lze jednotlivé třídy od sebe separovat s dostatečným rozptylem [13].

Metodu výběru příznaků lze rozdělit do třech kroků:

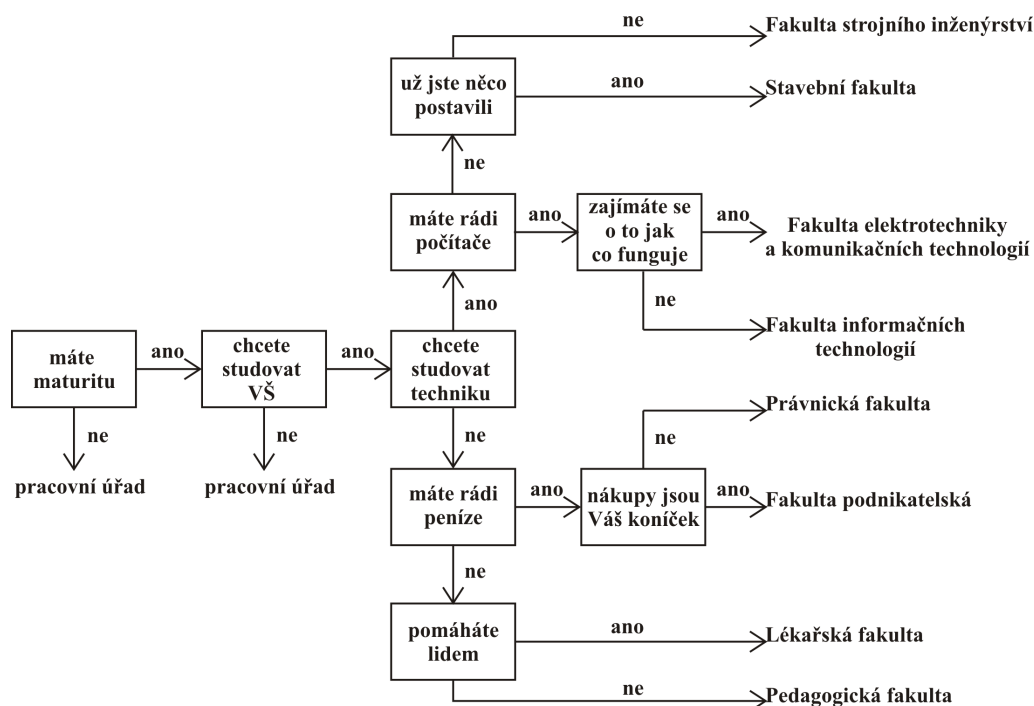
1. V prvním kroku je nutné určit vhodné příznakové vektory. Tyto vektory se používají jako reprezentativní množina mnohem větší množiny objektů, z které lze jednoznačně určit, do které třídy patří.
2. V dalším kroku jsou z příznakových vektorů na základě pozorování zjištěny charakteristické vlastnosti objektů.
3. V posledním kroku jsou zbylé objekty, podle charakteristických vlastností, rozděleny do příslušných tříd [23].

5.1 Učící se algoritmy

Učící se algoritmy slouží k automatizovanému výběru příznaků. Využívají různé algoritmy, které v různých aplikacích dosahují rozdílných výsledků. Proto je nutné, aby člověk, který je využívá, měl znalost dané problematiky a použil vhodný algoritmus pro danou úlohu [1].

5.1.1 Důvěryhodnost toho co bylo naučeno

Nejčastější chyba při použití učících se algoritmů pro trénování klasifikátoru je vybrání takových parametrů, které dávají nejlepší výsledky pro danou množinu dat, ale dále tento model není už testován na validačních datech. Při takovém přístupu dochází velice často k přetrénování klasifikátoru a daný klasifikátor není prakticky použitelný. K přetrénování klasifikátoru může dojít v tom případě, kdy se dvě třídy spolu navzájem prolínají. Přetrénovaný klasifikátor pak obě třídy rozdělí přesně podle těchto překrytých částí, i když ve skutečnosti to bývá často jinak. Proto je vždy nutné celou množinu rozdělit na data: trénovací, testovací a validační [1].



Obr. 5.1: Rozhodovací strom pro výběr dalšího studia

5.1.2 Rozhodovací stromy

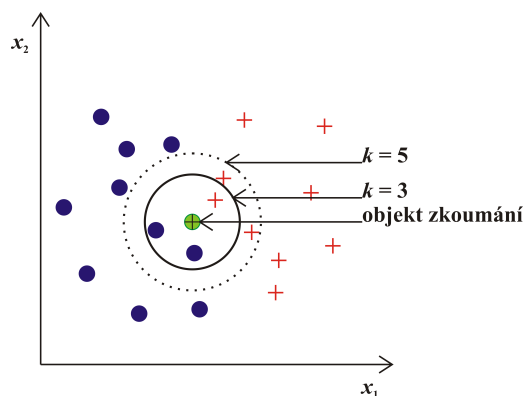
Rozhodovací stromy patří k nejjednodušším metodám výběru příznaků. Výhodou tohoto algoritmu je, že ho může používat i člověk, který není odborníkem na danou problematiku, avšak za cenu horších výsledků [1].

Algoritmů pro vytváření stromových struktur existuje celá řada jako například Huntův algoritmus, CART, ID3, SLIQ... Na obr. 5.1 je vidět příklad jednoduchého rozhodovacího stromu. Tento rozhodovací strom slouží k výběru vhodné vysoké školy [1].

5.1.3 Metoda k nejbližších sousedů

Metoda k nejbližších sousedů (k -NN – k Nearest-Neighbor) vychází z předpokladu, že zkoumaný prvek patří do stejné třídy, jako jeho k nejbližších sousedů. Pro změření vzdálenosti mezi prvky lze použít několik metrik (euklidovská, čebišova, korelační, kosinova...) [1].

Na obr. 5.2 je vidět klasifikace neznámého prvku. Pokud je nastavena hodnota $k = 3$, tak bude přiřazen do jiné třídy než když bude $k = 5$, proto je důležité, aby hodnota k byla určena správně. Důležitou podmínkou je volit k jako liché číslo, z důvodu jednoznačnosti klasifikace [1].



Obr. 5.2: Klasifikace pomocí k -NN

5.1.4 Umělé neuronové sítě

Umělé neuronové sítě (ANN – Artificial Neural Network) vychází z poznatků biologického chování neuronů, od kterých převzala jednoduchý matematický model. Neuronovou síť si lze představit, jako učící se algoritmus s adaptovatelnou schopností pro daný problém. [12].

Popis neuronu

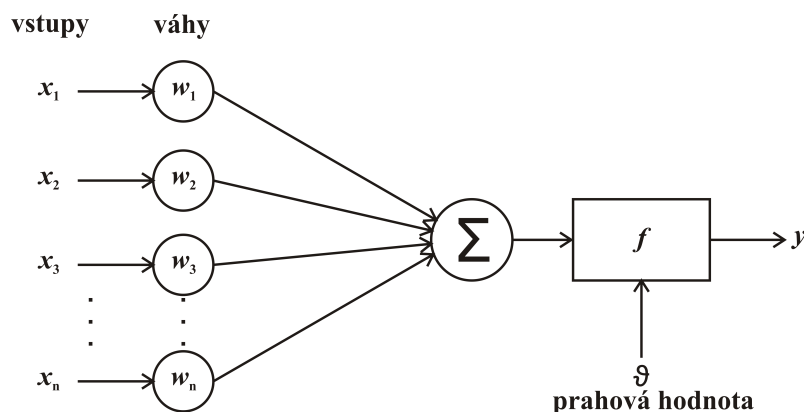
Neuron má n vstupů a jeden výstup (obr. 5.3), je popsán rovnicí:

$$y = f \left[\sum_{i=1}^n w_i x_i - \vartheta \right], \quad (5.1)$$

kde x_i je hodnota na i -tém vstupu, w_i je váha i -tého vstupu, ϑ je prahová hodnota a f je nelineární funkce [12].

Vrstvová struktura

Je jasné, že samotný neuron nemůže sám řešit složité úlohy, ale je jich potřeba více. Neurony jsou vzájemně propojeny a vzniká mezi nimi interakce. Propojeny jsou většinou do vrstev. Vrstvový model neuronů je dělen na vstupní vrstvu, která přijímá vstupní hodnoty pro zpracování a přivádí je na vstup každého neuronu v další vrstvě. Další n vrstev jsou skryté vrstvy, jejich počet závisí na složitosti problému. Poslední vrstva se nazývá výstupní a na jejím výstupu jsou výstupní hodnoty celého systému [12].



Obr. 5.3: Model neuronu

5.1.5 Algoritmy podpůrných vektorů

Algoritmy podpůrných vektorů (SVM – Support Vector Machine) jsou algoritmy patřící do skupiny jádrových algoritmů (kernel machines). Základním principem tohoto algoritmu je převod původního prostoru, kde jsou jednotlivé třídy od sebe lineárně neseparovatelné, do prostoru vyšší dimenze, kde lze pak jednotlivé třídy od sebe lineárně separovat [1, 25].

Nejsložitější úlohou tohoto algoritmu je nalezení co možná nejširší hranice mezi třídami. Tato optimální hranice se hledá pomocí kvadratického programování¹ [25, 24].

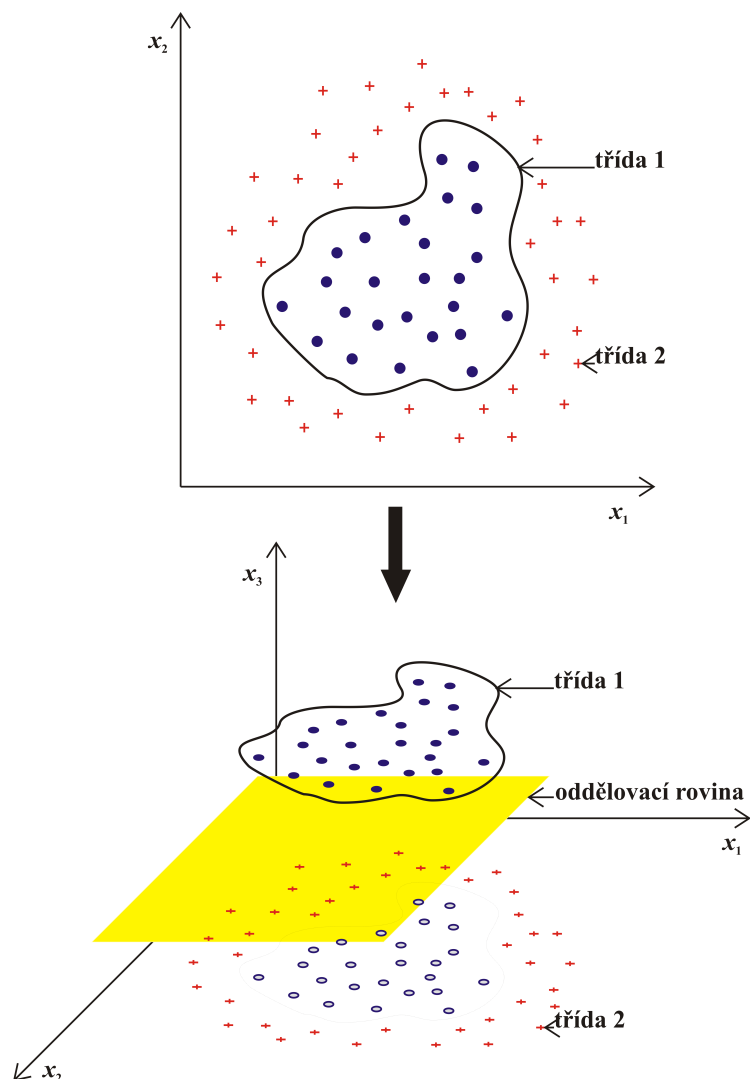
Na obr. 5.4 je vyobrazeno rozdělení dvou tříd posunutím původní dimenze do dimenze vyšší. V první části obrázku jsou vidět příznaky patřící do dvou nelineárně separovatelných tříd. V druhé části obrázku je přidána dimenze, která umožňuje tyto dvě třídy oddělit lineární rovinou [25].

Existuje několik variant jádrových algoritmů SVM (lineární, radiální, bodové, gaussovské ...) [1].

5.1.6 Cross-validace

Cross-validace je jednou z metod, podle které lze hodnotit kvalitu naučeného algoritmu na malém počtu prvků. Cross-validace pracuje zároveň s daty, které slouží k trénování tak i k testování. Proto dokáže věrohodně ohodnotit použitý algoritmus, i když je použito jen malé množství prvků [1]. Nejčastěji jsou používány dvě varianty této metody:

¹Kvadratické programování je speciální typ optimalizační matematické úlohy. Optimalizační problém spočívá v nalezení lineární funkce, která bude odpovídat kvadratické funkci několika proměnných.



Obr. 5.4: Oddělení dvou tříd pomocí SVM

1. „jeden vynechej“ (leave-one-out), jedná se o nejpřesnější metodu cross-validace, nevýhodou je větší výpočetní náročnost. Princip je takový, že se vybere vždy jen jeden prvek z celé množiny, který je určen pro testování a ostatní prvky jsou použity pro trénování. Tento postup se opakuje tolikrát, kolik je v dané množině prvků [1].
2. „vynechej k “ (k -fold), jedná se o metodu, kde se množina rozdělí na k podmnožin. Jedna z těchto podmnožin pak slouží jako testovací a ostatní $k - 1$ slouží jako trénovací podmnožiny. Tento postup se opakuje k -krát [1].

