



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV TELEKOMUNIKACÍ

DEPARTMENT OF TELECOMMUNICATIONS

NÁVRH PARAMETRŮ KVANTIFIKUJÍCÍ PORUCHY RESPIRACE U PACIENTŮ S PARKINSONOVOU NEMOCÍ

DEVELOPMENT OF FEATURES QUANTIFYING RESPIRATORY DYSFUNCTIONS IN
PARKINSON'S DISEASE PATIENTS

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Dominik Cvetler

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Daniel Kováč

BRNO 2023

Diplomová práce

magisterský navazující studijní program **Audio inženýrství**
specializace Zvuková produkce a nahrávání
Ústav telekomunikací

Student: Bc. Dominik Cvetler
Ročník: 2

ID: 195798
Akademický rok: 2022/23

NÁZEV TÉMATU:

Návrh parametrů kvantifikující poruchy respirace u pacientů s Parkinsonovou nemocí

POKyny PRO VYPRACOVÁNÍ:

Pacienti s Parkinsonovou nemocí (PN) se potýkají s obtížemi dýchání způsobenými třesem, ztuhlostí dýchacích svalů nebo bradykinezií. V řeči se pak tyto obtíže manifestují sníženou hlasitostí (hypofónie) nebo její variací v čase. Cílem práce je kvantifikovat tyto poruchy a porovnat nově navržené akustické parametry řeči pacientů s PN s parametry zdravých kontrol. Nejprve bude sestaven algoritmus pro automatizovanou detekci nádechů v řečovém signálu. Dále proběhne návrh a implementace parametrů na základě aktuálních poznatků vědy o respiračních dysfunkcích pacientů s PN. V posledním kroku budou parametry statisticky zpracovány.

DOPORUČENÁ LITERATURA:

- [1] D'ARRIGO, Andrea, Stefano FLORO, Francesca BARTESAGHI, Chiara CASELLATO, Giuseppe Francesco SFERRAZZA PAPA, Stefano CENTANNI, Alberto PRIORI a Tommaso BOCCI. Respiratory dysfunction in Parkinson's disease: a narrative review. ERJ Open Research [online]. 2020, 6(4), /erjor/6/4/00165-2020.atom [cit. 2022-05-11]. ISSN 2312-0541. Dostupné z: doi:10.1183/23120541.00165-2020
- [2] HLAVNIČKA, Jan, Roman ČMEJLA, Tereza TYKALOVÁ, Karel ŠONKA, Evžen RŮŽIČKA a Jan RUSZ. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. Scientific Reports [online]. 2017, 7(1) [cit. 2022-05-11]. ISSN 2045-2322. Dostupné z: doi:10.1038/s41598-017-00047-5

Termín zadání: 6.2.2023

Termín odevzdání: 19.5.2023

Vedoucí práce: Ing. Daniel Kováč

doc. Ing. Jiří Schimmel, Ph.D.
předseda rady studijního programu

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

V úvodu práce je krátce popsána Parkinsonova nemoc a hypokinetická dysartrie, která má negativní vliv na tvorbu řeči a způsobuje problémy s respirací během řeči u nemocných pacientů. Cílem práce je vytvoření algoritmu pro automatizovanou detekci nádechů a návrh parametrů pro kvantifikaci respiračních poruch u pacientů s Parkinsonovou nemocí. V prostředí MATLAB byly zpracovány nahrávky zkoumaných subjektů a vytvořen algoritmus pro detekci nádechů, u kterého byla použita metoda logistické regrese. Na základě predikovaných nádechů byly z nahrávek extrahovány navržené parametry, které byly následně statisticky analyzovány a porovnávány v rámci zdravých kontrol a pacientů s Parkinsonovou nemocí. Využitím modelu strojového učení bylo možno do jisté míry predikovat klinická data pacientů z navržených parametrů. Průměrná přesnost modelu pro predikci nádechů byla 0,85. Ze 14 navržených parametrů bylo 6 parametrů vhodných pro kvantifikaci respiračních poruch spojených s hypokinetickou dysartrií. Výsledkem práce je funkční algoritmus pro automatizovanou detekci nádechů v řečovém signálu a navržené parametry, které by mohly být užitečné pro kvantifikaci respiračních poruch u pacientů s Parkinsonovou nemocí.

Klíčová slova

Hypokinetická dysartrie, Parkinsonova nemoc, respirace, detekce nádechů, detektor řečové aktivity, statistická analýza, strojové učení

Abstract

In the beginning of the thesis, Parkinson's disease and hypokinetic dysarthria are briefly described, which have a negative effect on speech production and cause breathing problems during speech in sick patients. The aim of the thesis is to create an algorithm for automated detection of breaths and the design of parameters for the quantification of respiratory disorders in patients with Parkinson's disease. In the MATLAB environment, the recordings of the researched subjects were processed and an algorithm was created for the detection of breaths, which used the logistic regression method. Based on the predicted breaths, proposed parameters were extracted from the recordings, which were then statistically analyzed and compared in healthy controls and patients with Parkinson's disease. By using a machine learning model, it was possible to predict the clinical data of patients from the proposed parameters to a certain extent. The average accuracy of the model for predicting puffs was 0.85. Of the 14 proposed parameters, 6 were suitable for quantifying respiratory disorders associated with hypokinetic dysarthria. The result of the work is a functional algorithm for the automated detection of breaths in the speech signal and proposed parameters that could be useful for the quantification of respiratory disorders in patients with Parkinson's disease.

Keywords

Hypokinetic dysarthria, Parkinson's disease, respiration, breath detection, voice activity detector, statistical analysis, machine learning

Bibliografická citace

CVETLER, Dominik. *Návrh parametrů kvantifikující poruchy respirace u pacientů s Parkinsonovou nemocí*. Brno, 2023. Dostupné také z: <https://www.vut.cz/studenti/zav-prace/detail/151160>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav telekomunikací. Vedoucí práce Daniel Kováč.

Prohlášení autora o původnosti díla

Jméno a příjmení studenta: Dominik Cvetler

VUT ID studenta: 195798

Typ práce: Diplomová práce

Akademický rok: 2022/23

Téma závěrečné práce: NÁVRH PARAMETRŮ KVANTIFIKUJÍCÍ PORUCHY RESPIRACE U PACIENTŮ S PARKINSONOVOU NEMOCÍ

Prohlašuji, že svou závěrečnou práci jsem vypracoval samostatně pod vedením vedoucí/ho závěrečné 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 závěrečné práce dále prohlašuji, že v souvislosti s vytvořením této závěrečné 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 části druhé, hlavy VI. díl 4 Trestního zákoníku č. 40/2009 Sb.

V Brně dne: 19. května 2023

podpis autora

Poděkování

Děkuji vedoucímu práce, kterým byl Ing. Daniel Kováč, za kvalitní vedení práce, za odborný dohled a cenné rady v průběhu celého akademického roku. Děkuji také rodině a nejbližším za podporu během tvorby diplomové práce a po čas celého studia.

V Brně dne: 19. května 2023

podpis autora

Obsah

SEZNAM OBRÁZKŮ	9
SEZNAM TABULEK	10
SEZNAM ZKRATEK	11
ÚVOD	12
1. PARKINSONOVA NEMOC	14
1.1 HYPOKINETICKÁ DYSARTRIE.....	14
1.1.1 Oblast fonace	15
1.1.2 Oblast prozódie.....	15
1.1.3 Oblast artikulace.....	15
1.1.4 Oblast respirace.....	15
2. SOUČASNÝ STAV POZNÁNÍ (STATE OF THE ART)	17
2.1 DETEKCE NÁDECHŮ.....	17
2.2 KVANTIFIKACE PORUCH.....	19
3. METODIKA PRÁCE	21
3.1 DATABÁZE ZKOUMANÝCH SUBJEKTŮ.....	21
3.2 ANOTACE NAHRÁVEK.....	23
3.3 PŘEDZPRACOVÁNÍ ŘEČOVÉHO SIGNÁLU	24
3.3.1 Vzorkování a normalizace signálu	24
3.3.2 Preemfázový filtr	25
3.3.3 Segmentace signálu.....	26
3.4 AUTOMATIZOVANÁ DETEKCE NÁDECHŮ	26
3.4.1 Volba parametrů řeči	27
3.4.2 Vytvoření datasetu	28
3.4.3 Logistická regrese.....	29
3.4.4 Získání časových údajů.....	31
3.5 PARAMETRIZACE NAHRÁVEK	31
3.5.1 Navržené parametry pro kvantifikaci respiračních poruch:	31
3.5.2 Statistická analýza	32
3.5.3 Strojové učení.....	34
4. VÝSLEDKY	35
4.1 AUTOMATIZOVANÁ DETEKCE NÁDECHŮ	35
4.2 STATISTICKÁ ANALÝZA	40
4.3 STROJOVÉ UČENÍ	46
5. VYHODNOCENÍ VÝSLEDKŮ, DISKUSE	48
5.1 AUTOMATIZOVANÁ DETEKCE NÁDECHŮ	48
5.2 STATISTICKÁ ANALÝZA	50
5.3 POROVNÁNÍ DETEKOVANÝCH NÁDECHŮ S RUČNÍ ANOTACÍ.....	52
5.4 LIMITACE PRÁCE	53
6. ZÁVĚR	54

LITERATURA.....	55
SEZNAM PŘÍLOH.....	57
PŘÍLOHA 1 – OBSAH DISKU	57

SEZNAM OBRÁZKŮ

Obrázek 1: vliv HD na oblasti tvorby řeči	14
Obrázek 2: Rozložení věku zkoumaných subjektů	21
Obrázek 3: prostředí programu Wavesurfer.....	23
Obrázek 4: ukázka .lab souboru s časy nádechů	24
Obrázek 5: princip vzorkování.....	25
Obrázek 6: ROC křivka (subjekt K1003)	30
Obrázek 7: upravené hodnoty parametru n_breath po eliminaci vlivu věku a pohlaví	33
Obrázek 8: Průměrná ROC křivka vytvořeného modelu	38
Obrázek 9: Ukázka lab souboru s predikovanými časy pro subjekt P1088.	40
Obrázek 10: Rozložení parametrů s nejvyšším počtem pacientů s PN ležící mimo normy zdravých kontrol	44
Obrázek 11: Tři nejvýznamnější parametry pro predikci klinického skóre pro jednotlivé oblasti	47

SEZNAM TABULEK

Tabulka 1: Demografická data zkoumaných subjektů	21
Tabulka 2: Ukázka tabulky klinických dat pacientů	22
Tabulka 3: Průměrná klinická data	22
Tabulka 4: Tabulka výsledků křížové validace pro různé parametry řeči.....	35
Tabulka 5: Ukázka tabulky vypočítaných parametrů řeči.....	36
Tabulka 6: Ukázka tabulky výsledků predikce nádechů a metrik pro vyhodnocení modelu	37
Tabulka 7: Výsledky vyhodnocení klasifikačního modelu pro predikci nádechů	37
Tabulka 8: Ukázka tabulky predikovaných a anotovaných časů nádechů	38
Tabulka 9: Porovnání predikovaných časů (vlevo) a anotovaných časů (vpravo) nádechu pro subjekt K1006.....	39
Tabulka 10: Porovnání predikovaných časů (vlevo) a ručně anotovaných časů (vpravo) v sekundách pro subjekt P2005	40
Tabulka 11: Navržené parametry pro kvantifikaci respiračních poruch (prvních pět subjektů)	41
Tabulka 12: Výsledky statistické analýzy.....	43
Tabulka 13: Porovnání výsledků Shapiro-Wilkova testu, p-hodnoty	44
Tabulka 14: Pearsonův korelační koeficient pro navržené parametry	45
Tabulka 15: Hodnocení modelu strojového učení pro predikci klinických dat (průměr + směrodatná odchylka).....	46
Tabulka 16: Porovnání výsledků strojového učení	46

SEZNAM ZKRATEK

PN	Parkinsonova nemoc
HD	Hypokinetická dysartrie
VAD	Voice Activity Detector
MFCC	Mel-frequency Cepstral Coefficient
CNN	Convolutional Neural Network
LSTM	Long Short-Term Memory
GB	Gradient Boost
REM	Rapid Eye Movement
UPDRS	Unified Parkinson's Disease Rating Scale
ZK	Zdravá kontrola
LED	Levodopa Equivalent Dose
MMSE	Mini-Mental State Exam
ZCR	Zero-Crossing Rate
ZCC	Zero-Crossing Count
LSF	Line Spectral Frequencies
LPC	Linear Predictive Coding
ROC	Receiver Operating Characteristics
AUC	Area Under Curve
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
MAE	Mean Absolute Error
EER	Estimated Error Rate
ACC	Accuarcy
SEN	Sensitivita
SPEC	Specificita
CM	Confusion Matrix
MIN	Minimum
MAX	Maximum

ÚVOD

Více než 70 % pacientů s Parkinsonovou nemocí trpí hypokinetickou (extrapyramidovou) dysartrií, tedy poruchou motorické realizace řeči. Řeč pacientů bývá monotónní, zpomalená či zrychlující se a často méně zřetelná. Hypokinetická dysartrie je spojená s příznaky Parkinsonovy nemoci, přesněji ztuhlostí a zpomalenými pohyby svalstva. Hypokinetická dysartrie postihuje do jisté míry všechny oblasti tvorby řeči včetně respirace.

Současné metody pro kvantifikaci poruch spojených s hypokinetickou dysartrií se zaměřují spíše na řečový signál a především na řeč samotnou, než na oblast respirace, která bývá často opomíjena. V dnešní době není dostatečný nástroj pro kvantifikaci respiračních obtíží u pacientů s Parkinsonovou nemocí

Diplomová práce je zaměřena na návrh algoritmu pro automatizovanou detekci nádechů z řečového signálu a následně návrh parametrů pro kvantifikaci respiračních potíží u pacientů s Parkinsonovou nemocí.

Úvodní část práce je věnována popisu Parkinsonovy nemoci, zejména pak problémům s řečí nemocných pacientů, u kterých se vyskytuje hypokinetická dysartrie a její dopady na jednotlivé oblasti tvorby řeči. Hypokinetická dysartrie je souhrnné označení pro postižení motorických aspektů řeči u pacientů s Parkinsonovou nemocí. Dále je v úvodu popsán takzvaný současný stav poznání, tedy pohled na nejmodernější metody v oblasti zpracování řeči za účelem detekce nádechů z řečového signálu, detekce chorob z nádechů v řečovém signálu na základě akustických parametrů, strojového učení a umělé inteligence.

Třetí kapitola se zabývá metodikou příprav nahrávek zdravých kontrol a pacientů s Parkinsonovou nemocí k analýze v programu Wavesurfer a vytvoření databáze sledovaných subjektů, zpracování řečového signálu a jeho následnou analýzou, vytvoření algoritmu pro automatickou detekci nádechů pomocí strojového učení s využitím modelu logistické regrese a jeho popisem v programu MATLAB. Pro detekci nádechů byly vybrány parametry, používané v oblasti zpracování řeči, které jsou v této kapitole popsány. Následně jsou vyhodnoceny získané výsledky vytvořeného modelu na základě křížové validace pomocí specificity a sensitivity, přesnosti modelu a oblasti pod zobrazenou operační křivkou. Dále jsou zde popsány navržené parametry a metodika extrakce navržených parametrů pro kvantifikaci respiračních obtíží, statistická analýza pomocí statistických testů a metodika strojového učení.

Závěrečná část práce je věnována získaným výsledkům predikčního algoritmu pro detekci nádechů a výsledkům statistické analýzy. Tyto výsledky jsou zde vyhodnoceny a porovnány. V rámci diskuse výsledků jsou vybrány parametry, které se jeví jako vhodné pro objektivní kvantifikaci respiračních poruch u pacientů s Parkinsonovu nemocí. V rámci práce jsou také porovnávány výsledky statistických testů pro parametry

extrahované na základě automatizované detekce nádechů a pro parametry získané na základě ruční anotace nádechů. Dále jsou v rámci této sekce popsány limitace, které byly během práce objeveny. Závěr práce shrnuje dosažené cíle diplomové práce

1. PARKINSONOVA NEMOC

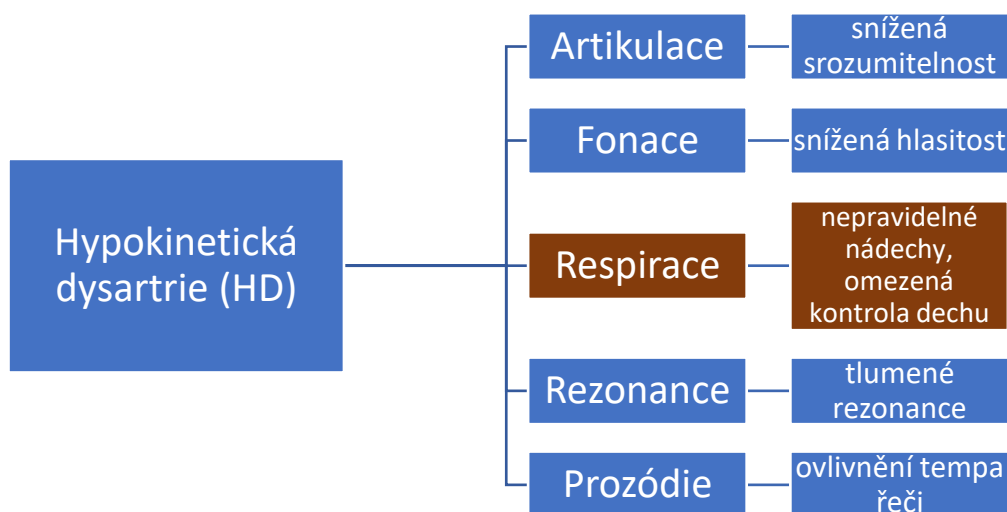
Parkinsonova nemoc (PN) je onemocnění mozku neurodegenerativního typu, které je zapříčiněno zejména nedostatkem dopaminu, který je do mozku dodáván, a je doprovázeno další řadou poruch v neurotransmitterových systémech nervové soustavy. Toto onemocnění bylo poprvé popsáno jako „třesavá obrna“ anglickým chirurgem Jamesem Parkinsonem v publikaci „An Essey on the Shaking Palsy“ (1817) [1].

Čtyřmi základními klinickými příznaky PN jsou: klidový třes (tremor), rigidita (tuhost, neohebnost), bradykineze (zpomalený pohyb) a postižení posturálních reflexů [2]. Parkinsonova nemoc postihuje řadu motorických i nemotorických funkcí, mimo jiné postihuje i motorické aspekty řeči. Dalším z projevů PN je respirační dysfunkce, způsobena oslabením dýchacích svalů, obstrukcí horních i dolních dýchacích cest [3].

V řečovém signálu se tyto příznaky projevují nižší srozumitelností, monotónností, kolísáním intenzity řeči, nemožností dokončit větu či slovo, nádechem uprostřed slova a dalšími obtížemi spojenými s respirací. Souhrnně lze motorické poruchy řeči označit jako hypokinetická dysartrie (HD). Poruchou řeči je postiženo průměrně 70-90 % pacientů s PN [2] [4].

1.1 Hypokinetická dysartrie

Termín dysartrie popisuje motorickou poruchu realizace řeči. Hypokinetická dysartrie vzniká narušením funkce bazálních ganglií a jejich spojů s centrální nervovou soustavou a její příčinou je PN. Může být také zapříčiněna traumatem nebo mozkovou příhodou, případně vlivem léků či drog [5]. Právě výše zmíněné příznaky (bradykineze, tremor a rigidita) mají negativní vliv na tvorbu řeči. Ztuhlost, zpomalenost pohybů a třes mají vliv na aspekty tvorby řeči jimiž jsou artikulace, prozódie, fonace a respirace [5].



Obrázek 1: vliv HD na oblasti tvorby řeči

1.1.1 Oblast fonace

Oblast fonace tedy procesu vytváření zvuku v hlasivkách a hrtanu je ovlivněna celkovou ztuhlostí svalstva v oblasti hrtanu. Nedokonalé ovládání hlasivek má za následek jejich neúplné sevření což může v kombinaci s respiračními obtížemi vést ke hlasovému třesu, zvýšenou střední hodnotu základního kmitočtu hlasu, hypofonii (snížená intenzita hlasu) a dysfonii (snížení kvality hlasu) [6].

1.1.2 Oblast prozódie

Oblast prozódie je oblast tvorby řeči zkoumající jevy jako přízvuk, melodie, tón, rytmus řeči a její frázování. Řečník s PN má obvykle pomalé a kolísavé tempo řeči, objevují se nepřírozně dlouhé pauzy či pauzy uprostřed slov, řeč je monotónní a může také postrádat emoční zabarvení. To má za následek celkové snížení srozumitelnosti. Oblast prozódie bývá společně s artikulací ovlivněna také Hypomimií, tedy omezení mimických funkcí, jenž je také častým příznakem PN [3] [6].

1.1.3 Oblast artikulace

Artikulace je oblast řeči zabývající se tvorbou hlásek pomocí mluvidel. Oblast artikulace je postižena nesprávnou funkcí svalstva podílejícího se na tvorbě hlásek (jazyk, rty, mimické svaly a svaly dásní). Hodnocení artikulace je prováděno pomocí tzv. diadochokinetických úloh, souborem úloh obsahující exploziva a samohlásky v rychlém sledu, při kterém dochází ke změnám postavení artikulačních orgánů. Snížená pohyblivost a ztuhlost znemožňují pacientovi tyto rychlé změny provádět a dochází tak k nižší srozumitelnosti a nepřírozenosti řeči [3] [6].

1.1.4 Oblast respirace

Další nedílnou součástí tvorby řečového signálu, která je ovlivněna Parkinsonovu nemocí, je respirace. Při nádechu je do plic vehnán vzduch, jehož proud při výdechu rozkmitává hlasivky. U pacientů s HD je narušena jak fáze nádechu, tak fáze výdechu.