5.1.7 Vyhodnocení učících se algoritmů

Vyhodnocení učících se algoritmů lze v některých jednoduchých případech porovnávat pomocí srovnávacích vztahů, avšak těchto případů je velice málo. Je nutné používat

jeden z sofistikovaných nástrojů. Nejčastěji používaným nástrojem je matice záměn (confusion matrix). Z matice záměn lze vyvodit vztahy z kterých lze objektivně posoudit kvalitu algoritmů. Tuto matici lze vidět v tab. 5.1 [1].

	Pravděpodobná třída		
Skutečná třída		pozitivní	negativní
	pozitivní	A	B
	negativní	C	D

Tab. 5.1: Matice záměn

Z této matice lze pak vyjádřit několik parametrů:

- Přesnost (accuracy):

$$\text{přesnost} = \frac{A + D}{A + B + C + D} \quad (5.2)$$

Tento parametr při velkém počtu prvků v jedné třídě a malém počtu prvků v druhé třídě nemá příliš velkou vypovídací hodnotu. Je vhodné ho používat v řádově stejně velkých třídách. Například existuje 500 prvků pro jednu třídu a její vyhodnocení je správné. V druhé třídě existuje 10 prvků, ale ty jsou vyhodnoceny špatně. Výsledek tohoto hodnocení je takový, že přesnost v tomto případě je 98% a přitom celá jedna třída byla vyhodnocena nesprávně [1].

- Senzitivita (recall) je vyjádření negativních případů, které byly klasifikovány jako pozitivní: [1]

$$\text{senzitivita} = \frac{A}{A + B} \quad (5.3)$$

- Specifičnost je vyjádření negativních případů ku všem negativním případům:

$$\text{specifičnost} = \frac{D}{C + D} \quad (5.4)$$

- Pozitivní předpovídající hodnota je vyjádření poměru mezi správně předpovězených případů ke všem pozitivně předpovězeným:

$$\text{PPV} = \frac{A}{A + C} \quad (5.5)$$

- Negativní předpovídající hodnota je vyjádření poměru mezi špatně předpověděných případů ke všem špatně předpovězeným:

$$\text{NPV} = \frac{D}{D + B} \quad (5.6)$$

- Jedna z nejlépe odpovídajících hodnot pro srovnání algoritmů je Matthewův korelační koeficient (MCC – Matthews Correlation Coefficient):

$$\text{MCC} = \frac{A \cdot D - C \cdot B}{\sqrt{(A + C) \cdot (A + B) \cdot (D + C) \cdot (D + B)}} \quad (5.7)$$

Pokud by jmenovatel měl být roven 0, nastaví se na 1. $\text{MCC} \in \langle -1; 1 \rangle$, kde 1 udává 100% predikci, 0 náhodnou predikci a -1 inverzní predikci [1].

5.2 Evoluční algoritmy

Evoluční algoritmy jsou založeny na principu evoluce v přírodě. V přírodě se provádí přirozený výběr jedince tak, že jen nejlépe adaptovaný jedinec může přežít. Podobně to probíhá i u evolučních algoritmů, kde do reprodukce vstupují jen jedinci s nejvyšší hodnotou fitness funkce² [2].

5.2.1 Genetické algoritmy

Genetické algoritmy jsou jednou z podmnožin evolučních algoritmů. Slouží k řešení složitých úloh pomocí poznatků z evoluce [2].

Používané operátory

V evoluci je definováno několik základních operátorů, které se podílí na výsledné změně populace. Operátory, které se vyskytují u genetických algoritmů jsou: selekce, křížení, mutace a rekombinace [2].

Selekce znamená výběr jedinců, kteří budou zahrnuti do nové populace. Upřednostňují se jedinci s vyšší hodnotou fitness funkce, ale zároveň musí být vybírání jedinci i s menší hodnotou fitness funkce, aby se zajistila rozmanitost populace [2]. Existuje několik metod výběru:

1. proporcionalní selekce (roulette wheel selection) – výběr závisí na hodnotě fitness funkce, aby nedocházelo k postupnému nahrazení celé populace jedincem s vysokou hodnotou fitness funkce, tak se původní fitness funkce upraví pomocí: [2]

- Komprimace fitness funkce (windowing) lze vyjádřit vztahem:

$$f'(i) = f(i) + \beta^t, \quad (5.8)$$

kde β vyjadřuje nejhorší fitness funkci v t -generaci.

- Sigma škálování (scaling) lze vyjádřit vztahem:

$$f'(i) = \max \left(f'(i) (-E(f) - c \cdot \sigma_f), 0 \right), \quad (5.9)$$

kde $c = 2$.

2. lineární uspořádání (linear ranking) – tento druh selekce vyžaduje setříděnou populaci, kde nejhorší jedinec má index 1 a nejlepší N . Výběr lze vyjádřit vztahem:

$$p(i) = \frac{2 - s}{N} + \frac{2i(s - 1)}{N(N - 1)}, \quad (5.10)$$

kde $i \in \{1, 2, \dots, N\}$ a s je hodnota selekční intenzity, obvykle 1,5 nebo 2.

²Fitness funkce vyjadřuje kvalitu daného řešení. Tzn. jedinec který má největší pravděpodobnost k přežití v dané populaci.

3. exponenciální uspořádání (exponential ranking) – tato selekce je podobná jako lineární uspořádání. Rozdíl je v tom, že rozložení pravděpodobnosti má exponenciální závislost. Výběr lze vyjádřit vztahem:

$$p(i) = \frac{1 - e^{-i}}{c}, \quad (5.11)$$

kde $i \in \{1, 2, \dots, N\}$ a c je normalizační konstanta, která vyjadřuje velikost populace, bývá volena v rozsahu $0 < c < 1$.

4. turnajová selekce (tournament selection) – výhodou je, že nepotřebuje mít seřazenou populaci. Výběr je proveden tak, že z N jedinců se vybere t jedinců, z těchto t jedinců se vybere jedinec s největší fitness funkcí [2].

Křížení znamená vytváření nového jedince na základě zkřížení genetických informací rodičů. Hlavním účelem křížení je vznik rozmanité populace. Genetické algoritmy používají dva druhy křížení: [2]

1. N -bodové – zde probíhá křížení v N bodech mezi dvěma rodiči. Výsledkem jsou dva potomci, kteří obsahují část z každého rodiče [2].
2. Uniformní křížení – křížení probíhá tak, že s pravděpodobností 50 % zdědí jeden gen po jednom rodiči a druhý gen po druhém. Tento druh křížení vnáší do procesu obranu před brzkou konvergencí [2].

Mutace znamená změnu genetické informace s velmi malou pravděpodobností. Pravděpodobnost mutace nesmí být příliš vysoká, aby nedošlo k nestabilitě vývoje další populací. Nejvýznamněji se projevuje u malých populací [2].

Rekombinace znamená obnovení populace, existují dva základní způsoby: [2]

1. S úplnou obnovou – rodiče jsou kompletně nahrazeni rodiči.
2. S částečnou obnovou – jeden potomek nahradí nejslabšího jedince.

V praxi jsou často uplatňovány oba zmíněné způsoby, kdy se nahradí například půlka populace [2].

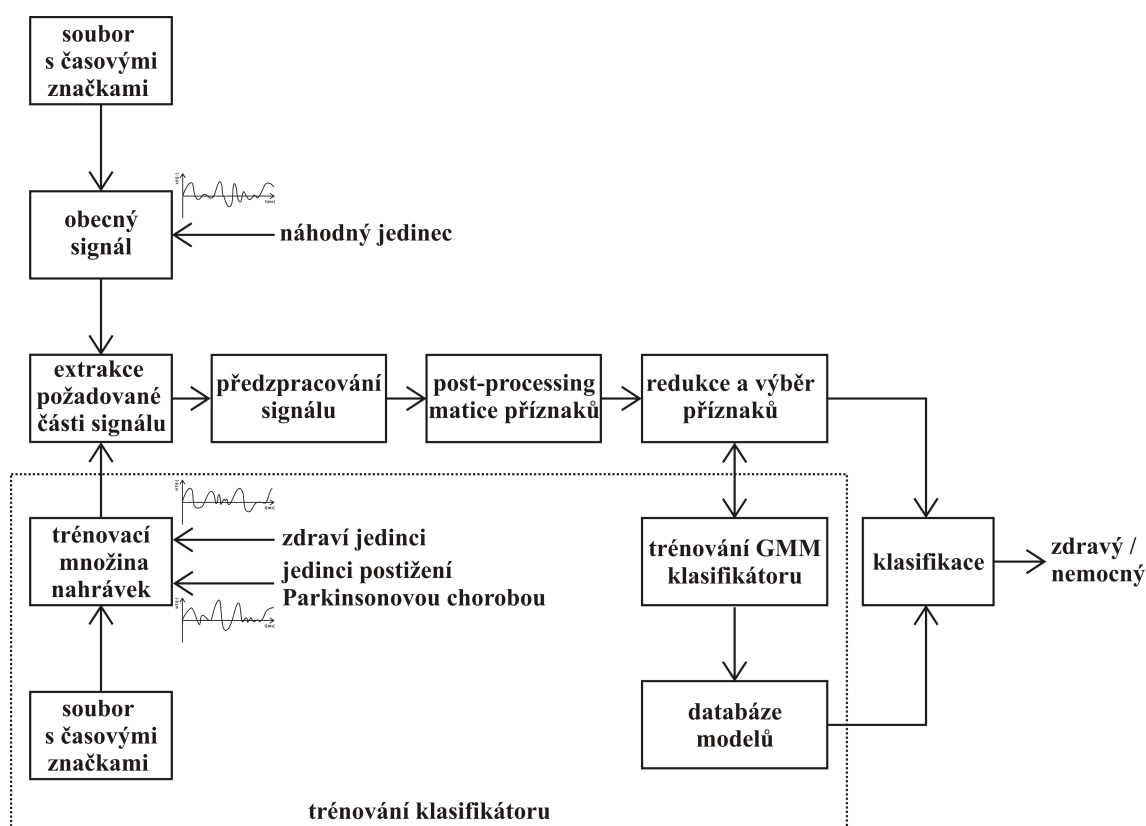
Proces genetických algoritmů

Celý proces genetických algoritmů lze shrnout do několika bodů: [2]

- inicializace, vytvoření náhodné populace
- ohodnocení všech jedinců pomocí fitness funkce
- vytvoření cyklu, ten bude probíhat dokud nedojde ke splnění určené podmínky
- inkrementace čísla populace
- selekce rodičů podle určené metody výběru
- rekombinace vybraných rodičů
- mutace potomků
- ohodnocení potomků podle fitness funkce
- obnova populace [2]

6 SYSTÉM PRO DIAGNÓZU PARKINSONOVY CHOROBY

V této části práce bude popsán systém pro diagnózu Parkinsonovy choroby, vytvořený na základě znalostí získaných z předchozích kapitol. Blokové schéma tohoto systému je vidět na obr. 6.1. Schéma je složeno ze dvou větších částí. První část se stará o výpočet samotných příznaků, byla vytvořena v programovém prostředí MATLAB (verze: 7.9.0.529 (R2009b), 64-bit). Druhá část se stará o natrénování klasifikátoru a následnou klasifikaci, byla vytvořena v programovém prostředí RapidMiner (verze: 5.1.006). Obě tyto části budou v následujícím textu důkladně popsány.



Obr. 6.1: Blokové schéma systému pro diagnózu Parkinsonovy choroby

6.1 Nahrávky

Diplomová práce „Moderní řečové příznaky používané při diagnóze chorob“ probíhá ve spolupráci s I. neurologickou klinikou v nemocnici u sv. Anny v Brně, která také poskytl nahrávky. Bohužel počet těchto nahrávek byl doposud velmi omezený

(tab. 6.1). Proto bylo nutné, k otestování funkčnosti klasifikátoru, vytvořit několik dalších kontrolních nahrávek (tab. 6.1). Tyto nahrávky byly pořízeny na Ústavu telekomunikací FEKT VUT v Brně.

	Muž	Žena
Zdravý	2	0
Pacient	11	9

Tab. 6.1: Počet nahrávek od fakultní nemocnice sv. Anny v Brně

	Muž	Žena
Zdravý	41	6
Pacient	11	9

Tab. 6.2: Počet nahrávek celkem

Tato skutečnost nese dva problémy, s kterými bylo v práci dále počítáno. Prvním problémem je, že nahrávky od fakultní nemocnice sv. Anny v Brně obsahují celou škálu cvičení, které byly s řečníky prováděny (viz. příloha B), ale nahrávky zaznamenané na FEKT VUT v Brně obsahují jen cvičení se samohláskami (7.1_1_a – 7.1_4_u), ty jsou také, nejčastěji využívány k diagnóze Parkinsonovy choroby, jak bylo uvedeno v kapitole 4. Druhým problémem je, že nahrávky pocházející z fakultní nemocnice u sv. Anny v Brně a nahrávky z FEKT VUT v Brně, byly nahrány v jiném akustickém prostředí a byl použit jiný mikrofon. Tyto dvě skutečnosti ovlivnili trénování a následnou klasifikaci, proto jsou prováděny jen za pomoci cvičení založených na samohláskách a na příznacích, které nejsou založeny na energii signálu.