Oblast respirace, která je pro účely práce stěžejní, je ovlivněna ztuhlostí dýchacích svalů, která vede ke snížení vitální kapacity plic, zvyšuje se frekvence nádechů, nemožnost ovlivnit nakládání s energií dechu. Objevují se nepravidelné, mělké nádechy a delší pauzy, při kterých se pacient snaží tzv. popadnout dech. Pacienti s HD mají tendenci zrychlovat tempo řeči a vysloví méně slov na jeden nádech než zdravé kontroly. Postižení respiračních svalů je v diagnostice HD často opomíjeno [3].

Řeč je komplexní prostředek pro sdělování informací a je silně individuální. Hypokinetická dysartrie se tak projevuje u každého pacienta odlišně a působí na jednotlivé oblasti tvorby řeči v různé míře [7].

A. D'Arrigo se se svým týmem věnuje shrnutí poznatků a vědomostí o respiračních dysfunkcích pacientů trpících PN. Článek vzniká také na popud opomíjení respiračních problémů v rámci diagnostiky a léčby Parkinsonovy nemoci. Zabývá se obstrukčními respiračními obtížemi postihující zejména horní cesty dýchací, dále se věnuje restriktivním dysfunkcím respirace a změnám v respiračních vzorcích. Studie se zaměřuje také na další potíže spojené s PN jako jsou spánková apnoe a pneumonie. V článku je zmíněn vliv dopaminergické terapie, korelace mezi pneumologickými léčivy a PN a léčby pomocí hloubkové mozkové stimulace. Autor popisuje artefakty, které se vyskytují u pacientů. Jeden z nich je popsán jako “respirační třes”, který je zapříčiněn tremorem a bradykinezií, dále zmiňuje stridor = sípavý, pískavý zvuk během nádechů a výdechů. Stridor se také může projevovat jako zvuk chrčení či pískání. [8]

Často pozorovanými jevy u pacientů s PN oproti zdravým kontrolám jsou počet nádechů, změny v dechovém vzorci, rozdílná práce s intenzitou řeči, rozdílná délka nádechů a výdechů [9]. Respirační a plicní dysfunkce u osob s PN jsou především v pozdější, rozvinuté fázi nemoci jedním z hlavních faktorů vedoucí k mortalitě a morbiditě pacientů. Nemoc bývá doprovázena zvýšenou dušností, obtížemi s popadnutím dechu a lapáním po dechu [10].

2. SOUČASNÝ STAV POZNÁNÍ (STATE OF THE ART)

Zpracování řeči je v dnešní době všudypřítomná technologická disciplína, která je hojně využívána ve všech inteligentních zařízeních od telefonů, počítačů, chytrých hodinek a spotřebičů až po využití v oblasti medicíny, kdy například zpracování řeči napomáhalo v diagnostice onemocnění či právě při kvantifikaci určitých onemocnění či zdravotních poruch [11].

2.1 Detekce nádechů

Článek **Automatic Detection of Breath Using Voice Activity Detection and SVM Classifier with Application on News Reports** [12] z roku 2019 pojednává o automatické detekci nádechů z řečového signálu s použitím detektoru řečové aktivity a klasifikátoru využívajícím metodu podpurných vektorů. Pomocí detektoru řečové aktivity (VAD) byly ze signálu vybrány segmenty řeči, které VAD označil jako ne-řečové. Pro práci byl použit VAD z knihovny Covarep. Pomocí MFCC byl vytvořen kepstrogram sloužící jako vstup pro klasifikátor rozhodující, zdali se jedná o nádech či nikoliv. Trénovací set dat obsahoval 131 segmentů dechu a 154 nedečových segmentů. Výsledky klasifikace byly upraveny a odstraněny segmenty, které obsahovaly nádechy kratší než 100 ms. Výsledkem je *speech breathing rate* (*sbr*), volně přeloženo jako míra nádechů v řeči, počítaná dle vzorce:

$$sbr = \frac{N-1}{bs_N - bs_1} \quad (1)$$

kde N představuje počet nádechů a bs_N , bs_1 představuje počátek posledního, respektive prvního nádechu.

Navrhovaná metoda s použitím VAD měla lepší výsledky (F1 Score, přesnost). Výhodou použití VAD je možnost vyřazení segmentů obsahující řeč, které jsou pro detekci nádechů zbytečné. Zvyšuje se tím i rychlost algoritmu [12]. Podobný postup s použitím VAD pro vyloučení řečových segmentů je využít i v následující práci.

V. S. Nallanthingal se v článku [13] z roku 2019 zabývá detekcí nádechů z nahrávky konverzace pomocí hlubokého učení. Práce porovnává výsledky dvou typů neuronových sítí, a to konvoluční neuronovou sítí (CNN) a Long short-term memory neuronovou sítí (LSTM). Vstupem pro neuronové sítě je spektrogram signálu a log Mel spektrogram. Analýzou dechového signálu jsou získány dva parametry dechu, a to *Breathing rate*, který udává průměrný počet nádechů za minutu a *Tidal volume* určující množství vzduchu dodaného do plic během normálního nádechu. Nejlepších výsledků bylo dosaženo

použitím LSTM neuronové sítě a log Mel spektrogramu jako vstupního parametru oproti klasickému spektrogramu [13].

Článek **Project Achoo: A Practical Model and Application for COVID-19 Detection From Recordings of Breath, Voice, and Cough** [14] vydaný v únoru roku 2022 se věnuje detekci onemocnění COVID-19 ze zvukových nahrávek pořízených na spotřebitelských zařízeních. K detekci je kromě parametrů řeči využito i nádechů a kašle. Práce popisuje využití neuronových sítí pro detekci onemocnění COVID-19 společně s metodami odstranění šumu a zlepšení kvality audio signálu, jelikož nahrávky pořízené například skrze telefon nedosahují takové kvality jako nahrávky pořízené v kontrolovaném prostředí na profesionální techniku. V práci je využita kombinace CNN a Gradient Boosting (GB), ke klasifikaci slouží Mel spektrogramy vypočítané ze segmentů pořízených nahrávek [14].

Článek **An Overview on Audio, Signal, Speech, & Language Processing for COVID-19** [15] z května roku 2020 popisuje metody moderní technologie v oblasti zpracování řeči s využitím umělé inteligence k rozpoznání příznaků nemoci COVID-19. Pro detekci kašle specifického pro COVID-19 je nutné shromáždit množství COVID a ne-COVID zvukových nahrávek kašle. Pro tyto účely začaly vznikat knihovny, do kterých je možné pomocí telefonního mikrofону nahrát zvukové záznamy kašle. Kašel lze považovat za specifický případ respirace. V článku je zmíněna studie pro automatické měření počtu nádechů během řeči, kdy na základě kepstrogramu a použití metody podpůrných vektorů vykazoval model dobré výsledky pro parametr speech-breathing rate, který udává počet nádechů za minutu [16]. V další studii [17], zaměřené na detekci dechového signálu během řeči, dosahoval model strojového učení sensitivity 91,2 % pro detekci respirační události a MAE 1,01 nádechů za minutu při měření míry respirace (počtu nádechů za minutu). Tento model využívá pro detekci nádechu spektrogramy a kombinaci konvoluční a rekurentní neuronové sítě. Za zmínku stojí například aplikace detekující covidový kašel [18] nebo tzv. „Chat-boty“, které dokáží rozpoznat nemoc s přesností až 90 %. Chat-bot pojmenovaný *Symptoma* [19] dokáže rozpoznat okolo 20 000 nemocí s velice vysokou přesností a COVID-19 rozpozná onemocnění s podobnými symptomy s přesností až 96,32 % [15].

2.2 Kvantifikace poruch

Hlavnička et al. [20] představují metodu pro automatizovanou segmentaci a analýzu důležitých parametrů řeči, které jsou zapříčiněny respiračním deficitem, dysfonií, špatnou artikulací a dysrytmií u pacientů s PN. V práci jsou také zahrnuti pacienti s poruchou Rapid Eye Movement (REM) fáze spánku.

Pro analýzu řeči je použit detektor řečové aktivity a hlavními parametry řeči jsou energie, zero-crossing rate, melovské keprální koeficienty a autokorelační funkce. Na základě těchto parametrů jsou za pomoci segmentace nahrávek vytvořeny segmenty znělé či neznělé řeči, segmenty respirace a segmenty pauz.

V práci je zkoumáno několik aspektů řeči, z nichž je pro tuto práci nejdůležitější oblast respirace. V této oblasti byly definovány čtyři parametry hodnotící kvalitu respirace. Těmito parametry jsou: relativní hlasitost respirace, míra respirace udávající počet nádechů za určitý čas, délka pauzy před a po nádechu a intervaly mezi nádechy.

Tyto parametry, společně s dalšími parametry jednotlivých aspektů řeči jsou analyzovány za pomoci statistické analýzy. Parametr kvantifikující hlasitost respirace nevykazuje při čtení textu žádné odchylky mezi lidmi s PN a zdravými kontrolami. Mírné odchylky vykazuje parametr v rámci nahraného monologu, kdy je průměrná hlasitost nádechu nižší u zdravých kontrol. V rámci čtené úlohy nevykazuje odchylky ani parametr kvantifikující délku pauzy před nebo po nádechu. V rámci monologu je však tato pauza v průměru o 30 ms delší u pacientů s PN. Zdravé kontroly vykazovaly v průměru o jeden nádech více během určitého časového intervalu, a to jak při čtení textu, tak při monologu. Míra respirace pacientů s PN byla totožná u obou úloh (čtení a monolog) a téměř totožná v porovnání s mírou respirace zdravých kontrol při čtení textu. Při monologu však byla míra respirace u zdravých kontrol o trochu menší.

Ačkoliv byl v rámci práce vyvinut kvalitní detektor řeči a nádechů, navržené parametry pro hodnocení poruch řeči nevykazovaly žádné významné korelace s klinickým skóre, hodnocením UPDRS III nebo významné odchylky parametrů hodnotící respiraci porovnáním zdravých kontrol a pacientů s PN.

Z výše uvedených článků je patrné, že pro detekci nádechů z řečového signálu a celkově číslicové zpracování řeči je v moderní době standardem využití strojového učení. Jedním z nejčastěji užívaných parametrů řeči pro detekci nádechů je MFCC a Mel spektrum, užívají se však i další parametry jako například míra průchodu signálu nulovou hodnotou za určitý úsek, celkový počet průchodů signálu nulovou hodnotou, krátkodobá energie nebo Percepční lineární predikční koeficienty (PLP) [21] [22].

Zpracování řeči, následná analýza řečového signálu a jeho aspektů jako je plynulost, energie, tempo, respirace a jiné mohou být užitečným nástrojem pro detekci a diagnózu řady nemocí. Analýza dechu v řečovém signálu může být využita pro prognózu řady respiračních onemocnění jako je astma, COVID-19 či výše zmíněná hypokinetická dysartrie. Pro detekci nádechů lze také kromě řečového signálu využít moderní metody

jako například vysokorychlostní termální kameru, respirační monitorovací systém či různě modifikované dýchací masky.

Nedostatkem v dnešní době je opomíjení respiračních obtíží, diagnostiky těchto obtíží ve spojení s Parkinsonovu nemocí. V případě objektivního hodnocení v rámci diagnostiky či kvantifikace poruch pomocí strojového učení je na prvním místě extrakce parametrů na základě řečového signálu, nikoli na základě respiračního signálu. Této problematice by pomohl vývoj dokonalého detektoru nádechů/řeči a zaměření pozornosti na respirační dysfunkce a včasná diagnostika, které by mohl napomocť nástroj pro kvantifikaci respiračních poruch a dysfunkcí.

I přes to, že je v dnešní době zpracování řečového signálu a jeho analýza na velmi vysoké úrovni a je využíváno řadou chytrých zařízení a má přesah v rámci oborů, analýzy samotného nádechu a následného zpracování a použití v oblasti medicíny není zatím plně prozkoumáno a využíváno.

Na základě provedené rešerše výše uvedených článků a literatury, byly navrhnuty a definovány parametry, které by mohly být využity pro kvantifikaci respiračních poruch pacientů s PN.

3. METODIKA PRÁCE

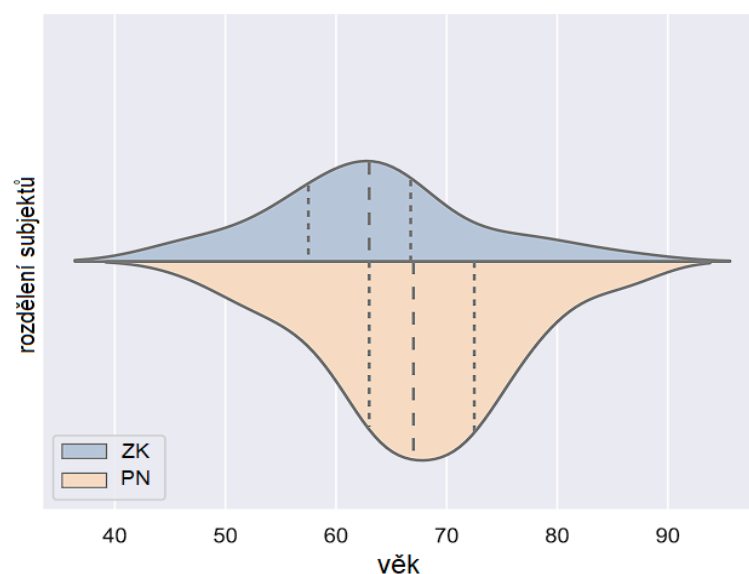
Následující kapitola popisuje metodiku semestrální práce zahrnující vytvoření a popis databáze zkoumaných subjektů, ruční anotace nádechů v programu wavesurfer, zpracování řečového signálu (normalizace, preemfáze, segmentace, získání parametrů řeči) v programu MATLAB, využití logistické regrese k predikci nádechů a popis skriptu, pomocí kterého lze převést segmenty predikovaných nádechů na čas. Dále je zde popsána parametrizace nahrávek na základě predikovaných časů, návrh parametrů kvantifikující respirační poruchy a provedení statistické analýzy v programovacím jazyce Python.

3.1 Databáze zkoumaných subjektů

Pro účely práce bylo zkoumáno 136 subjektů z nichž 91 lidí trpí Parkinsonovou nemocí a 45 je zdravých kontrol. Poměr žen a mužů je v případě zdravých kontrol vyrovnaný, u nemocných pacientů je zastoupení mužského pohlaví rovno 60 % z celkového počtu zkoumaných subjektů s PN. Průměrný věk zkoumaných subjektů je roven 66,1 let, přitom nejnižší věk je roven 45. Nejstarší subjekt je starý 87 let, věkový modus je roven 65. Délka nemoci u osob s PN se pohybuje od 9–18 let, průměrná délka doby nemoci se pohybuje okolo 14 let.

	Ženy	Muži	celkově
ZK	24	21	45
PN	35	56	91
celkově	59	78	136

Tabulka 1: Demografická data zkoumaných subjektů



Obrázek 2: Rozložení věku zkoumaných subjektů

Klinická data zkoumaných subjektů jsou obsažena v Tabulkách 2 a 3, které obsahují společně s informacemi o pohlaví a věku také údaje o délce trvání nemoci, míře medikace nemocných pacientů, hodnoty na škálách Unified Parkinson's Disease Rating Scale – part 3 (UPDRS III) a celkový dysartrický index (DX index), které slouží ke kvantifikaci poruch spojených s PN. DX index je hodnocení vytvořené na základě 3F testu, jehož výsledkem je dysartrický profil, který je tvořený součtem bodů z testů faciokineze, fonorespirace a fonetiky [23]. Dále je v Tabulce 3 uvedena průměrná hodnota denní dávky léku Levodopa (LED), který je nejčastěji užíván v léčbě PN a průměrné výsledky krátkého testu kognitivních funkcí (MMSE). Zdravé kontroly jsou označeny ID: **K**1003 – **K**XXX, pro pacienty s PN začíná označení písmenem **P**1006 – **P**XXX.

Společně s tabulkou obsahující klinická data subjektů je k dispozici 136 zvukových nahrávek označených pomocí ID zkoumaných subjektů, které byli pořizeny během čtení textu. Nahrávky jsou dlouhé v průměru 60 sekund.

ID	diagnoza	pohlavi	vek	trvani_ner	LED	UPDRSIII	UPDRS IV	overall DX
K1003	ZK	F	63					87
K1004	ZK	F	65					90
K1005	ZK	F	59					90
K1006	ZK	F	64					89
K1007	ZK	F	59					88
P1106	PN	F	72	12	907	52	4	61
P1108	PN	F	63	13	1354	34	5	62
P1109	PN	F	65	9	228	37	9	30
P1110	PN	F	62	12	1993	31	0	73
P2001	PN	M	65	13	1680	31	4	72
P2003	PN	M	67	13	1768	12	2	86
P2004	PN	M	80	18	300	41	2	79

Tabulka 2: Ukázka tabulky klinických dat pacientů

	ZK	PN
Délka nemoci [rok]	-	14.2 ± 2.8
LED [mg]	-	984.8 ± 552.0
faciokineze	27.9 ± 1.8	24.6 ± 3.4
fonorespirace	28.5 ± 1.5	24.1 ± 3.6
fonetika	29.5 ± 1.0	25.7 ± 3.7
celkový DX index	85.8 ± 3.2	74.4 ± 9.1
UPDRS III	-	24.9 ± 9.8
MMSE	28.2 ± 1.2	28.0 ± 1.5

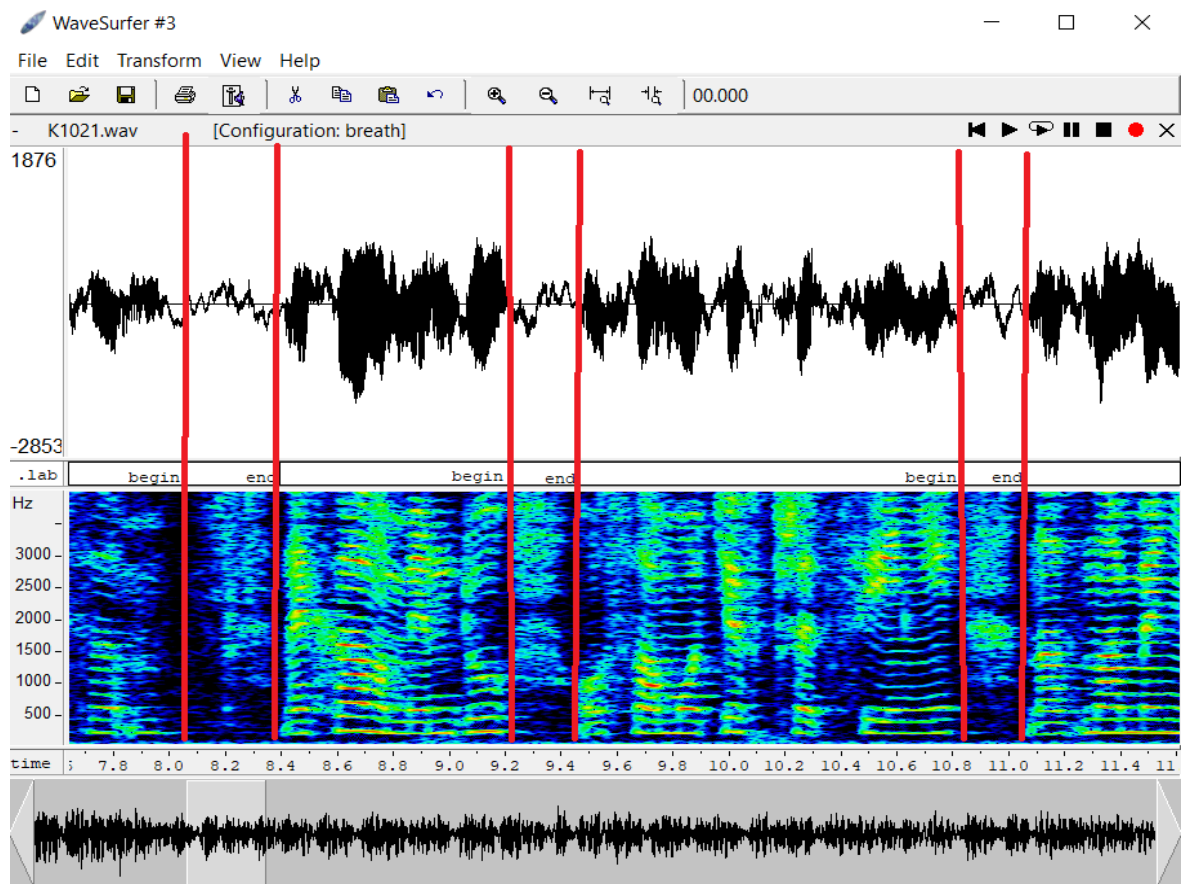
Tabulka 3: Průměrná klinická data

3.2 Anotace nahrávek

Prvním krokem práce byla ruční anotace nádechů v jednotlivých zvukových nahrávkách zkoumaných subjektů. Anotace byla provedeno pomocí jednoduchého open source programu Wavesurfer [24] (obr. 3), který slouží pro práci se zvukem, jeho vizualizaci pomocí spektrogramů a časového průběhu a slouží také pro analýzu zvuku a řeči.

Po nahrání zvukového souboru do programu Wavesurfer je v horní části okna zobrazen časový průběh amplitudy řečového signálu a ve spodní části je zobrazeno frekvenční spektrum signálu. Uprostřed se nachází úzké pole sloužící k výpisu poznámek či komentářů, v tomto případě pro označení nádechů.