6.2 Výpočet příznaků

Celý proces výpočtu příznaků se skládá z několika funkcí napsaných v programovém prostředí MATLAB. Do tohoto prostředí je implementován volně stažitelný konzolový program Praat. Praat je určen k výpočtu několika nejběžněji využívaným příznaků ve zpracování řeči ($F_0, F_1, F_2, \dots$). Výstupem celého procesu je soubor typu *.xls. Tento soubor obsahuje ve sloupcích hodnoty jednotlivých příznaků, v řádcích jsou názvy nahrávek a jednotlivé sešity nesou názvy podle jednotlivých cvičení. Ukázka části tabulky je vidět na obr. 6.2. Celkově je počítáno 75 běžných příznaků a 5 globálních příznaků pro diagnózu Parkinsonovy choroby. Přehled a popis všech vypočítaných příznaků je uveden v příloze C.1. Každá z nahrávek nese název podle příslušné logiky:

- první písmeno značí řečníka K – zdravý, P – nemocný člověk

- první číslo značí pohlaví řečníka 1 – žena, 2 – muž
- další písmena a číslice rozlišují jednotlivé nahrávky

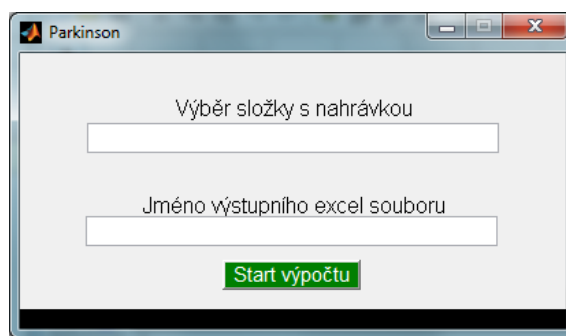
Například: „P1006zemanova“ – jedná se o pacienta, ženského pohlaví.

	A	B	C	D	E	F	G	H	I	J	K
1	Jmena/Pr	F0_media	F0_mean	F0_std	F0_min	F0_max	F0_max_n	relF0VR	relF0SD	voicing_fr	voicing_nl
2	K0100.wav	99,579	100,624	3,382	97,165	109,792	12,627	12,5487	3,361027	0	0
3	K0101.wav	106,848	108,061	6,056	100,75	120,33	19,58	18,1194	5,604242	15,385	0
4	K0102.wav	113,751	112,883	6,647	98,686	126,167	27,481	24,34468	5,888398	0	0
5	K0103.wav	116,202	116,811	2,013	114,837	123,813	8,976	7,684208	1,723297	4,762	0
6	K0104.wav	136,944	134,57	15,51	78,22	150,635	72,415	53,81214	11,5256	8	0
7	K0105.wav	105,361	105,333	1,596	102,327	107,781	5,454	5,177864	1,515195	0	0
8	K0106.wav	113,781	113,737	3,236	109,083	126,513	17,43	15,32483	2,84516	0	0
9	K0107.wav	106,373	107,498	3,597	102,195	114,713	12,518	11,64487	3,346109	0	0
10	K0108.wav	128,423	132,102	13,221	114,755	157,873	43,118	32,63993	10,00818	3,125	0
11	K0109.wav	111,94	117,156	13,821	99,737	139,714	39,977	34,12288	11,79709	0	0
12	K0110.wav	120,213	127,419	20,876	104,315	168,79	64,475	50,60077	16,38374	0	0
13	K0111.wav	139,968	139,905	5,648	131,985	149,357	17,372	12,417	4,037025	9,524	0
14	K0112.wav	154,161	153,977	8,223	129,278	168,482	39,204	25,46095	5,340408	0	0
15	K0113.wav	117,116	119,606	6,357	113,019	136,488	23,469	19,62193	5,314951	7,895	0
16	K0114.wav	121,527	121,266	1,594	116,282	124,193	7,911	6,523675	1,314466	2,128	0
17	K0115.wav	101,72	102,104	4,744	94,622	110,975	16,353	16,01602	4,646243	2,273	0

Obr. 6.2: Část vygenerované tabulky příznaků

6.2.1 Grafické prostředí

Jak bylo popsáno v minulé kapitole, celý program je složen z několika funkcí, proto bylo vytvořeno jednoduché grafické prostředí pomocí, kterého lze celý proces výpočtu příznaků snadno spustit (obr. 6.3).



Obr. 6.3: Grafické prostředí pro výpočet příznaků

Do řádku „Výběr složky s nahrávkou“ se doplní absolutní cesta ke složce obsahující nahrávky a příslušné *.lab soubory. A do řádku „Jméno výstupního excel souboru“ se doplní jméno výstupní excel souboru. Ten pak bude obsahovat všechny vypočítané příznaky.

6.2.2 Extrakce požadované části signálu

Nejdříve je nutné extrahovat část signálu, pro část pak budou vypočteny jednotlivé příznaky. V tomto případě se jedná o úseky reprezentující jednotlivá cvičení prováděná s řečníky. Extrakce probíhá ze souboru typu `*.wav` na základě přiloženého souboru typu `*.lab`, obsahující časové značky začátku a konce jednotlivých cvičení. Oba názvy souborů se spolu musí shodovat.

6.2.3 Předzpracování signálu

Extrahovaný signál je nutné před vlastním zpracováním nejdříve upravit do „použitelné podoby“ tj. odstranit rušivou stejnosměrnou složku, segmentovat signál a další úpravy popsáné v kapitole 2. Praat se sám stará o předzpracování signálu pro počítané příznaky. O ostatní příznaky se starají funkce napsané v programovém prostředí MATLAB.

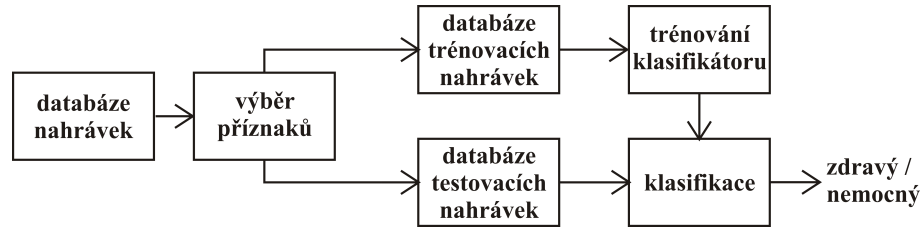
6.2.4 Post-processing matice příznaků

Některé z příznaků popsaných v kapitole 4, sloužících k diagnóze Parkinsonovy choroby potřebují ke svému výpočtu více než jednu hodnotu určitého příznaku. Matice příznaků obsahuje tyto jednotlivé hodnoty. Blok post-processing matice příznaků vybírá z matice jednotlivé hodnoty a následně z nich počítá složitější příznaky (např. pro příznak FCR se vyberou z matice příznaků hodnoty formantů F_{1a} , F_{1i} , F_{1u} , F_{2a} , F_{2i} , a F_{2u}).

6.3 Testování

Výsledná klasifikace závisí na několika faktorech. Nejdůležitějším faktorem je kvalitní a hlavně velká databáze nahrávek, kterou je nutné rozdělit na trénovací a testovací. Následně se testuje použitý algoritmus. Tím se částečně odstraňuje nezávislost testovacích dat, proto je nutné mít navíc data validační. Pokud je databáze nahrávek malá, jako je to v tomto případě, je nutné použít některou s komplexnějších metod ohodnocení. Jednou z těchto metod je cross-validace. Další podmínkou dosažení dobrých výsledků je vybrání vhodných příznaků pro řešení dané problematiky. Blokové schéma klasifikátoru je vidět na obr. 6.4.

Seznam a popis příznaků použitých pro trénování a klasifikaci jsou uvedeny v příloze C.2. V tabulce jsou uvedeny pouze všechny typy příznaků. Ve skutečnosti se při tomto procesu používá mnohem více příznaků, kde jednotlivé typy příznaků náleží příslušnému cvičení, tzn. počet příznaků je roven součinu typu příznaků s počtem cvičení, příklad je uveden v tab. 6.3.



Obr. 6.4: Blokové schéma klasifikátoru

cvičení					
příznaky		7.1.1_a	7.1.1_e	...	7.2.1_u
	F_0 median	F_0 median 1 a	F_0 median 1 e	...	F_0 median 2 u
	F_0 mean	F_0 mean 1 a	F_0 mean 1 e	...	F_0 mean 2 u
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
	F_2b max–min	F_2b max–xmin 1 a	F_2b max–min 1 e	...	F_2b max–min 2 u

Tab. 6.3: Příklad všech příznaků pro klasifikaci

6.3.1 Databáze nahrávek

Nejdříve je nutné podotknout, že trénování i klasifikace probíhá pouze pro mužské řečníky a to z důvodu malého počtu nahrávek od ženských řečníků. Při požití nahrávek od ženských řečníků byly výsledky velice zkreslené a nevypovídající. Z tohoto důvodu nejsou v práci dále uváděny. Avšak po obdržení většího počtu nahrávek lze aplikovat celý výpočet i pro ženské řečníky.

Pro trénování a klasifikaci byly použity cvičení 7.1.1_a, 7.1.1_e, 7.1.1_i, 7.1.1_o, 7.1.1_u, 7.1.2_a, 7.1.2_e, 7.1.2_i, 7.1.2_o, 7.1.2_u, důvodem tohoto výběru je skutečnost, že pouze tato cvičení byla absolvována všemi pacienty a všemi kontrolními mluvčími.

6.3.2 Výběr příznaků

Důležitou funkcí každého klasifikátoru je vybrat relevantní příznaky k dané problematice. Výběr příznaků probíhá pomocí operátoru **Optimize Selection** vycházejícího z genetických algoritmů popsaných v kapitole 5.2.1. Po vyzkoušení několika nastavení se atributy ustálily na hodnotách:

- velikost populace – 5
- maximální počet generací – 30
- typ selekce – turnajová
- část populace určená pro selekci – 0,5
- inicializační změna mutace populace – 1,0

- pravděpodobnost křížení – 0,5
- typ křížení – uniformní

Při tomto nastavení udává výstup typu `performance` hodnoty v matici záměn uvedené v tab. 6.4.

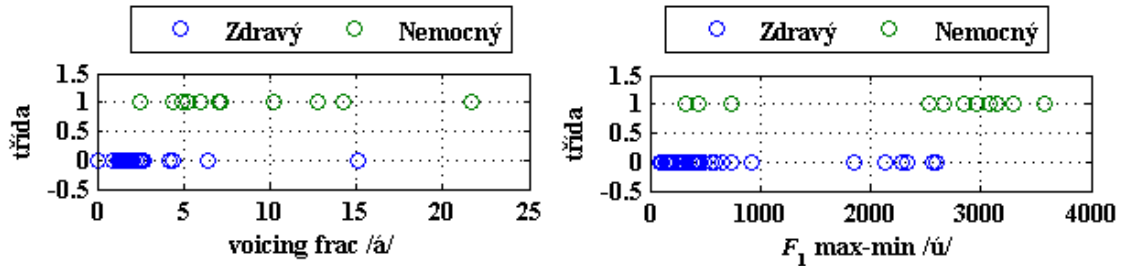
	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	41	1
	Negativní	0	10

Tab. 6.4: Matice záměn pro výběr příznaků

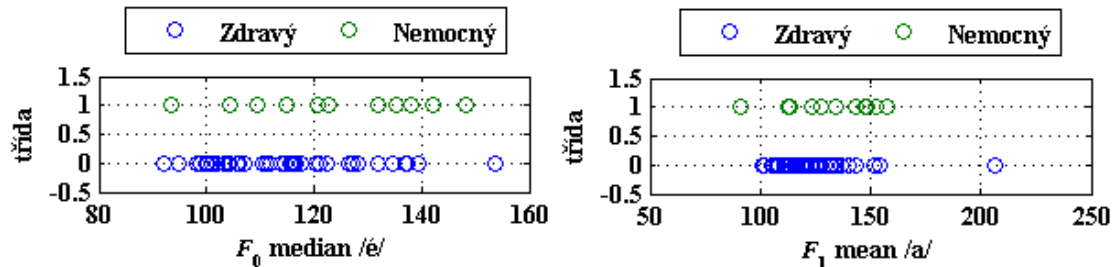
Přesnost se počítá podle vztahu 5.2 a její výsledek je 98,1 %. Lépe vypovídající algoritmus je Matthewův korelační koeficient počítaný podle vztahu 5.7. Výsledek MCC je 94,2 %.

Výsledek výběru – z 490 příznaků bylo vybráno 223 nejlepších příznaků pro trénování a klasifikaci.

Na obr. 6.5 jsou vidět dva vybrané příznaky, použité pro trénování a klasifikaci. Překrytí obou tříd je velice malé. Naopak na obr. 6.6 jsou vidět dva příznaky, které z důvodu velkého překrytí obou tříd nebyly vybrány.



Obr. 6.5: Příznaky vhodné pro klasifikaci



Obr. 6.6: Příznaky nevhodné pro klasifikaci

Realizace výběru příznaků v RapidMineru

Výběr příznaků se skládá z několika navzájem propojených operátorů. Nejdůležitějším operátorem je **Optimize selection**, nastavený podle výše zmíněných parametrů. Na vstup tohoto operátoru jsou přivedeny pomocí **Retrieve** importované hodnoty z vytvořeného excel souboru. **Optimize selection** využívá corss-validace. Corss-validace je reprezentovaná operátorem **X-Validation**. V důsledku malého počtu nahrávek je použita cross-validace typu leave-one-out, popsána v kapitole 5.1.6. Operátor **X-Validation** využívá učící se algoritmus SVM nastavený s lineárním jádrovým algoritmem popsáným v kapitole 5.1.5, reprezentovaný operátorem **Support Vector Machine**. Na vstup **Support Vector Machine** jsou přivedeny importované hodnoty. SVM vytvoří model, který se aplikuje a ohodnotí. Takto ohodnocená data jsou využita pro testování dat pomocí operátor **X-Validation**. Výstupem **X-Validation** je vektor hodnotící dané řešení. Tento vektor je poskytnut **Optimize selection**, ten na jeho základě vybere příslušné příznaky. Vybrané příznaky se nakonec zapíše pomocí operátoru **Write Weights** do externího souboru *.wgt.

6.3.3 Testování algoritmů

Pro trénování a následnou klasifikaci byly vybrány 3 algoritmy a 3 množiny příznaků.

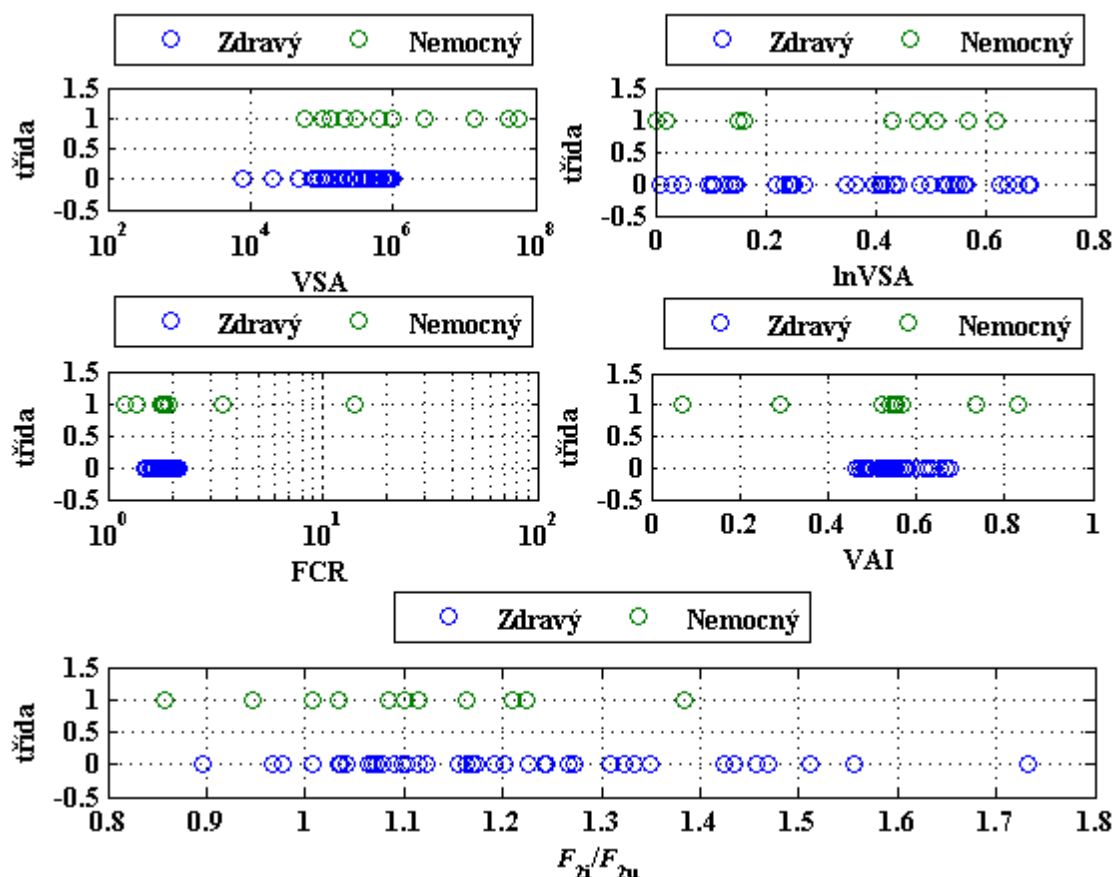
Přehled množin příznaků

- Všechny – množina obsahující všech 490 příznaků.
- Výběr – množina obsahující 223 vybraných příznaků.
- Parkinson – množina obsahující 5 příznaků (VSA, lnVSA, FCR, VAI a F_{2u}/F_{2i}). Rozložení nahrávek do příslušných tříd pomocí těchto příznaků je vidět na obr. 6.7.

Přehled použitých algoritmů a jejich parametrů

V RapidMineru je možné nastavovat mnoho parametrů u jednotlivých algoritmů. Nastavovány byly jen nejdůležitější parametry pro daný algoritmus:

1. k -nejbližších sousedů
 - volba k
2. Algoritmy podpůrných vektorů
 - jádrový algoritmus
3. Umělé neuronové sítě
 - velikost skryté vrstvy
 - počet tréninkových cyklů
 - rychlost učení



Obr. 6.7: Rozdělení nahrávek pomocí příznaků VSA, lnVSA, FCR, VAI a F_{2u}/F_{2i}

Realizace klasifikátoru v RapidMineru

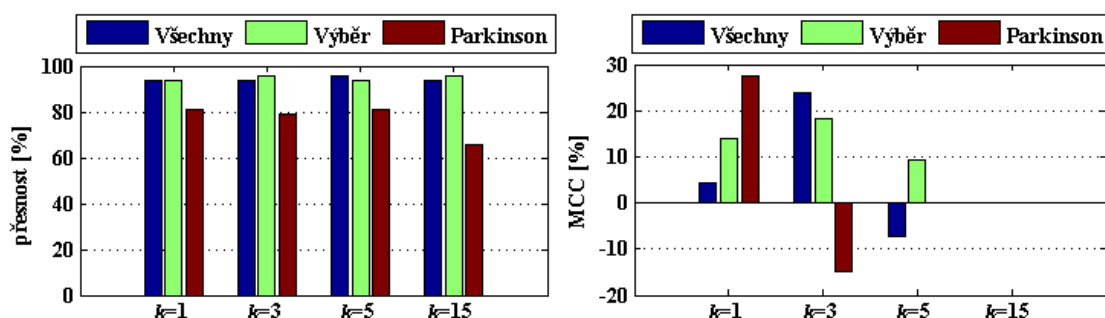
Obdobně jak tomu bylo u výběru příznaků, tak i klasifikátor je složen z několika dílčích operátorů. Typ klasifikátoru se liší pouze obměnou jednoho operátoru.

V tomto případě je ústředním operátorem klasifikátoru **X-Validation**, ten je opět nastaven na typ validace *leave-one-out*. Na jeho vstup je přivedena trénovací/testovací množina nahrávek pomocí **Retrieve**. **X-Validation** využívá jednoho z trojce operátorů podle toho jestli se jedná o klasifikátor typu *k*-NN (operátor **k-NN**), SVM (operátor **Support Vector Machine**) a nebo ANN (operátor **Neural Net**). Na vstup vybraného operátoru je přivedena trénovací/testovací množina nahrávek a jeho výstupem je vytvořený model. Ten je následně použit, ohodnocen a předán **X-Validation**. Jeho výstupem je vybraný model následně aplikovaný pro vyhodnocení validační množiny. Aplikaci modelu zajišťuje operátor **Apply Model** a následné vyhodnocení je provedeno pomocí operátoru **Performance**. Výstupem celého procesu je matice záměn pro validační množinu nahrávek.

$k=1$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	35	8
	Negativní	3	1
$k=3$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	36	7
	Negativní	2	2
$k=5$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	37	9
	Negativní	1	0
$k=15$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9
	Negativní	0	0

$k=1$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0
$k=3$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0
$k=5$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0
$k=15$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0

Tab. 6.5: Matice záměn k -NN klasifikátoru (všechny příznaky)



Obr. 6.8: Přesnost a MCC k -NN klasifikátoru

k -NN klasifikátor

Algoritmus k -NN popsaný v kapitole 5.1.3 byl použit pro různé hodnoty $k = 1, 3, 5, 15$ (k musí být liché číslo z důvodu jednoznačnosti). V tab. 6.5 jsou jednotlivé matice záměn pro různé k s použitím množiny příznaků „všechny“, v tab. 6.6 jsou matice záměn množiny příznaků „výběr“ a v tab. 6.7 jsou uvedeny matice záměn množiny příznaků „Parkinson“. Levá strana tabulek reprezentuje trénovací/testovací množinu a pravá strana reprezentuje validační množinu nahrávek.

Tab. 6.8 obsahuje výsledky výpočtů přesnosti a MCC pro všechny nastavení k -NN klasifikátoru. Tyto výsledky jsou pak dále vyjádřené graficky na obr. 6.8.

$k=1$	Pravděpodobná třída			$k=1$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	34	7		Pozitivní	3	2
	Negativní	4	2		Negativní	0	0
$k=3$	Pravděpodobná třída			$k=3$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	35	7		Pozitivní	3	2
	Negativní	3	2		Negativní	0	0
$k=5$	Pravděpodobná třída			$k=5$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	36	8		Pozitivní	3	2
	Negativní	2	1		Negativní	0	0
$k=15$	Pravděpodobná třída			$k=15$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9		Pozitivní	3	2
	Negativní	0	0		Negativní	0	0

Tab. 6.6: Matice záměn k -NN klasifikátoru (vybrané příznaky)

$k=1$	Pravděpodobná třída			$k=1$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	32	5		Pozitivní	3	2
	Negativní	6	4		Negativní	0	0
$k=3$	Pravděpodobná třída			$k=3$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	34	9		Pozitivní	3	2
	Negativní	4	0		Negativní	0	0
$k=5$	Pravděpodobná třída			$k=5$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9		Pozitivní	3	2
	Negativní	0	0		Negativní	0	0
$k=15$	Pravděpodobná třída			$k=15$	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9		Pozitivní	3	2
	Negativní	0	0		Negativní	0	0

Tab. 6.7: Matice záměn k -NN klasifikátoru (Parkinsonovy příznaky)

$k=1$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	76,6 %	76,6 %	76,6 %
MCC	4,5 %	13,8 %	27,5 %
$k=3$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	80,9 %	78,7 %	72,3 %
MCC	23,9 %	18,3 %	-14,8 %
$k=5$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	78,7 %	78,7 %	80,9 %
MCC	-7,1 %	9,4 %	0 %
$k=15$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	80,9 %	80,9 %	80,9 %
MCC	0 %	0 %	0 %

$k=1$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	60,0 %	60,0 %	60,0 %
MCC	0 %	0 %	0 %
$k=3$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	60,0 %	60,0 %	60,0 %
MCC	0 %	0 %	0 %
$k=5$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	60,0 %	60,0 %	60,0 %
MCC	0 %	0 %	0 %
$k=15$	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	60,0 %	60,0 %	60,0 %
MCC	0 %	0 %	0 %

Tab. 6.8: Vyhodnocení kvality klasifikace k -NN klasifikátoru

Lineární	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	2
	Negativní	0	7
Polynom.	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	6
	Negativní	0	3

Lineární	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	1
	Negativní	0	1
Polynom.	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0

Tab. 6.9: Matice záměn SVM klasifikátoru (všechny příznaky)

SVM klasifikátor

SVM algoritmus popsáný v kapitole 5.1.5 byl použit pro dva typy jádrových algoritmů: lineární a polynomiální. V tab. 6.9, 6.10 a 6.11 jsou stejně jako v předchozím případě matice záměn pro jednotlivé nastavení. Tyto tabulky nesou stejnou logiku jako v předchozím případě (levá strana trénovací/testovací, pravá strana validační množina nahrávek).

Tab. 6.12 obsahuje výsledky přesnosti a MCC. V tomto případě pro obě dvě nastavení jádrových algoritmů. Tyto výsledky jsou pak dále vyjádřené graficky na obr. 6.9.