Detekce nádechů při anotaci byla provedena zejména na základě sluchu, nádech lze však detekovat také pomocí vizualizace v programu wavesurfer. Během nádechu lze v časové oblasti pozorovat nižší počet průchodů amplitudy řečového signálu nulovou hodnotou oproti řeči, ve spektrální oblasti lze během nádechu pozorovat absenci základní frekvence neboli fundamentu a šumový charakter spektra v oblasti od 1 do 4 kHz (obr. 3). Detekce nádechů na základě vizuální stránky je však podmíněna kvalitou nahrávky. V případě méně kvalitní nahrávky je schopnost detekce pomocí zraku snížena.



Obrázek 3: prostředí programu Wavesurfer

Anotace nahrávek spočívala v označení začátku (značka begin) a konce (značka end) nádechu v nahrávkách čteného textu. Označeny byly pouze nádechy, které byly jasné a zřetelné bez nežádoucích rušivých zvuků okolí.

Během fáze anotace nahrávek bylo vyřazeno z databáze celkem 14 nahrávek, z původně dostupných 150, z důvodu nízké kvality. Nahrávky obsahovaly nežádoucí zvuky například vyzvánění telefonu, narušení nahrávání třetí osobou, tikání hodin nebo byly nahrávky nekvalitně nahrány (vysoký obsah šumu a malá energie užitečného signálu).

Výsledkem anotace je textový soubor ve formátu „ID“.lab (K1003.lab) s označenými začátky a konci nádechů v sekundách. Z časových údajů byly dále vytvořeny segmenty nádechu (obr.3).

```
0.0000000 5.5255837 begin
5.5255837 5.8433048 end
5.8433048 9.9598646 begin
9.9598646 10.2085159 end
10.2085159 12.1378655 begin
12.1378655 12.6259587 end
12.6259587 19.8138222 begin
19.8138222 20.0394502 end
```

Obrázek 4: ukázka .lab souboru s časy nádechů

3.3 Předzpracování řečového signálu

Pro analýzu nahrávek a následné vytvoření algoritmu pro detekci nádechů bylo potřeba zvukový signál nejprve zpracovat na jednotlivé vzorky pomocí vzorkování, normalizovat a následně segmentovat pomocí váhovacího okna na krátké rámce o délce N vzorků, ze kterých lze vypočítat jednotlivé parametry řeči. Ke zpracování nahrávek byl použit program MATLAB R2022b vyvinut společností MathWorks.

3.3.1 Vzorkování a normalizace signálu

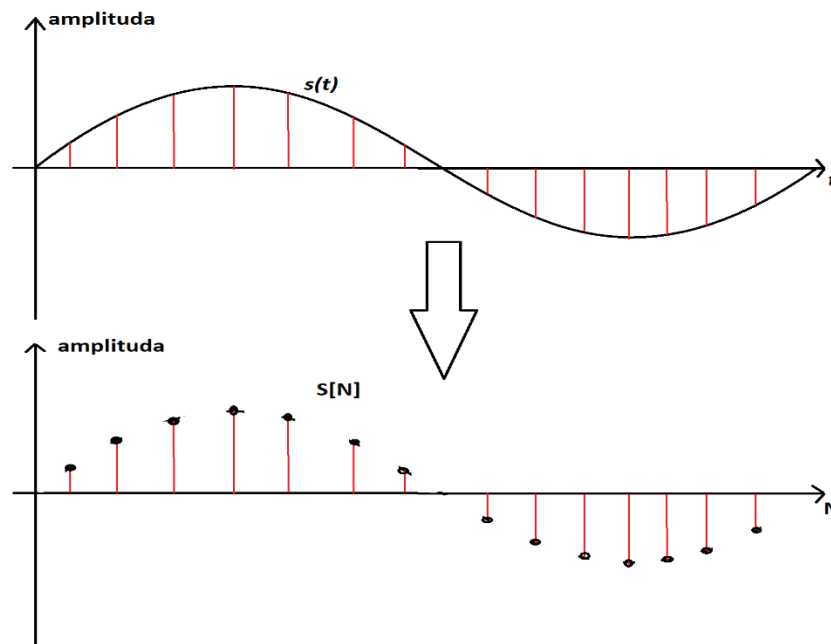
Prvním krokem zpracování nahrávek řeči je jejich **vzorkování**. Tento proces představuje transformaci signálu $x(t)$, který je spojitý v čase, na posloupnost vzorků $x_{[n]} = x[nT]$, které jsou diskrétní v čase a kde $n \in \mathbb{Z}$ a T je vzorkovací perioda.

V prostředí MATLAB je načtení vzorků signálu provedeno pomocí příkazu:

```
[y,Fs] = audioread(filename) % filename = název souboru
```

Výstupní proměnná y obsahuje navzorkovaný zvukový soubor formátu wav a proměnná F_s je rovna vzorkovací frekvenci, kterou byla nahrávka navzorkována.

Princip vzorkování je popsán na obrázku 5.



Obrázek 5: princip vzorkování

Dále byla provedena **normalizace**. Normalizace je úprava, která aplikuje zesílení či zeslabení celého signálu tak, aby maximální hodnota amplitudy dosahovala určité hodnoty. Vzhledem k použití zesílení/zeslabení na celý signál, zůstává odstup signálu od šumu stejný jako před normalizací. Stejně je tomu tak i v případě dynamického rozsahu. V programu MATLAB je normalizace provedena pomocí jednoduchého příkazu:

```
y = y/abs(max(y)); % y = navzorkovaný zvukový signál
```

3.3.2 Preemfázový filtr

Preemfázový číslicový filtr je využíván pro zmírnění negativních vlivů kvantizačního šumu na řečový signál. Ačkoliv je podstatná část energie signálu pod frekvenční hranicí 300 Hz, užitečné informace jsou obsaženy ve frekvenční oblasti nad 300 Hz. Vzhledem k povaze kvantizačního šumu, tedy jeho rovnoměrného spektra, jsou negativně ovlivněny vyšší složky spektra řeči se slabší energií.

Preemfázový filtr je filtr typu horní propust (FIR) s přenosovou funkcí:

$$H(z) = 1 - a_x z^{-1} \quad (2)$$

Koeficient a_x bývá zpravidla volen v rozsahu $0,9 - 1$ [21]. Pro tuto práci byla zvolena hodnota koeficientu $a_x = 0,92$. Filtrace řečového signálu je v prostředí MATLAB realizována pomocí příkazu:

```
alfa = 0.92;  
x_preem = filter([1 -alfa], [1 0], x); % x = zvukový signal (normalizován)
```

3.3.3 Segmentace signálu

Řečový signál je zpracováván pomocí metod krátkodobé analýzy, na základě předpokladu, že se řečový signál a jeho vlastnosti mění v úsecích krátké délky dostatečně pomalu.

Signál je rozdělen na úseky o délce N pomocí váhovacích oken, nejčastěji používanými okny jsou **pravoúhlé okno** a **Hammingovo okno**. Délka okna bývá při zpracování řeči v rozsahu od 10 do 30 milisekund.

Hammingovo okno je definováno:

$$\begin{aligned} w[n] &= 0,54 - 0,46 \cos \left[n \frac{2\pi}{N} \right], & \text{pro } n = 0, 1, \dots, N-1, \\ w[n] &= 0, & \text{pro ostatní } n. \end{aligned} \quad (3)$$

Okna se mohou či nemusí překrývat, jejich překrytím lze dosáhnout přesnějších výsledků výpočtu řečových parametrů, narůstají však nároky na výpočet [21].

Pro účely práce bylo při segmentaci řečového signálu použito Hammingovo okno o délce 25 ms, s překrytím 50 %. Jeden segment tedy obsahuje 400 vzorků, při vzorkovací frekvenci 16000 Hz. V prostředí MATLAB je k segmentaci použita pomocná funkce *segmentation.m* (viz příloha). Vstupními parametry funkce jsou navzorkovaný a normalizovaný signál řeči, typ váhovacího okna společně s jeho délkou a míra překrytí oken.

3.4 Automatizovaná detekce nádechů

Tato kapitola je věnována zdrojovému kódu a funkci jednotlivých skriptů pro automatizovanou detekci nádechů v nahrávkách řeči. Celý algoritmus detekce nádechů se skládá ze tří skriptů napsaných v programu MATLAB.

3.4.1 Volba parametrů řeči

Ze získaných segmentů řečového signálu lze vypočítat parametry řeči. Pro účely automatizované detekce nádechů byly počítány Melovské kepstrální koeficienty, energie signálu, průchod amplitudy signálu nulovou hodnotou (zero-crossing count/zero-crossing rate) a řádkové spektrální frekvence neboli LSF.

MFCC neboli Melovské kepstrální koeficienty jsou jedním ze základních parametrů pro analýzu řečového signálu (viz [12], [13], [14], [21]). MFCC spočívá ve výpočtu reálného kepstra signálu, který má nelineárně transformovanou kmitočtovou osu ve spektrální oblasti. Kmitočtová osa je transformována na melovskou stupnici [21].

Výpočet MFCC je v programu MATLAB uskutečněn pomocí příkazu:

```
coeffs = mfcc(Seg, fs, Window=wind, OverlapLength=winover, NumCoeffs=17)
```

Pro účely práce bylo vypočítáno 17 koeficientů. Při detekci nádechů pomocí logistické regrese se nejlépe osvědčilo použití prvních 13 melovských kepstrálních koeficientů.

Energie diskrétního signálu v rámci jednotlivých segmentů je rovna

$$E = |s[n]|^2. \quad (4)$$

Energie je vhodným parametrem pro analýzu řeči a detekci nádechů, během nádechu je energie signálu mnohem menší než v případě řeči.

Parametr **zero-crossing rate** (ZCR) vyjadřuje míru s jakou přechází signál přes nulovou hodnotu z pozitivních hodnot do negativních či naopak. ZCR je vyjádřeno jako podíl počtu průchodů signálu nulovou hodnotou a počtu vzorků v segmentu.

Zero-crossing count (ZCC) pak udává počet těchto průchodů. V programu MATLAB je výpočet těchto parametrů proveden pomocí příkazu

```
[rate, count]=zerocrossrate(x) %x = vstupní signál
```

Míra průchodu nulou je parametr vhodný pro detekci nádechů, jelikož během nádechu a pauz je ZCR a ZCC menší než v případě řeči. Řeč pacienta s PN obsahuje více pauz, někdy jsou pauzy delší než u zdravých kontrol.

LSF neboli řádkové spektrální frekvence či jsou dalším parametrem používaným pro zpracování řeči. Jejich výpočet je proveden z koeficientů lineární predikce. Bylo zjištěno, že LSF má lepší využití ve zpracování řeči než LPC koeficienty vzhledem k lepší interpolaci a menší náchylnosti na kvantovací šum [25].

V prostředí MATLAB jsou nejprve vypočítány LPC koeficienty, ze kterých je následně vypočítán parametr LSF pomocí příkazů

```
a = lpc(x, 9); % x = vstupní signál
lsf = poly2lsf(a); % a = koeficienty LPC
```

Pro účely práce bylo vypočítáno 9 LPC koeficientů, ze kterých byly získány LSF.