Lineární	Pravděpodobná třída			Lineární	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	2		Pozitivní	3	1
	Negativní	0	7		Negativní	0	1
Polynom.	Pravděpodobná třída			Polynom.	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	7		Pozitivní	3	2
	Negativní	0	2		Negativní	0	0

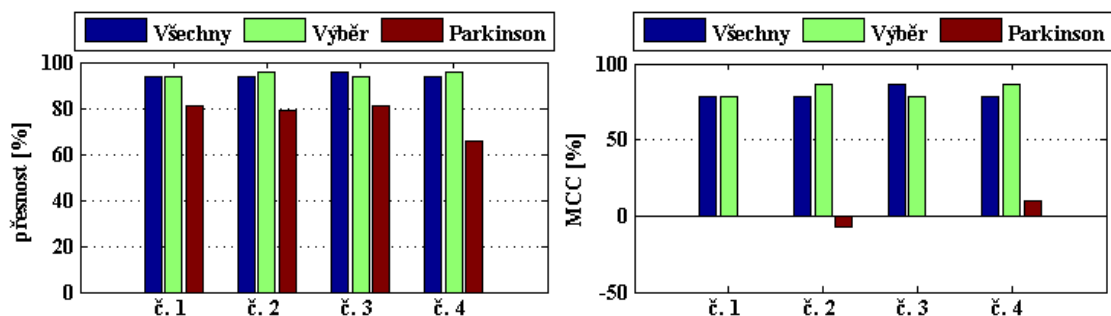
Tab. 6.10: Matice záměn SVM klasifikátoru (vybrané příznaky)

Lineární	Pravděpodobná třída			Lineární	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9		Pozitivní	3	2
	Negativní	0	0		Negativní	0	0
Polynom.	Pravděpodobná třída			Polynom.	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní	Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9		Pozitivní	3	2
	Negativní	0	0		Negativní	0	0

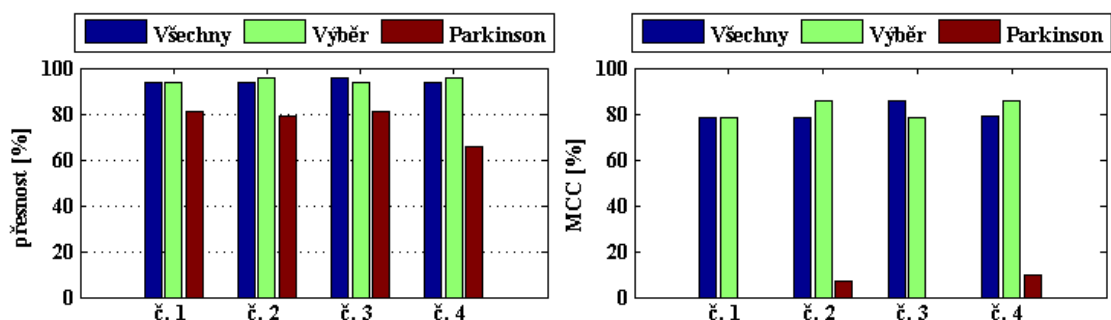
Tab. 6.11: Matice záměn SVM klasifikátoru (Parkinsonovy příznaky)

Lineární	Příznaky			Lineární	Příznaky		
	Všechny	Vybrané	Parkinson		Všechny	Vybrané	Parkinson
Přesnost	95,7 %	95,7 %	80,9 %	Přesnost	80,0 %	80,0 %	60,0 %
MCC	86,0 %	86,0 %	0 %	MCC	61,2 %	61,2 %	0 %
Polynom.	Příznaky			Polynom.	Příznaky		
	Všechny	Vybrané	Parkinson		Všechny	Vybrané	Parkinson
Přesnost	87,2 %	95,7 %	80,9 %	Přesnost	60,0 %	60,0 %	60,0 %
MCC	53,7 %	86,0 %	0 %	MCC	0 %	0 %	0 %

Tab. 6.12: Vyhodnocení kvality klasifikace SVM klasifikátoru



Obr. 6.9: Přesnost a MCC SVM klasifikátoru



Obr. 6.10: Přesnost a MCC ANN klasifikátoru

ANN klasifikátor

ANN algoritmus popsáný v kapitole 5.1.4 využívá umělé neuronové sítě. U ANN byly měněny tři parametry: počet neuronů ve skryté vrstvě, počet tréninkových cyklů a rychlost učení neuronů. Byly použity 4 různá nastavení:

1. Nastavení klasifikátoru č. 1:
 - Velikost skryté vrstvy – 1
 - Počet tréninkových cyklů – 10
 - Rychlost učení neuronů – 0,3
2. Nastavení klasifikátoru č. 2:
 - Velikost skryté vrstvy – 1
 - Počet tréninkových cyklů – 100
 - Rychlost učení neuronů – 0,5
3. Nastavení klasifikátoru č. 3:
 - Velikost skryté vrstvy – 5
 - Počet tréninkových cyklů – 10
 - Rychlost učení neuronů – 0,3
4. Nastavení klasifikátoru č. 4:
 - Velikost skryté vrstvy – 15
 - Počet tréninkových cyklů – 1000
 - Rychlost učení neuronů – 0,5

Pro jednotlivé nastavení (č. 1 – č. 4) byly jejich výsledky opět zapsány do matic záměn zapsaných v tab. 6.13, 6.14 a 6.15. Tyto tabulky nesou stejnou logiku jako v předchozích dvou případech (levá strana trénovací/testovací, pravá strana validační množina nahrávek).

Tab. 6.16 obsahuje výsledky přesnosti a MCC pro všechny nastavení ANN klasifikátoru. Tyto výsledky jsou pak dále graficky vyjádřené na obr. 6.10.

č. 1	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	3
	Negativní	0	6
č. 2	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	3
	Negativní	0	6
č. 3	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	2
	Negativní	0	7
č. 4	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	37	2
	Negativní	1	7

č. 1	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	1
	Negativní	0	1
č. 2	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	0
	Negativní	0	2
č. 3	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	0
	Negativní	0	2
č. 4	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	0
	Negativní	0	2

Tab. 6.13: Matice záměn ANN klasifikátoru (všechny příznaky)

č. 1	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	3
	Negativní	0	6
č. 2	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	2
	Negativní	0	7
č. 3	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	3
	Negativní	0	6
č. 4	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	2
	Negativní	0	7

č. 1	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	1
	Negativní	0	1
č. 2	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	1
	Negativní	0	1
č. 3	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	1
	Negativní	0	1
č. 4	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	1
	Negativní	0	1

Tab. 6.14: Matice záměn ANN klasifikátoru (vybrané příznaky)

č. 1	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9
	Negativní	0	0
č. 2	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	37	9
	Negativní	1	0
č. 3	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	38	9
	Negativní	0	0
č. 4	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	30	8
	Negativní	8	1

č. 1	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0
č. 2	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0
č. 3	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0
č. 4	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	3	2
	Negativní	0	0

Tab. 6.15: Matice záměn ANN klasifikátoru (Parkinsonovy příznaky)

č. 1	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	93,6 %	93,6 %	80,9 %
MCC	78,6 %	78,6 %	0 %
č. 2	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	93,6 %	95,7 %	78,7 %
MCC	78,6 %	86,0 %	-7,1 %
č. 3	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	95,7 %	93,6 %	80,9 %
MCC	86,0 %	78,6 %	0 %
č. 4	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	93,6 %	95,7 %	66,0 %
MCC	78,7 %	86,0 %	-9,9 %

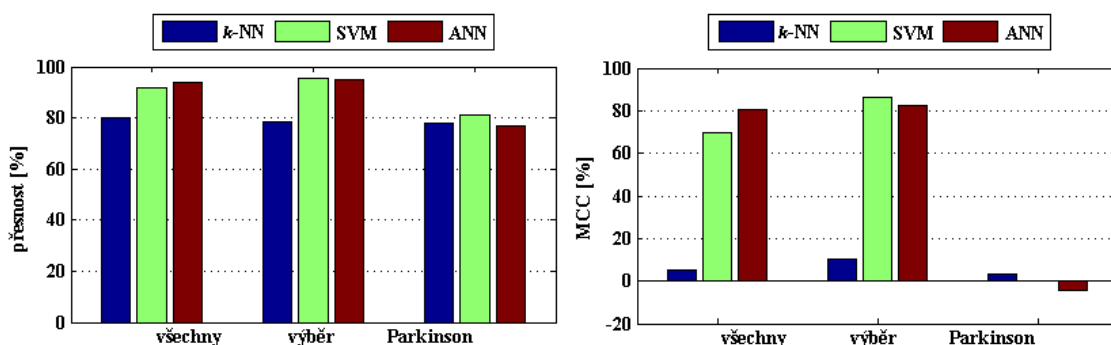
č. 1	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	80,0 %	80,0 %	60,0 %
MCC	61,2 %	61,2 %	0 %
č. 2	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	100,0 %	80,0 %	60,0 %
MCC	100 %	61,2 %	0 %
č. 3	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	100,0 %	80,0 %	60,0 %
MCC	100 %	61,2 %	0 %
č. 4	Příznaky		
	Všechny	Vybrané	Parkinson
Přesnost	100,0 %	80,0 %	60,0 %
MCC	100 %	61,2 %	0 %

Tab. 6.16: Vyhodnocení kvality klasifikace ANN klasifikátoru

6.3.4 Zhodnocení výsledků jednotlivých měření

Tato kapitola se zabývá zhodnocením výsledků trénovací/testovací množině příznaků.

Levá strana obr. 6.11 vyjadřuje průměrnou přesnost klasifikátorů. Nejpřesnější klasifikátor pro všechny tři množiny příznaků „všechny“, „výběr“ a „Parkinson“ byl ANN a SVM. V parametru přesnosti dosáhl dokonce podobných výsledků i k -NN klasifikátor, který je nejjednodušší ze všech tří. Těchto dobrých výsledků dosáhl zejména proto, že množina zdravých řečníků je téměř čtyřikrát větší než množina nemocných. Z tohoto plyne, jak bylo popsáno v kapitole 5.1.7 je parametr přesnost zavádějící.



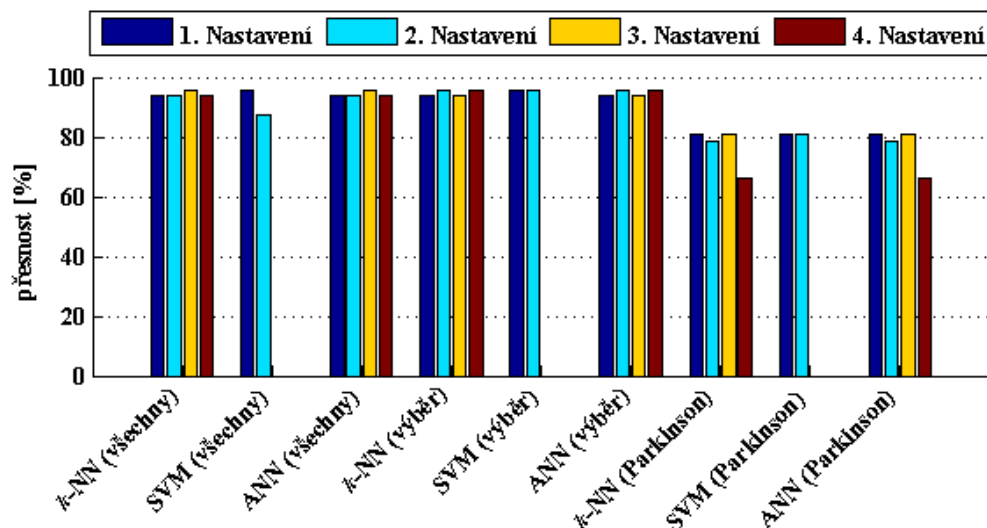
Obr. 6.11: Průměrná přesnost a MCC všech klasifikátorů.

MCC dává objektivnější výsledky hodnocení pro různě velké množiny oproti parametru přesnosti. Průměrné výsledky MCC klasifikátorů jsou vidět na pravé straně obr. 6.11. Podle hodnocení MCC se opět ukázal jako nejlepší klasifikátor ANN a SVM. Zde je dobře vidět neobjektivnost přesnosti, podle tohoto měření klasifikátor k -NN přiděluje jednotlivé nahrávky zcela náhodně do daných tříd. Vůbec nejlepších výsledků bylo v průměru dosaženo při použití množiny příznaků „výběr“ a naopak množina „Parkinson“ je zcela nevhodná. Tato skutečnost je způsobena malou množinou trénovacích nahrávek s velkým rozptylem hodnot. Rozptyl hodnot je na obr. 6.7.

Celkové hodnocení výsledků parametru přesnosti všech nastavení klasifikátorů je vidět na obr. 6.12.

Pro množinu příznaků „všechny“ a „výběr“ jsou výsledky pro všechny tři typy klasifikátorů i jejich nastavení velice podobné. Největší rozdíl nastal při klasifikaci pomocí množiny příznaků „Parkinson“, tato množina udává nejhorší výsledky. Nejhorším klasifikátorem byl k -NN nastaven na hodnotu $k = 15$ a klasifikátor ANN v nastavení č. 4.

Celkové hodnocení parametru MCC všech nastavení klasifikátorů je vidět na obr. 6.13. Jak bylo zmíněno výše parametr MCC více vypovídá skutečnosti oproti přesnosti u takového počtu nahrávek.



Obr. 6.12: Přesnost klasifikátorů pro testovací množinu nahrávek

Při použití množiny příznaků „všechny“, dosahoval klasifikátor k -NN nejlepších výsledků pro hodnotu $k = 3$, ale i tak kvalita klasifikace je nedostačující. Výrazně lepších výsledků dosáhly oba dva zbývající. V tomto případě je vidět velký rozdíl mezi SVM algoritmem při použití rozdílných jádrových algoritmů, kde lineární je téměř o 30 % lepší, než polynominální.

U další množiny „výběr“ dává k -NN klasifikátor opět nejhorší výsledky. SVM dává podobné výsledky pro obě nastavení jádrových algoritmů. SVM a ANN klasifikátor dosáhl podobných výsledků pro všechna 4 nastavení (č. 1 – č. 4).

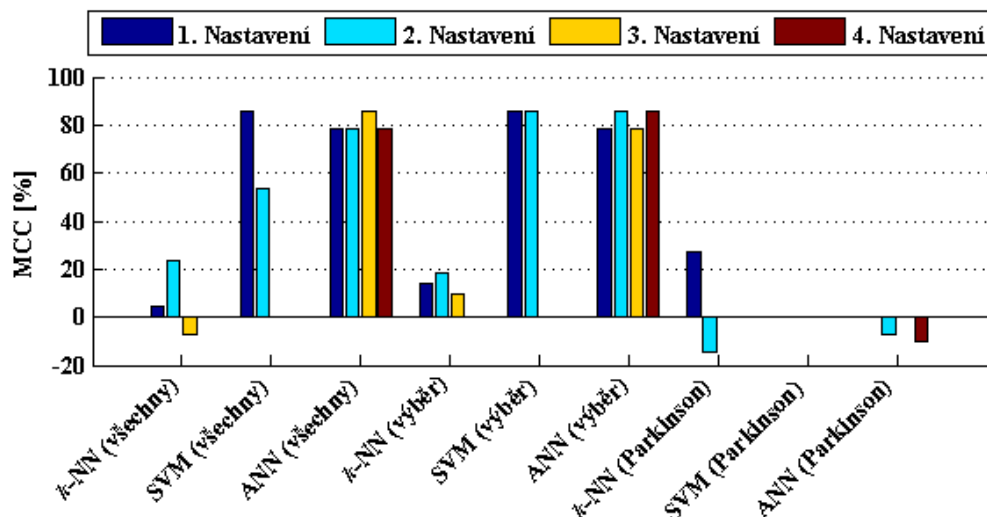
V poslední množině příznaků „Parkinson“ k -NN klasifikátor dosáhl nejlepších a zároveň nejhorších výsledků. Nejlepší výsledek byl dosažen při nastavení $k = 1$ a naopak nejhorších výsledků bylo dosaženo při nastavení $k = 3$. Ale ani při nastavení $k = 1$ nedosahují výsledky hranici 50 %, kde se klasifikace dá považovat za objektivní. Oba dva zbývající klasifikátory vykazují buď náhodnou možnou klasifikaci, nebo dokonce klasifikátor ANN inklinuje k inverzní klasifikaci.

Globální příznaky pro diagnózu Parkinsonovi choroby

Pro klasifikaci byly samostatně vyzkoušeny i globální příznaky (VSA, lnVSA, FCR...), ale dávaly ve všech nastavení klasifikátorů nulové hodnoty MCC, proto dále nejsou rozebírány.

Vyhodnocení výsledků validační množiny nahrávek

V předchozí kapitole byly vyhodnoceny všechny klasifikátory pro tři množiny příznaků za použití trénovací/testovací množiny nahrávek. Tato kapitole se zabývá



Obr. 6.13: MCC klasifikátorů pro testovací množinu nahrávek

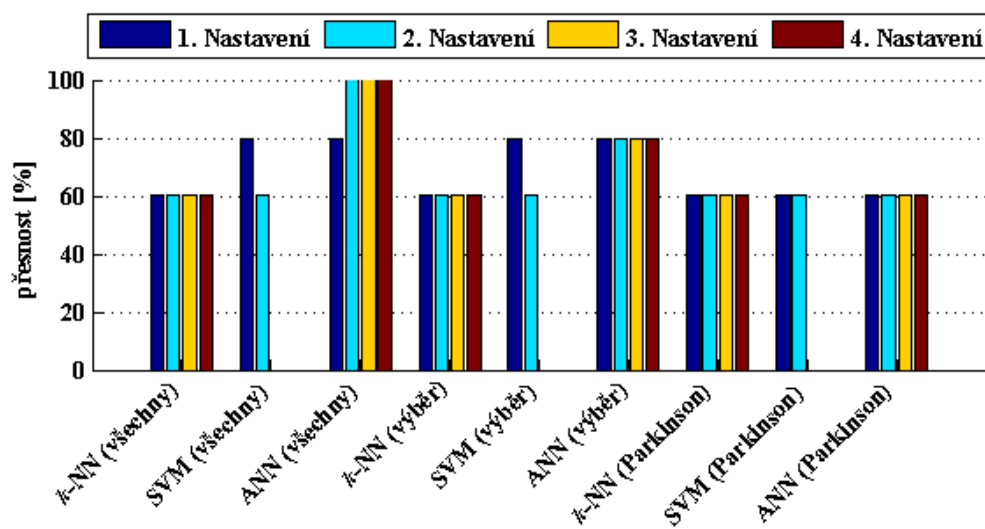
hodnocením klasifikátorů podle validační množiny nahrávek. I v tomto případě byly použity stejné tři množiny příznaků. Použitá validační množina byla z důvodu celkového množství nahrávek velice malá (3 zdravý řečníci a 2 nemocní řečníci). Proto ohodnocení klasifikátorů, může být do jisté míry nepřesné.

Na obr. 6.14 je vidět přesnost klasifikátorů při hodnocení validační množiny nahrávek. Každý z použitých klasifikátorů dokázal přesně vyhodnotit zdravé řečníky to odpovídá 60 % přesnosti a to pro všechny tři množiny příznaků. Jediný klasifikátor ANN v nastavení č. 2–č.4 dokázal vyhodnotit správně celou validační množinu nahrávek a to u množiny příznaků „všechny“. Klasifikátor SVM dokázal správně rozeznat jednoho z nemocných řečníků, ale jen při nastavení s lineárním jádrovým algoritmem a to pro množinu příznaků „všechny“ i „výběr“. Přesnost pro množinu „Parkinson“ dopadla podobně jako u testovací množiny, kde byly ve většině případů rozpoznáni jen zdraví řečníci.

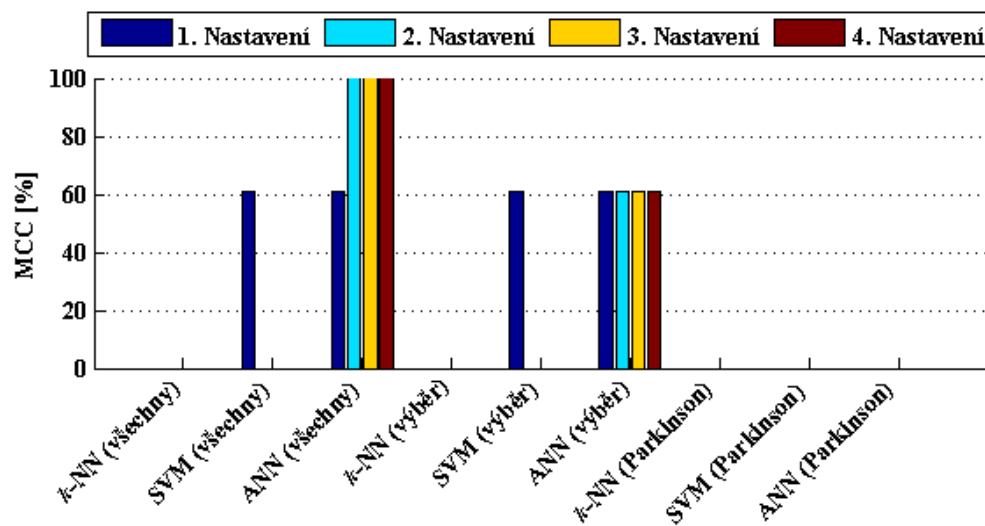
Další obr. 6.15 reprezentuje MCC pro validační množinu nahrávek. Tady se ukázalo, že klasifikátor k -NN je nepoužitelný a naopak klasifikátor ANN dává přesné výsledky pro množinu příznaku „všechny“. Výsledky jednotlivých klasifikátorů odpovídají přesnosti tzn. k -NN má ve všech případech nulovou hodnotu, SVM je použitelný jen s lineárním jádrovým algoritmem a ANN má přesné výsledky pro nastavení č. 2–č. 4 množiny příznaků „všechny“.

6.4 Vyhodnocení příznaků

V předchozí kapitole byly ohodnoceny celé množiny příznaků použité při klasifikaci. V této kapitole budou vyhodnoceny jednotlivé příznaky.



Obr. 6.14: Přesnost klasifikátorů pro validační množinu nahrávek



Obr. 6.15: MCC klasifikátorů pro validační množinu nahrávek

6.4.1 Realizace ohodnocení příznaků v RapidMineru

Ohodnocení příznaků v RapidMineru bylo vytvořeno velice podobně, jak výběr příznaků. Jen se místo operátoru `Optimize Selection` použil operátor `Optimize Weights`. Tento operátor využívá stejně jako u výběru příznaků `X-Validation` (leave-one-out) a ten využívá `Support Vector Machine` s lineárním jádrovým algoritmem.

Po vyzkoušení celého množství nastavení `Optimize Weights`, po kterých vždy byla zkontrolována matice záměn, se nastavení ustálilo na hodnotách:

- velikost populace – 5
- maximální počet generací – 30
- typ selekce – turnajová
- část populace určená pro selekci – 0,6
- mutace – 0,4
- pravděpodobnost křížení – 0,5
- typ křížení – uniformní

Při tomto nastavení udává výstup typu `performance` hodnoty v matici záměn uvedené v tab. 6.17. Přesnost tohoto řešení je 96,2 % a MCC je 88,3 %.

	Pravděpodobná třída		
Skutečná třída		Pozitivní	Negativní
	Pozitivní	41	2
	Negativní	0	9

Tab. 6.17: Matice záměn pro ohodnocení příznaky

Výstupem celého procesu je ohodnocení všech příznaků hodnotou od 0 do 1. Příznaky blíží se 1 jsou nejdůležitější pro klasifikaci a naopak příznaky blíží se k 0 nemají na pro klasifikaci téměř žádný význam. Tab. 6.18 obsahuje 5 nejlepších a 5 nejhorších příznaků pro diagnózu Parkinsonovi choroby.

Nejlepší		Nejhorší	
Příznak	Váha	Příznak	Váha
F_2 max /í/	1	F_1 mean /ú/	0
F_0 max–min /e/	0,998	hnr nh /ó/	0,001
F_2 min /e/	0,997	F_2 mean /ó/	0,002
voicing frac /a/	0,996	F_2 b var /í/	0,006
F_2 var /ú/	0,992	F_2 b mean /í/	0,007

Tab. 6.18: 5 nejlepších a 5 nejhorších příznaků

7 ZÁVĚR

Tato práce se zabývá diagnózou Parkinsonovy choroby pomocí analýzy řečového signálu. V první části práce 1 je vysvětlen samotný princip tvorby řeči. Zejména zapojení jednotlivých orgánů, popis vzniku jednotlivých fonémů a prozodie řeči. Tyto poznatky tvoří základ pro samotné pochopení toho, co se děje u lidí postižených touto nemocí.

V kapitole 2 je pak rozebrána problematika zpracování řečového signálu, který není možné přímo zpracovávat a to zejména z důvodu rychlosti výpočetní kapacity a kapacity úložných médií. Teoreticky se jedná o nekonečně dlouhý nestacionární proces, z kterého lze statistické vlastnosti jen velice těžko odhadnout. Proto je nutné, před vlastním zpracováním tohoto signálu, jej předzpracovat a parametrizovat.

Poslední teoretická část 4 této práce, se zabývá Parkinsonovou chorobou. U lidí s Parkinsonovou chorobou je postiženo celé pohybové ústrojí, které se právě mimo jiné podílí i na tvorbě řeči. Tato skutečnost vede k mnoha odchylkám v řeči, podle kterých je pak možné diagnostikovat Parkinsonovu chorobu. Dále zde jsou rozebrány příznaky sloužící k diagnóze Parkinsonovy choroby. Těchto příznaků existuje celá řada. Většina z nich je založeno na prozodických vlastnostech řeči. V této oblasti dosahují nejlepších výsledků globální příznaky, které jsou počítány z více dílčích příznaků, například FCR, VSA nebo lnVSA.

Všechny tyto teoretické poznatky pak byly uplatněny v poslední části diplomové práce, která se zabývá realizací výpočtu příznaků pomocí programového prostředí MATLAB a jejich následné uplatnění pro trénování, testování klasifikátorů a následné vyhodnocení v programovém prostředí RapidMiner.

Celkově je vypočteno 75 běžně používaných příznaků ve zpracování řeči (F_0 , F_1 , jitter...), dále je vypočteno 5 globálních příznaků pro diagnózu Parkinsonovy choroby (VSA, lnVSA, FCR, VAI, F_{2i}/F_{2u}). Z těchto příznaků jsou vytvořeny tři skupiny příznaků („všechny“, „výběr“ a „Parkinson“) ty jsou pak testovány pomocí tří klasifikátorů (k -NN, SVM a ANN) pro různé nastavení. Jednotlivé klasifikátory z důvodu věrohodnosti (problém přetrénování) byly zkoušeny pro testovací i validační množinou nahrávek.

Pro testovací množinu nahrávek dávala nejlepší výsledky množina příznaků „výběr“, tento výsledek odpovídá i teoretickému rozboru který byl popsán v kapitole 5. Tato množina byla následována množinou „všechny“. Nejhorší výsledky poskytla množina „Parkinson“, podle této množiny nebylo vůbec možné obě třídy od sebe odlišit. To bylo pravděpodobně způsobeno tím, že tyto příznaky jsou navrženy pro anglický jazyk a také velmi omezeným počtem nahrávek.

Zde použitá validační množina byla velice malá z důvodu celkově malého počtu nahrávek. Proto mohou být výsledky zkreslené. V tomto případě se ukázala jako

nejlepší množina příznaků „všechny“. A stejně jako u testovací množiny i zde se množina příznaků „Parkinson“ ukázala být nepoužitelnou.

Nakonec byl vytvořen systém pro zhodnocení všech 80 vypočtených příznaků. Jako vůbec nejlepším příznakem se ukázala být maximální hodnota F_2 samohlásky /í/ a naopak nejhorším byla střední hodnota F_1 samohlásky /ú/.