V prvním skriptu *Breath_process.m* jsou zpracovány nahrávky řečového signálu, vytvoření vektoru nádechů z labelů jednotlivých nahrávek, použití VAD pro vyřazení řečových segmentů a výpočet parametrů řeči.

Druhý skript *Log_Reg.m* obsahuje zdrojový kód s modelem logistické regrese, pomocí kterého probíhá predikce nádechů v řečovém signálu. Skript dále obsahuje vyhodnocení modelu pomocí křížové validace.

Poslední skript *Time_back.m* slouží k výpočtu času predikovaných nádechů ze segmentů a k porovnání s ručně anotovanými časy.

3.4.2 Vytvoření datasetu

V úvodní části skriptu jsou v prostředí MATLAB předzpracovány nahrávky podle metod popsanych v kapitole 3.3 a vytvořeny vektory nádechů na základě délky vektoru navzorkované nahrávky, anotovaných časů nádechu a vzorkovací frekvence. Vektor nádechů obsahuje hodnotu nula na místě, kde se nachází řeč, místa nádechu jsou naopak označena číslem jedna.

Dalším krokem je použití detektoru řečové aktivity (VAD) pro detekci vzorků obsahující řečový signál. Pro účely práce byl použit VAD založený na výpočtu spektrální energie [26]. Funkce je volána pomocí příkazu:

```
[activeNum, pos] = vad_YW(speech, fs, plotMode, Buffer, activeMax, energyTh2).
```

Kde vstupní parametr *speech* představuje navzorkovaný signál nahrávky řeči, *fs* vzorkovací frekvence, *plotMode* umožňuje zobrazení výstupu VAD, *Buffer* je hodnota zadávaná v sekundách, umožňuje aplikaci vyrovnávací paměti na obě strany řečového bloku, *activeMax* určuje maximální délku řečového bloku v sekundách, *energyTh2* je vstupním parametrem, který určuje hodnotu energie, od které se bude řeč detekovat. Výstupy funkce jsou *activeNum* udávající počet vzorků označených jako řeč a *pos* udávající pozici (index) těchto vzorků.

Pro účely práce bylo potřeba dosažení co nejmenšího překrytí vzorků označených jako řeč se vzorky anotovaných nádechů. Hlavní myšlenkou tohoto kroku je omezení datasetu pro logistickou regresi pouze na neřečové segmenty. To bylo dosaženo pomocí

jednoduchého for cyklu, ve kterém se zvyšovala hodnota $energy_{Th2}$ do chvíle, kdy byla hodnota překrytých vzorků menší než 1600, což odpovídá 0,1 s z celkové nahrávky. Maximální délka řeči byla nastavena na 12 sekund.

V následujícím kroku byl aplikován preemfázový filtr na řečový signál, jeho segmentace, segmentace výstupu VAD a segmentace vektoru anotovaných nádechů. Na základě získaných indexů z výstupu VAD bylo možné vyloučit segmenty, které byly pomocí VAD označeny jako řeč. Tento postup je velmi podobný postupu z kapitoly 2 [12].

Poslední část skriptu obsahuje výpočet parametrů řeči uvedených v kapitole 3.3 společně s vektorem segmentů s nádechy, které byly ručně anotovány a jejich uložení do tabulky (viz tabulka 2) pro další zpracování.

Celý skript je součástí for cyklu, který postupně provádí výše uvedené příkazy pro každý lab soubor ve složce *Labels* (viz příloha). Po spuštění skriptu pak stačí pouze vybrat tuto složku souborů.

3.4.3 Logistická regrese

V počáteční sekci skriptu jsou nahrány výsledky výpočtů řečových parametrů a z nich vytvořený dataset, sloužící jako vstupní data pro algoritmus na detekci nádechů pomocí logistické regrese.

Vyhodnocení modelu je provedeno na základě křížové validace, přesněji pomocí upravené validační metody Leave-one-out Cross Validation. Jako testovací data jsou vybrány řečové parametry jednoho zkoumaného subjektu, trénovacími daty modelu jsou parametry řeči zdravých kontrol.

Pomocí for cyklu je z datasetu vybrán vždy jeden subjekt a jeho parametry, v trénovací sadě jsou pak tyto parametry vymazány.

Vytvoření modelu logistické regrese je v prostředí MATLAB uskutečněno pomocí příkazu:

```
B = mnrfit(Dataset_in, Dataset_target);
```

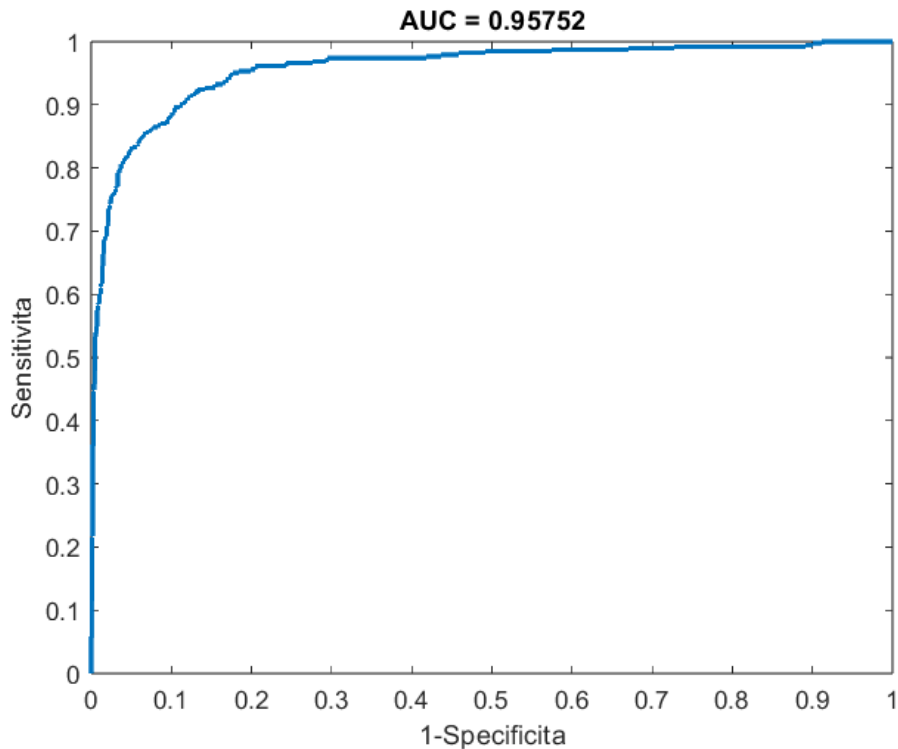
kde *Dataset_in* obsahuje řečové parametry trénovacího datasetu a *Dataset_target* obsahuje vektor se segmenty označenými jako nádech na základě ruční anotace. Výstupem je matice *B* obsahující odhady koeficientů logistické regrese.

Pro výpočet pravděpodobností výskytu nádechu v rámci segmentů slouží příkaz:

```
pihat = mnrval(B, Dataset_test);
```

kde vstupní parametr *Dataset_test* obsahuje parametry řeči právě zkoumaného subjektu z testovacího datasetu a parametr *B* obsahuje koeficienty logistické regrese.

Ze získaných výsledků predikovaných časů nádechu je možné zobrazit křivku operační charakteristiky (ROC) a z ní vypočítat plochu pod její křivkou (AUC) (obr.6), která slouží k hodnocení vytvořeného modelu.



Obrázek 6: ROC křivka (subjekt K1003)

Pomocí predikovaných a anotovaných vzorků nádechu byla vypočítána matice záměn, tedy TP = pravdivě pozitivní, TN = pravdivě negativní, FP = falešně pozitivní a FN = falešně negativní. Z těchto hodnot byla vypočítána přesnost modelu, sensitivita a specificita, ze kterých byla vyobrazena ROC křivka a spočítána hodnota AUC.

Všechny výsledky jsou uloženy do tabulky s názvem *Tabulka_Cross_val.mat* (viz příloha). Pomocí for cyklu jsou automaticky upravovány vstupní datasety pro vytvoření modelu na základě právě zkoumaného subjektu. Ze získaných metrik byl vypočten jejich průměr pro celkové hodnocení modelu a pomocí funkcí *rocmetrics* a *average*, které jsou součástí „Statistics and Machine Learning Toolbox“, byla vyobrazena průměrná ROC křivka.

3.4.4 Získání časových údajů

Poslední skript slouží k vyjádření predikovaných nádechů v čase. Většina obsahu skriptu je totožná s obsahem ze skriptu *Breath_process.m*. Odlišuje se až v závěrečné části, ve které jsou nahrána data z přechozího skriptu, která obsahují vektory segmentů s predikovaných nádechů.

Na základě pozice (indexů) segmentů predikovaných modelem jako nádech a známé pozice ne-řečových segmentů vybraných na základě výstupu z VAD lze určit časy predikovaných nádechů. K tomu je zapotřebí znát také délku váhovacího okna při segmentaci a hodnotu překrytí oken. Dalším krokem bylo odstranění špatně predikovaných nádechů, které jsou kratší než 100 milisekund a omezení maximální délky nádechu na 1,5 sekund.

Výsledkem je tabulka 7 obsahující ID subjektů, časy predikovaných nádechů a časy ručně anotovaných nádechů pro jejich porovnání. Dále je výsledkem XXXX.lab soubor, který obsahuje uložené anotované časy ve stejném formátu jako anotace z programu wavesurfer, pro každý zkoumaný subjekt.

3.5 Parametrizace nahrávek

Na základě vygenerovaných anotací časů nádechů jsou, za pomoci skriptu *get_features.m* v prostředí MATLAB, získány navržené parametry pro kvantifikaci respiračních poruch. Vstupem tohoto skriptu je složka obsahující .lab soubory s vygenerovanými časy nádechů a složka obsahující .wav soubory s nahrávkami, ze kterých parametry získávají. Tyto parametry byly navrženy na základě provedené rešerše publikací z kapitoly 2. Výsledkem je tabulka *features.xlsx* obsahující celkově 14 parametrů pro každý subjekt.

3.5.1 Navržené parametry pro kvantifikaci respiračních poruch:

- **n_breath** = celkový počet nádechů během čtení textové úlohy.
- **breath_rate** = počet nádechů vztažený na celkovou dobu čtení textové úlohy.
- **speech_length_mean** = průměrná doba trvání řeči po nádechu.
- **speech_length_std** = směrodatná odchylka doby trvání řeči po nádechu vztažená k průměrné hodnotě.
- **breath_length_mean** = průměrná délka trvání nádechu.
- **breath_length_std** = směrodatná odchylka doby trvání nádechu vztažená k průměrné délce nádechu.
- **pause_before_mean** = průměrná délka pauzy mezi koncem řeči a začátkem nádechu.
- **pause_before_std** = směrodatná odchylka délky pauzy mezi koncem řeči a začátkem nádechu vztažená k průměrné hodnotě.