Bohužel všechny dosažené výsledky jsou ovlivněny určitou chybou z důvodu malé databáze nahrávek.

V navazující práci lze využít všech zmíněných teoretických a praktických poznatků. Nejdůležitějším aspektem práce bude rozšíření databáze nahrávek. Další možností je vytvořit větší množství příznaků a použít nové klasifikátory.

LITERATURA

- [1] BURGET, Radim. Dolování znalostí z báze dat. *Teoretická informatika* [online]. Brno: Ústav telekomunikací, Vysoké učení technické v Brně, 3.11.2010, [cit. 2011-05-18]. Dostupné z WWW: <<https://www.vutbr.cz/elearning/mod/resource/view.php?id=114594>>.
- [2] BURGET, Radim; KARÁSEK, Jan. Optimalizace. *Teoretická informatika* [online]. Brno: Ústav telekomunikací, Vysoké učení technické v Brně, 25.11.2010, [cit. 2011-05-18]. Dostupné z WWW: <<https://www.vutbr.cz/elearning/mod/resource/view.php?id=116467>>.
- [3] CANTINIAUX, Stéphanie, et al. Comparative analysis of gait and speech in Parkinson's disease: hypokinetic or dysrhythmic disorders?. *J Neural Neurosurg Psychiatry*. 2010, 81, s. 177–184.
- [4] ČERNOCKÝ, Jan. *Zpracování řečových signálů*. Brno: Vysoké učení technické v Brně, 2006. 128 s.
- [5] FISCHER, Emily; GOBERMAN, Alexander M. Voice onset in Parkinson disease. *Journal of Communication Disorders*. 2010, 43, s. 21–34.
- [6] GOBERMAN, Alexander M. Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson Disease. *Medical science*. 2005, 11, s. 109–116.
- [7] GOBERMAN, Alexander M.; COELHO, Carl A.; ROBB, Michael P. Phonotary chracteristics of Parkinsonian speech before and after morning medication: the ON and OFF states. *Journal of Communication Disorders*. 2002, 35, s. 217–239.
- [8] GOBERMAN, Alexander M.; COELHO, Carl A.; ROBB, Michael P. Prosodic Characteristics of Parkinsonian Speech: The Effect of Levodopa-Based Medication. *Journal of Medical Speech-Language Pathology*. 2005, 13, s. 51–68.
- [9] GOBERMAN, Alexander M.; MCMILLAN, Jessica. Relative Speech Timing in Parkinson Disease. *Issues in Communication Science and Disorders*. 2005, 32, s. 22–29.
- [10] HO, Aileen K.; BRADSHAW, John L.; IANSEK, Robert. For Better or Worse: The Effect of Levodopa on Speech in Parkinson's Disease. *Movement Disorders*. 2008, 23, s. 574–580.

- [11] KRČMOVÁ, Marie. *Fonetika a fonologie* [online]. Brno: Filozofická fakulta Masarykovy univerzity, 2008 [cit. 2010-11-08]. Dostupné z WWW: <<http://is.muni.cz/do/1499/el/estud/ff/js08/fonetika/ucebnice/index.html>>.
- [12] MOLNÁR, Karol. Úvod do problematiky umělých neuronových sítí. *Elektrorevue* [online]. 22.2.2000, 2000/13, [cit. 2011-05-18]. Dostupné z WWW: <<http://www.elektrorevue.cz/clanky/00013/index.html>>.
- [13] PSUTKA, Josef, et al. *Mluvíme s počítačem česky*. Vid. 1. Praha: Academia, 2006. 752 s. ISBN 80-200-1309-1.
- [14] RABINER, Lawrence ; JUANG, Biing-Hwang. *FUNDAMENTALS OF SPEECH RECOGNITION*. New Jersey: Prentice Hall PTR, 1993. 497 s. ISBN 0-13-285826-6.
- [15] RABINER, L.R.; SCHAFER, R.W. *Digital Processing of Speech Signals*. New Jersey: Prentice Hall PTR, 1978. 509 s. ISBN 0-13-243603-1.
- [16] REDDY, Raj. *SPOKEN LANGUAGE PROCESING: A Guide to Theory, Algorithm, and System Development*. New Jersey: Prentice Hall PTR, 2001. 931 s. ISBN 0-13-022616-5.
- [17] SAPIR, Shimon, at al. Formant Centralization Ratio (FCR): A proposal for a new acoustic measure of dysarthric speech. *J Neural Transm.* 2010, 114, s. 1–20. Dostupný také z WWW: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2821466/>>.
- [18] SAPIR, Shimon; RAMING, Lorraine; FOX, Cynthia. Speech and swallowing disorders in Parkinson disease. *Current Opinion in Otolaryngology & Head and Neck Surgery*. 2008, 16, s. 205–210.
- [19] SKODDA, Sabine; RINSCHKE, Heiko ; SCHLEGEL, Uwe. Progression of Dysprosody in Parkinson's Disease Over Time—A Longitudinal Study. *Movement Disorders*. 2009, 24, s. 716–722.
- [20] SKODDA, Sabine; VISSER, Wenke ; SCHLEGEL, Uwe. Short- and long-term dopaminergic effects on dysarthria in early Parkinson's disease. *J Neural Transm.* 2010, 117, s. 197–205.
- [21] SMÉKAL, Zdeněk. *Číslíkové zpracování řeči*. Brno: Vysoké učení technické v Brně, 2009. 122 s.

- [22] ZAMIŠKOVÁ, Gabriela, at al. Poruchy řeči u Parkinsonovy nemoci. *Neurologie v praxi* [online]. 2010, č. 2, [cit. 2010-11-08]. Dostupný z WWW: <<http://www.neurologiepropraxi.cz/pdfs/neu/2010/02/10.pdf>>.
- [23] Diskriminační analýza. In *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikipedia Foundation, 18.3.2010, last modified on 5.7.2010 [cit. 2010-11-20]. Dostupné z WWW: <http://cs.wikipedia.org/wiki/Diskriminační_analýza>.
- [24] Quadratic programming. In *Wikipedia: the free encyclopedia* [online]. St. Petersburg (Florida): Wikipedia Foundation, 24.7.2006, last modified on 13.5.2010 [cit. 2010-11-20]. Dostupné z WWW: <http://en.wikipedia.org/wiki/Quadratic_programming>.
- [25] ŽIŽKA, Jan. Support Vector Machine. *Strojové učení* [online]. Brno: Fakulta informatiky, Masarykova univerzita, 9.12.2004, [cit. 2011-05-18]. Dostupné z WWW: <http://is.muni.cz/el/1433/podzim2006/PA034/09_SVM.pdf?fakulta=1433;obdobi=3523;kod=PA034>.

SEZNAM SYMBOLŮ, VELIČIN A ZKRATEK

$\mathbb{R}$	množina reálných čísel
$\mathbb{N}$	množina celých čísel
ANN	umělé neuronové sítě – Artificial Neural Network
ED_{xx}	Euklidova vzdálenost – Euclidean Distance
F_0 VR	rozsah rozptylu F_0 – Variation Range
rel F_0 VR	relativní kolísání F_0 – Relative Variation Range
rel F_0 SD	relativní směrodatná odchylka F_0 – Relative Standard Deviation
FCR	centralizační poměr formantů – Formant Centralization Ratio
GMM	smíšené Gaussovy modely – Gaussian Mixture Models
ISD	doba mezi dvěma pauzami – Interpause-Speech Duration
k -NN	k nejbližších sousedů – k Nearest-Neighbor
LnVSA	logaritmická vokální oblast hlasového traktu – Logarithmic Vowel Space Area
LPC	lineární predikční kódování – Linear Predictive Coding
PPV	negativní předpovídající hodnota – Negative Predictive Value
NSR	čistá rychlost řeči – Net Speech Rate
MCC	Matthewův korelační koeficient – Matthews Correlation Coefficient
MFCC	melovské keprální koeficienty – Mel-frequency Cepstral Coefficients
PPV	pozitivní předpovídající hodnota – Positive Predictive Value
PLP	percepční lineárně predikční koeficienty – Perceptual Linear Predictive Coefficients
SPIR	počet mezislovních pauz za minutu – Speech Index of Rhythmicity
SVM	algoritmy podpůrných vektorů – Support Vector Machines
TSR	celková rychlost řeči – Total Speech Rate
VAI	artikulační index samohlásky – Vowel Articulation Index

VOT	doba navození řeči – Voice Onset Time
VSA	vokální oblast hlasového traktu – Vowel Space Area
$A(z)$	přenosová funkce analyzovaného filtru
a_p	váhové koeficienty
c	normalizační konstanta
$e[n]$	chyba predikce
F_0	základní kmitočet [Hz]
F_i	kmitočet formantů [Hz]
f	nelineární funkce
f_b	kmitočet ve stupnici bark [bark]
f_m	kmitočet ve stupnici mel [mel]
$f'(i)$	proporcionální selekce
$H(z)$	přenosová funkce
$p(i)$	lineární uspořádání
$S[f]$	spektrum řečového signálu
$\mathbf{S_B}$	matice pro rozptyl mezi třídami
$\mathbf{S_T}$	meziprvková autokovariační matice
$\mathbf{S_W}$	matice pro rozptyl uvnitř tříd
s	selekční intenzita
$s[n]$	řečový signál
$s'[n]$	modifikovaný řečový signál
$\tilde{s}(n)$	predikovaný signál
$W[f]$	spektrum okénkové funkce
$w[n]$	okénková funkce
w_i	váha i -tého vstupu

$X[f]$	výsledné spektrum
x_i	hodnota i -tého vstupu
z	komplexní proměná Z-transformace
β^t	nejhorší fitness funkce v t -generaci
Γ_P	Toeplitzova matice
$\gamma_{xx}[n]$	autokorelační koeficienty
μ_s	střední hodnota signálu
π	Ludolfovo číslo
ϑ	prahová hodnota
σ_f	sigma škálování

SEZNAM PŘÍLOH

A	Obsah přiloženého DVD	67
A.1	obsah adresáře DiplomovaPrace	67
A.1.1	obsah adresáře Matlab	67
A.1.2	obsah adresáře RapidMiner	67
B	Přehled cvičení	68
C	Příznaky	72
C.1	Všechny vypočítané příznaky	72
C.2	Příznaky použité pro klasifikaci	75

A OBSAH PŘILOŽENÉHO DVD

Přiložené DVD obsahuje elektronickou verzi diplomové práce ve formátu `.pdf`, konzolový program Praat, veškeré matlab funkce potřebné k výpočtu příznaků, RapidMiner repozitář se všemi importovanými hodnotami, vytvořené klasifikátory, vytvořený výběr příznaků a vytvořené váhování příznaků.

Nahrávky nejsou zatím volně šiřitelné a tudíž nejsou na DVD. V případě zájmu je možné tyto nahrávky vyžádat na Ústavu telekomunikací, FEKT, VUT v Brně.

Výpočet příznaků lze spustit pomocí grafického prostředí, které se zavolá v prostředí MATLAB pomocí příkazu `Parkinson`.

Výpočty v RapidMineru lze spustit po otevření příslušného projektu.

Seznam všech důležitých souborů a složek je uveden níže.

A.1 obsah adresáře DiplomovaPrace

Matlab

RapidMiner

DiplomovaPrace.pdf

A.1.1 obsah adresáře Matlab

Funkce – složka obsahující veškeré Matlab funkce, konzolový program Praat a soubor `priznaky.xls` obsahující vypočtené příznaky.

A.1.2 obsah adresáře RapidMiner

Databaze – složka obsahující veškeré naimportované příznaky.

Finalni – složka obsahující jednotlivá řešení.

vyber.wgt – vybrané příznaky pro diagnózu Parkinsonovy choroby.

B PŘEHLED CVIČENÍ

č. cv.	účel cvičení	náplň cvičení
9.4	srozumitelnost četby textu	Přečtěte nahlas a pozorně o čem je tento text. TEXT čte pacient („ <i>I na tom, že člověk si opatří psa. . .</i> “)
9.2	srozumitelnost předříkávaných vět	Opakujte po mě věty. 9.2.1 Do čtvrt hodiny tam byla smršť. 9.2.2 Prohovořte to s ním dopodrobna. 9.2.3 Při ústupu pluku duní bubny. 9.2.4 Kuchařští učenci nejsou jak zlatníci. 9.2.5 Celý večer se učí počítat
8.1	udržení rytmu v rytmičtém celku	Přečtěte báseň s přednesem. („ <i>Chcete vidět. . .</i> “) TEXT čte pacient báseň
8.2	základní větné intonační vzorce	Přečtěte věty, červeně máte napsáno jakým způsobem je máte přečíst. TEXT čte pacient intonace vět („ <i>Prostřete k obědu?</i> “) 8.2.1 Prostřete k obědu? 8.2.2 Prostřete k obědu! 8.2.3 Prostřete k obědu.
8.3	přemísťování kontrastního přízvuku	Přečtěte důrazně zvýrazněná slova. TEXT čte pacient intonace vět („ <i>Z okapu teče špinavá rezavá voda.</i> “) 8.3.1 Z OKAPU teče špinavá rezavá voda 8.3.2 Z okapu teče ŠPINAVÁ rezavá voda. 8.3.3 Z okapu teče špinavá rezavá VODA.
8.4	intonální variabilita	Přečtěte věty, červeně máte napsáno jakým způsobem je máte přečíst. TEXT čte pacient („ <i>Ted' musíš být chvíli trpělivý. . .</i> “) 8.4.1 Ted' musíš být chvíli trpělivý, než to dokončíme. 8.4.2 Tak dáš mi už konečně pokoj! 8.4.3 Už mě to nebaví, dej mi už konečně pokoj! 8.4.4 Tak co, jak to dopadlo?

č. cv.	účel cvičení	náplň cvičení
9.1	srozumitelnost předřikávaných slov	Opakujte po mě slova: 9.1.1 rychlonožka 9.1.2 marnotratný 9.1.3 horolezectví 9.1.4 stříbrotepec 9.1.5 železobetonový 9.1.6 zákonodárce 9.1.7 horkovzdušný 9.1.8 strastiplná 9.1.9 záviděníhodný 9.1.10 československý 9.1.11 bíložravec 9.1.12 pětikoruna 9.1.13 dalekohled 9.1.14 teploměr 9.1.15 krátkozrakost 9.1.16 granthotel
7.1	přesnost opakování samohlásek / artikulace	Opakujte po mě, dodržujte pauzy mezi jednotlivými hláskami: 1. nejprve všechny krátké samohlásky 7.1.1_a a 7.1.1_e e 7.1.1_i i 7.1.1_o o 7.1.1_u u 2. dlouhé samohlásky 7.1.2_a á 7.1.2_e é 7.1.2_i í 7.1.2_o ó 7.1.2_u ú

č. cv.	účel cvičení	náplň cvičení
7.1	přesnost opakování samohlásek / artikulace	Opakujte po mě, dodržujte pauzy mezi jednotlivými hláskami: 3. dlouhé samohlásky co nejhlasitěji 7.1.3_a á 7.1.3_e é 7.1.3_i í 7.1.3_o ó 7.1.3_u ú 4. dlouhé samohlásky co nejtíšeji (ale ne šepot) 7.1.4_a á 7.1.4_e é 7.1.4_i í 7.1.4_o ó 7.1.4_u ú
7.4	diadochokineze s fonací	Říkejte rychle a srozumitelně. 1 (méně než 6x za 3s nebo částečně nepřesné) 7.4.1 ba-ba-ba... 7.4.2 p-t-k... 7.4.3 pa-ta-ka...
2.3	posouvání mandibulou	Posunujte bradou střídavě 3x vpravo a 3x vlevo. 2 (posune 1,5cm)
2.4	kroužení mandibulou	Pohybujte bradou dokola
5.1	výdrž expirace při delší fonaci hlásky	Vydržte co nejdéle prodloužené mmm se zavřenými ústy. 1 (méně než 15s) zavřená ústa
5.2	výdrž expirace při delší fonaci hlásky	Vydržte co nejdéle prodloužené 1. ííí 2. ááá s otevřenými ústy. 1 (5-10s) pootevřená ústa 5.2.1 ííí 5.2.2 ááá
5.3	synchro. respirace s fonací	Řekněte prodloužené fííííí. 2(5s fonace a je možno rozlišit neznělou a znělou hlásku) 5.3 fííí 5.3.1 rýýý

č. cv.	účel cvičení	náplň cvičení
9.2.1	sledování změny tempa a prozodie	Opakujte po mně věty. 9.2.1.1 Do čtvrt hodiny tam byla smršť. 9.2.1.2 Prohovořte to s ním dopodrobna. 9.2.1.3 Při ústupu pluku duní bubny. 9.2.1.4 Kuchařští učni nejsou jak zlatničtí. 9.2.1.5 Celý večer se učí sčítat.
9.4.1	sledování změny tempa a prozodie	Přečtěte nahlas a pozorně o čem je tento text. TEXT čte pacient („ <i>I na tom, že člověk si opatří psa. . .</i> “)
10	pro výpočet VOT	Opakujte po mně slova. 10.1.1 bába 10.2.1 dikobraz 10.3.1 guma 10.4.1 potok 10.5.1 téma 10.6.1 kýbl 10.1.2 bába 10.2.2 dikobraz 10.3.2 guma 10.4.2 potok 10.5.2 téma 10.6.2 kýbl 10.1.3 bába 10.2.3 dikobraz 10.3.3 guma 10.4.3 potok 10.5.3 téma 10.6.3 kýbl 10.1.4 bába 10.2.4 dikobraz 10.3.4 guma 10.4.4 potok 10.5.4 téma 10.6.4 kýbl

C PŘÍZNAKY

C.1 Všechny vypočítané příznaky

příznaky	jednotky	popis příznaků
F_0 median	[Hz]	medián F_0
F_0 mean	[Hz]	střední hodnota F_0
F_0 std	[Hz]	směrodatná odchylka F_0
F_0 min	[Hz]	minimální hodnota F_0
F_0 max	[Hz]	maximální hodnota F_0
F_0 max–min	[Hz]	rozdíl maximální a minimální hodnoty F_0
rel F_0 VR	[%]	relativní kolísání F_0
rel F_0 SD	[%]	relativní směrodatná odchylka F_0
voicing frac	[%]	shluk neznělých úseku
voicing nbreaks	[-]	počet pauz
voicing degree	[%]	počet pauz v %
jitter local	[%]	lokální kolísání periody
jitter localabs	[s]	absolutní hodnota lokální kolísání periody
jitter rap	[%]	průměrný rozdíl mezi periodou a průměrem dvou sousedních period
jitter ppq5	[%]	průměrný rozdíl mezi periodou a průměrem čtyř sousedních period
jitter ddp	[%]	průměrný rozdíl mezi za sebou jdoucích rozdílů period
shimmer local	[%]	lokální kolísání amplitudy
shimmer localdb	[dB]	lokální kolísání amplitudy v dB
shimmer apq3	[%]	průměrný rozdíl mezi amplitudou a průměrem dvou sousedních amplitud
shimmer apq5	[%]	průměrný rozdíl mezi amplitudou a průměrem čtyř sousedních amplitud
shimmer apq11	[%]	průměrný rozdíl mezi amplitudou a průměrem deseti sousedních amplitud
shimmer ddp	[%]	průměrný rozdíl mezi za sebou jdoucích rozdílů amplitud

příznaky	jednotky	popis příznaků
hnr ut	[—]	střední hodnota autokorace
hnr nh	[—]	střední hodnota poměru signál–šum
hnr hn	[dB]	střední hodnota poměru signál–šum v dB
F_1 mean	[Hz]	střední hodnota F_1
F_1 var	[Hz]	rozptyl F_1
F_1 std	[Hz]	směrodatná odchylka F_1
F_1 max	[Hz]	maximální hodnota F_1
F_1 min	[Hz]	minimální hodnota F_1
F_1 med	[Hz]	medián F_1
F_1 max–min	[Hz]	rozdíl maximální a minimální hodnoty F_1
F_1 b mean	[Hz]	střední hodnota šířky pásma F_1
F_1 b var	[Hz]	rozptyl šířky pásma F_1
F_1 b std	[Hz]	směrodatná odchylka šířky pásma F_1
F_1 b max	[Hz]	maximální hodnota šířky pásma F_1
F_1 b min	[Hz]	minimální hodnota šířky pásma F_1
F_1 b med	[Hz]	medián šířky pásma F_1
F_1 b max–min	[Hz]	rozdíl maximální a minimální hodnoty šířky pásma F_1
F_2 mean	[Hz]	střední hodnota F_2
F_2 var	[Hz]	rozptyl F_2
F_2 std	[Hz]	směrodatná odchylka F_2
F_2 max	[Hz]	maximální hodnota F_2
F_2 min	[Hz]	minimální hodnota F_2
F_2 med	[Hz]	medián F_2
F_2 max–min	[Hz]	rozdíl maximální a minimální hodnoty F_2
F_2 b mean	[Hz]	střední hodnota šířky pásma F_2
F_2 b var	[Hz]	rozptyl šířky pásma F_2
F_2 b std	[Hz]	směrodatná odchylka šířky pásma F_2
F_2 b max	[Hz]	maximální hodnota šířky pásma F_2
F_2 b min	[Hz]	minimální hodnota šířky pásma F_2
F_2 b med	[Hz]	medián šířky pásma F_2
F_2 b max–min	[Hz]	rozdíl maximální a minimální hodnoty šířky pásma F_2
F_3 mean	[Hz]	střední hodnota šířky pásma F_3
F_3 var	[Hz]	rozptyl šířky pásma F_3

příznaky	jednotky	popis příznaků
F_3 std	[Hz]	směrodatná odchylka šířky pásma F_3
F_3 max	[Hz]	maximální hodnota šířky pásma F_3
F_3 min	[Hz]	minimální hodnota šířky pásma F_3
F_3 med	[Hz]	medián šířky pásma F_3
F_3 max–min	[Hz]	rozdíl maximální a minimální hodnoty šířky pásma F_3
F_3 mean	[Hz]	střední hodnota šířky pásma F_3
F_3 var	[Hz]	rozptyl šířky pásma F_3
F_3 std	[Hz]	směrodatná odchylka šířky pásma F_3
F_3 max	[Hz]	maximální hodnota šířky pásma F_3
F_3 min	[Hz]	minimální hodnota šířky pásma F_3
F_3 med	[Hz]	medián šířky pásma F_3
F_3 max–min	[Hz]	rozdíl maximální a minimální hodnoty šířky pásma F_3
E max	[—]	maximální hodnota energie
E min	[—]	minimální hodnota energie
E mean	[—]	střední hodnota energie
E std	[—]	směrodatná odchylka energie
E median	[—]	medián energie
E max–min	[—]	rozdíl maximální a minimální hodnoty energie
TKEO	[—]	Teager-Kaisrův operátor energie
DelkaLab	[s]	délka jednotlivého cvičení v sekundách
FCR	[—]	centralizační poměr formantů
VSA	[—]	vokální oblast hlasového traktu
lnVSA	[—]	ln vokální oblast hlasového traktu
VAI	[—]	artikulační index samohlásky
F_{2i}/F_{2u}	[—]	podíl F_{2i} a F_{2u}

C.2 Příznaky použité pro klasifikaci

příznaky	jednotky	popis příznaků
F_0 median	[Hz]	medián F_0
F_0 mean	[Hz]	střední hodnota F_0
F_0 std	[Hz]	směrodatná odchylka F_0
F_0 min	[Hz]	minimální hodnota F_0
F_0 max	[Hz]	maximální hodnota F_0
F_0 max–min	[Hz]	rozdíl maximální a minimální hodnoty F_0
rel F_0 VR	[%]	kolísání směrodatné odchylky F_0
rel F_0 SD	[%]	kolísání rozdílu maximální a minimální hodnoty F_0
voicing frac	[%]	shluk neznělých úseku
voicing nbreaks	[-]	počet pauz
voicing degree	[%]	počet pauz v %
jitter local	[%]	lokální kolísání periody
jitter localabs	[s]	absolutní hodnota lokální kolísání periody
jitter rap	[%]	průměrný rozdíl mezi periodou a průměrem dvou sousedních period
jitter ppq5	[%]	průměrný rozdíl mezi periodou a průměrem čtyř sousedních period
jitter ddp	[%]	průměrný rozdíl mezi za sebou jdoucích rozdílů period
hnr ut	[—]	střední hodnota autokorelace
hnr nh	[—]	střední hodnota poměru signál–šum
hnr hn	[dB]	střední hodnota poměru signál–šum v dB
F_1 mean	[Hz]	střední hodnota F_1
F_1 var	[Hz]	rozptyl F_1
F_1 std	[Hz]	směrodatná odchylka F_1
F_1 max	[Hz]	maximální hodnota F_1
F_1 min	[Hz]	minimální hodnota F_1
F_1 med	[Hz]	medián F_1
F_1 max–min	[Hz]	rozdíl maximální a minimální hodnoty F_1
F_1 b mean	[Hz]	střední hodnota šířky pásma F_1
F_1 b var	[Hz]	rozptyl šířky pásma F_1
F_1 b std	[Hz]	směrodatná odchylka šířky pásma F_1

příznaky	jednotky	popis příznaků
F_1 b max	[Hz]	maximální hodnota šířky pásma F_1
F_1 b min	[Hz]	minimální hodnota šířky pásma F_1
F_1 b med	[Hz]	medián šířky pásma F_1
F_1 b max–min	[Hz]	rozdíl maximální a minimální hodnoty šířky pásma F_1
F_2 mean	[Hz]	střední hodnota F_2
F_2 var	[Hz]	rozptyl F_2
F_2 std	[Hz]	směrodatná odchylka F_2
F_2 max	[Hz]	maximální hodnota F_2
F_2 min	[Hz]	minimální hodnota F_2
F_2 med	[Hz]	medián F_2
F_2 max–min	[Hz]	rozdíl maximální a minimální hodnoty F_2
F_2 b mean	[Hz]	střední hodnota šířky pásma F_2
F_2 b var	[Hz]	rozptyl šířky pásma F_2
F_2 b std	[Hz]	směrodatná odchylka šířky pásma F_2
F_2 b max	[Hz]	maximální hodnota šířky pásma F_2
F_2 b min	[Hz]	minimální hodnota šířky pásma F_2
F_2 b med	[Hz]	medián šířky pásma F_2
F_2 b max–min	[Hz]	rozdíl maximální a minimální hodnoty šířky pásma F_2