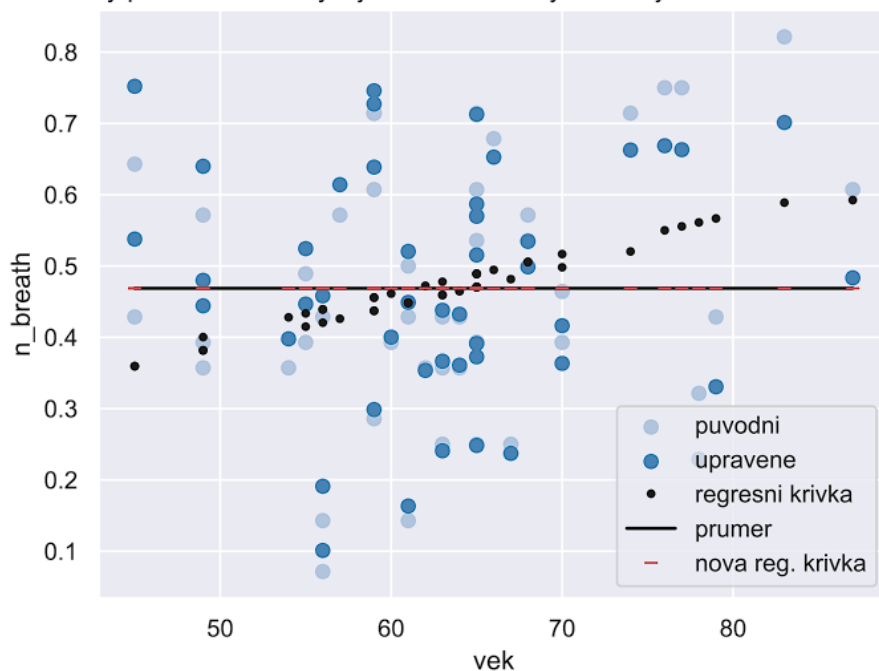
- **pause_after_mean** = průměrná délka pauzy mezi koncem nádechu a začátkem řeči.
- **pause_after_std** = směrodatná odchylka délky pauzy mezi koncem nádechu a začátkem řeči.
- **zcr_mean** = průměrná hodnota míry průchodu signálu nulou během nádechu.
- **E_speech_std** = směrodatná odchylka energie řeči po nádechu vztažená k průměrné hodnotě.
- **E_breath_std** = směrodatná odchylka energie nádechu vztažená k její průměrné hodnotě.
- **E_breath_norm_mean** = průměrná energie nádechu ve vztahu k energii řeči.

3.5.2 Statistická analýza

Navržené parametry pro kvantifikaci respiračních poruch u pacientů s PN, získané z časů predikovaných algoritmem pro automatizovanou detekci nádechů, byly podrobeny statistickým testům z oblasti statistické analýzy. Přesněji se jedná o test normality (Shapiro-Wilkův test) pro testování, zdali mají parametry normální rozdělení. Na základě výsledků testu normality byl dále zvolen neparametrický Mann-Whitney test. Posledním testem statistické analýzy byl zvolen výpočet Pearsonova korelačního koeficientu pro měření míry závislosti mezi dvěma veličinami.

Statistické zpracování výsledků a celková analýza byla provedena pomocí skriptů (viz příloha) ve vývojovém prostředí PyCharm pro programovací jazyk Python. Tabulka *features.xlsx* obsahující navržené parametry je v prvním kroku zpracována skriptem *adjust_features.py*, kterým se eliminuje vliv věku a pohlaví subjektů na jednotlivé parametry za užití lineární regrese (viz obr.7).

Hodnoty parametru se zvyšují s věkem a ženy dosahují nižších hodnot než muži.



Obrázek 7: upravené hodnoty parametru n_breath po eliminaci vlivu věku a pohlaví

Výsledkem skriptu je tabulka upravených parametrů *features_adj.xlsx* sloužící jako vstupní data pro další skripty provádějící statistické testy.

Pomocí skriptu *get_statistics.py* je vypočítán Shapiro-Wilkův test normálního rozdělení pro jednotlivé parametry a Mann-Whitneyův neparametrický test. Dále jsou pomocí skriptu vypočítány průměrné hodnoty, medián a směrodatná odchylka pro jednotlivé parametry a porovnání těchto parametrů mezi zdravými kontrolami a pacienty s PN. Posledním krokem tohoto skriptu je výpočet míry pacientů s PN, kteří leží mimo normální rozdělení na základě normálního rozložení hodnot zdravých kontrol, pro každý parametr. V rámci statistické analýzy byly vytvořené normy, založené na parametrech zdravých kontrol a bylo sledováno o kolik pacienti s PN překračují tyto normy. Byly stanovené spodní (T_{low}) a horní (T_{up}) prahové hodnoty pomocí rovnic:

$$T_{low} = Q_1 - 1,5 \cdot IQR, \quad (5)$$

$$T_{up} = Q_3 + 1,5 \cdot IQR, \quad (6)$$

kde Q_1 a Q_3 jsou spodní a horní kvartily a IQR je mezikvartilové rozpětí.

Výstupem skriptu je tabulka *stats_results.xlsx* obsahující výsledky výše uvedených statistických testů.

Dalším statistickým testem je výpočet pearsonova korelačního koeficientu, který měří sílu lineární závislosti mezi dvěma veličinami. Přesněji je zde měřena míra lineární závislosti navržených parametrů na klinických datech zkoumaných subjektů.

Výpočet pearsonova korelačního koeficientu je proveden pomocí skriptu *correlations.py*. Vstupními daty tohoto skriptu jsou tabulky klinických hodnot pacientů a tabulka s navrženými parametry *features_adj.xlsx*. Výsledkem je tabulka *pearson.xlsx* obsahující výsledky korelací jednotlivých parametrů s klinickými daty.

Pomocí skriptu *data_info.py* jsou vypočítána demografická data subjektů z datasetu a jejich průměrné hodnoty klinických dat, zejména pak data hodnotící skóre poruch řeči v jednotlivých oblastech a porovnání těchto hodnot mezi zdravými kontrolami a nemocnými pacienty.

3.5.3 Strojové učení

Skript *XGBoost.py* slouží k predikci klinických dat pacientů s PN a zdravých kontrol na základě strojového učení a využití algoritmu XGBoost. Vstupem tohoto skriptu je tabulka navržených parametrů, u kterých byl odstraněn vliv věku a pohlaví *features_adj.xlsx* a tabulka *clinical_data.xlsx* obsahující klinická data zkoumaných subjektů. Výstupem tohoto skriptu je dvojice tabulek.

První tabulka *ML_performance* obsahuje metriky mean absolute error (MAE), tedy průměrná absolutní chyba modelu a EER – estimated error rate měřící efektivitu modelu. Čím menší hodnota EER, tím přesnější prediktivní model je. Výpočet EER pomocí rovnice

$$EER = \frac{MAE}{R}, \quad (7)$$

kde R reprezentuje rozsah hodnot (stupnice klinických hodnot) v trénovacím setu.

Druhá tabulka *feature_importances.xlsx* obsahuje koeficienty jednotlivých parametrů pro každou ze skupin oblasti řeči a klinických dat, přesněji oblast fonorespirace, faciokineze, fonetika, dále pro hodnocení Parkinsonovy nemoci pomocí UPDRS III a MMSE nebo celkového DX indexu pro hodnocení HD. Model strojového učení byl nastaven na regresi a jeho validace proběhla využitím desetinasobné křížové validace s dvaceti opakováními.

4. VÝSLEDKY

Kapitola 4 obsahuje výsledky jednotlivých částí diplomové práce, kterými jsou vytvoření algoritmu pro automatizovanou detekci nádechů ze zvukových nahrávek, na kterých je zaznamenán úryvek čteného textu, který je v souladu s požadavky na textové úlohy v rámci hodnocení hypokinetické dysartrie.

Další částí diplomové práce je návrh parametrů pro kvantifikaci respiračních poruch u pacientů s Parkinsonovou nemocí a následná statistická analýza těchto parametrů

4.1 Automatizovaná detekce nádechů

Následující podkapitola obsahuje výsledky získané z jednotlivých částí vytvořeného algoritmu pro automatizovanou detekci nádechů v řečovém signálu.

Výsledkem první části je tabulka (tab. 5) vypočítaných parametrů řeči, které slouží jako vstupní data pro automatizovanou detekci nádechů.

Výsledkem druhé části je tabulka (tab.6) obsahující predikce nádechů společně s metrikami pro vyhodnocení modelu.

Výsledek poslední části obsahuje tabulku (tab.8) predikovaných a ručně anotovaných časů nádechu.

Z důvodu velkého množství dat jsou některé uvedené tabulky vyobrazeny pouze z části, kompletní tabulky s výsledky jsou dostupné v příloze.

Tabulka 4 obsahuje výsledky testování výběru vhodné kombinace parametrů pro zpracování řeči, které jsou popsány v kapitole 3.4.1. V tabulce jsou zobrazeny průměrné hodnoty metrik hodnotících kvalitu modelu pro detekci nádechů, pro různé kombinace parametrů řeči. Pro testování modelu byla použita desetinásobná křížová validace. Zvýrazněný řádek reprezentuje nejvhodnější kombinaci parametrů pro zpracování řeči.

	ACC_avg	AUC_avg	SEN_avg	SPEC_avg
14MFCC,E,ZC	0,8753	0,9147	0,516	0,9504
14MFCC,ZC	0,8721	0,912	0,503	0,9489
18MFCC,ZC	0,8719	0,9111	0,5041	0,9486
18MFCC,E,ZC	0,8751	0,9141	0,517	0,9501
14MFCC,E,ZC,LFS	0,8807	0,9193	0,5447	0,9513
18MFCC,E,ZC,LFS	0,8806	0,9193	0,5446	0,9512

Tabulka 4: Tabulka výsledků křížové validace pro různé parametry řeči

Tabulka 5 obsahuje data s počtem překrytí vzorků, které byly VAD označeny jako řeč se vzorky anotovaných nádechů, počet anotovaných nádechů a jejich průměrnou dobu. Dále jsou v tabulce uloženy data neřečových segmentů a vypočítané parametry řeči, které jsou popsány v kapitole 3.4.1. Posledním sloupcem tabulky je vektor segmentů s anotovanými nádechy.

ID	Počet překrytí [vzorky]	Počet nádechů	Průměrná doba nádechu [s]	Neřečové segmenty	MFCC	Energie	ZCR	ZCC	LSF	nádech
K1003	1555	15	0.3275	400x2567	2567x18	2567x1	2567x1	2567x1	2567x9	2567x1
K1004	1009	12	0.3775	400x681	681x18	681x1	681x1	681x1	681x9	681x1
K1005	0	15	0.2686	400x2020	2020x18	2020x1	2020x1	2020x1	2020x9	2020x1
K1006	1119	12	0.4243	400x1861	1861x18	1861x1	1861x1	1861x1	1861x9	1861x1
K1007	0	16	0.3813	400x2797	2797x18	2797x1	2797x1	2797x1	2797x9	2797x1
K1012	819	18	0.2803	400x4128	4128x18	4128x1	4128x1	4128x1	4128x9	4128x1

Tabulka 5: Ukázka tabulky vypočítaných parametrů řeči

Tabulka číslo 6 obsahuje výsledky logistické regrese společně s metrikami pro vyhodnocení modelu. Druhý sloupec tabulky obsahuje matici predikcí zkoumaných jevů v jednotlivých segmentech, třetí sloupec tabulky pak obsahuje vektor segmentů s predikovanými nádechy. V desátém sloupci se nacházejí matice záměn, zbylé sloupce tabulky obsahují metriky pro vyhodnocení kvality modelu.

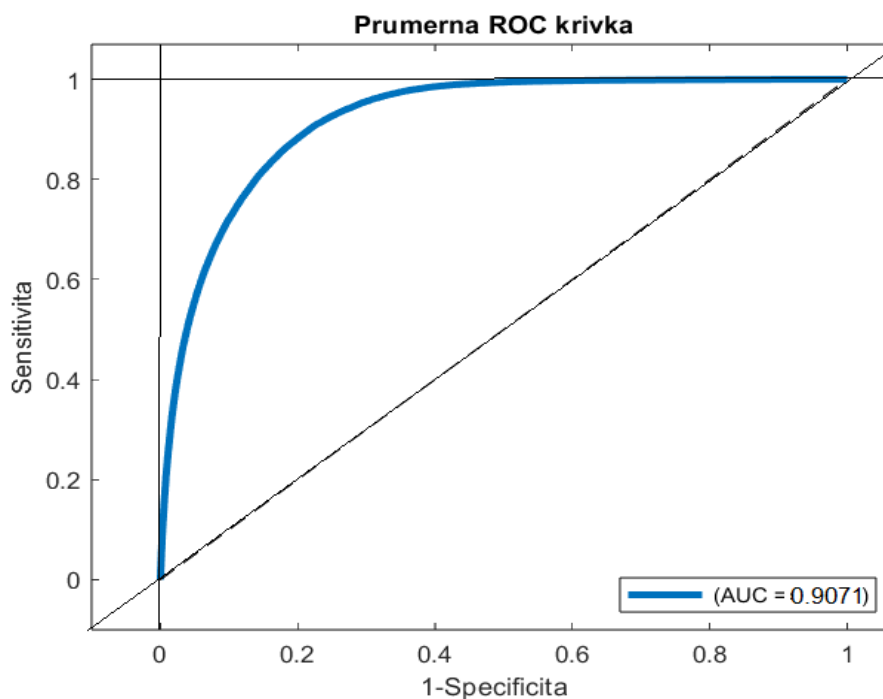
ID	pihat	yhat	AUC	ACC	TP	TN	FP	FN	CM	SEN	SPEC
P2029	4463x2	4463x1	0,93	0,88	164	3746	87	466	2x2	0,26	0,98
P2030	3149x2	3149x1	0,92	0,89	90	2721	23	315	2x2	0,22	0,99
P2032	3063x2	3063x1	0,97	0,94	66	2822	13	162	2x2	0,29	0,99
P2034	1387x2	1387x1	0,95	0,88	162	1058	15	152	2x2	0,52	0,99
P2037	4431x2	4431x1	0,94	0,90	456	3554	123	298	2x2	0,60	0,97
P2038	1919x2	1919x1	0,86	0,79	552	968	199	200	2x2	0,73	0,83
P2039	3011x2	3011x1	0,86	0,80	142	2271	92	506	2x2	0,22	0,96
P2042	1493x2	1493x1	0,68	0,66	143	841	57	452	2x2	0,24	0,94

Tabulka 6: Ukázka tabulky výsledků predikce nádechů a metrik pro vyhodnocení modelu

V tabulce 7 jsou statisticky porovnány výsledné hodnoty metrik pro vyhodnocení modelu, kterými jsou přesnost (ACC), plocha pod ROC křivkou (AUC), senzitivita (SEN) a specifická (SPEC). Tabulka obsahuje přehled o maximální a minimální hodnotě jednotlivých metrik, jejich průměrné hodnoty, medián a směrodatná odchylka.

	ACC		AUC		SEN		SPEC	
	Hodnota	ID	Hodnota	ID	Hodnota	ID	Hodnota	ID
MIN	0,6001	P2018	0,5801	P2005	0,0054	P2005	0,8068	P2055
MAX	0,9481	K2046	0,9818	K1024	0,7340	P2038	0,9993	P1031
Směrodatná odchylka	0,0792		0,0654		0,1837		0,0306	
Medián	0,8780		0,9255		0,3502		0,9850	
Průměr	0,8497		0,9071		0,3531		0,9741	

Tabulka 7: Výsledky vyhodnocení klasifikačního modelu pro predikci nádechů



Obrázek 8: Průměrná ROC křivka vytvořeného modelu

Dalším výsledkem je tabulka 8 obsahující časy detekovaných nádechů na základě predikce ve druhém sloupci, ve třetím sloupci jsou predikované časy po korekci, která zahrnovala odstranění nádechů kratších než 100 ms a omezení nádechů na maximální délku 1,5 sekund. Časy ručně označených nádechů se nachází ve čtvrtém sloupci. Hodnoty v pátém a šestém sloupci udávají počet, o kolik nádechů detekoval algoritmus více či méně oproti anotovaným nádechům před a po korekci.

ID	časy - predikce	časy – predikce po korekci	časy - anotace	počet nádechů - rozdíl	počet nádechů – rozdíl po korekci
K1024	16x2	11x2	30x1	1	-4
K1025	16x2	12x2	32x1	0	-4
K1026	15x2	12x2	32x1	-1	-4
K1029	17x2	16x2	32x1	1	0
K1031	19x2	17x2	32x1	3	1
K1036	18x2	16x2	26x1	5	3
K1037	20x2	20x2	34x1	3	3
K1039	12x2	12x2	24x1	0	0
K1040	24x2	20x2	40x1	4	0

Tabulka 8: Ukázka tabulky predikovaných a anotovaných časů nádechů

Tabulky 9 a 10 obsahují výsledky predikce nádechů a porovnání predikovaných (vlevo) a anotovaných (vpravo) časů nádechů v sekundách. V levém sloupci tabulek se nacházejí začátky časů nádechu a v pravém sloupci jejich konec. Výsledkem algoritmu pro predikci nádechů je kromě výše uvedené tabulky také lab soubor s predikovanými časy, který je vygenerován pro každý subjekt a odpovídá stejnému formátu jako má vstupní .lab soubor (viz obr. 4) s anotovanými nádechy. Na obrázku 9 je zobrazena ukázka výsledného souboru s predikovanými časy pro subjekt P1088.

časy nádechů – predikce [s]		
po korekci		
	začátek	konec
1.	4.8500	5.4375
2.	10.4370	10.9625
3.	13.9000	14.0750
4.	21.0625	21.4125
5.	24.1000	24.8250
6.	26.0500	27.2000
7.	28.6875	29.1750
8.	34.8625	35.2500
9.	38.7750	39.0000
10.	43.4375	43.8750
11.	46.3375	46.6250
12.	52.2625	52.5000

časy nádechů – anotace [s]		
	začátek	konec
1.	4.9914	5.3460
2.	10.3604	10.9176
3.	13.7540	14.1224
4.	16.0380	16.5721
5.	21.1353	21.4945
6.	24.6441	24.9480
7.	28.6915	29.1382
8.	34.9493	35.3683
9.	38.7343	39.0382
10.	43.4080	43.4080
11.	46.3596	46.7556
12.	52.1845	52.5022

Tabulka 9: Porovnání predikovaných časů (vlevo) a anotovaných časů (vpravo) nádechu pro subjekt K1006.

časy nádechů – predikce po korekci [s]		
	začátek	Konec
1.	4.7500	5.0375
2.	20.9375	21.8370
3.	25.2750	25.6125
4.	29.1750	29.6000
5.	32.2125	33.2125
6.	36.2125	37.0125
7.	43.9125	44.4875

časy nádechů – anotace [s]		
	začátek	konec
1.	7.2937	7.8325
2.	15.9090	16.5675
3.	19.3487	19.6112
4.	25.2657	25.8413
5.	29.1428	29.5895
6.	36.4043	37.1319
7.	40.0650	40.3137
8.	43.7764	44.4302
9.	46.0879	46.4241
10.	50.6097	50.9044

Tabulka 10: Porovnání predikovaných časů (vlevo) a ručně anotovaných časů (vpravo) v sekundách pro subjekt P2005

```

0.0000000 4.5875000 begin
4.5875000 4.8375000 end
4.8375000 17.1500000 begin
17.1500000 17.5625000 end
17.5625000 22.4750000 begin
22.4750000 22.8250000 end
22.8250000 27.3625000 begin
27.3625000 27.5625000 end
27.5625000 37.4375000 begin
37.4375000 37.7375000 end

```

Obrázek 9: Ukázka lab souboru s predikovanými časy pro subjekt P1088.

4.2 Statistická analýza

Následující část práce obsahuje výsledky výpočtu navržených parametrů pro kvantifikaci respiračních poruch pacientů s PN, které byly získány z audio nahrávek subjektů na základě časů nádechů, predikovaných pomocí výše popsaného algoritmu.

Dále jsou zde prezentovány výsledky statistické analýzy a statistických testů, kterými byly testovány a porovnávány navržené parametry.

Tabulka 11 zobrazuje navržené parametry pro kvantifikaci respiračních poruch, které byly popsány v kapitole 3.5.1. Zobrazeno jsou zde parametry pro prvních pět subjektů, obsah celé tabulky je dostupný v příloze. Vypočítané parametry jsou následně statisticky analyzovány (tab. 12 a 14) a slouží jako vstupní data modelu strojového učení pro predikci klinického skóre (tab. 15).

ID	n_breath [-]	breath_rate [-]	speech_length_mean [s]	speech_length_std [s]	breath_length_mean [s]
K1003	10	0,165	5,839	0,741	0,225
K1004	11	0,178	5,099	0,602	0,332
K1005	17	0,322	2,266	0,461	0,415
K1006	12	0,222	3,825	0,487	0,465
K1007	20	0,301	2,604	0,419	0,486

ID	breath_length_std [s]	pause_before_mean [s]	pause_before_std [s]	pause_after_mean [s]
K1003	0,236	0,014	2,024	0,084
K1004	0,573	0,196	0,879	0,082
K1005	0,682	0,111	0,769	0,032
K1006	0,581	0,118	0,417	0,061
K1007	0,889	0,106	0,725	0,092

ID	pause_after_std [s]	zcr_mean [-]	E_speech_std [-]	E_breath_std [-]	E_breath_norm_mean [-]
K1003	0,940	0,026	0,846	1,564	0,048
K1004	0,907	0,013	0,658	0,703	0,033
K1005	1,364	0,028	1,064	1,018	0,028
K1006	1,348	0,018	0,544	1,117	0,073
K1007	0,808	0,023	1,124	1,595	0,150

Tabulka 11: Navržené parametry pro kvantifikaci respiračních poruch (prvních pět subjektů)

Tabulka číslo 12 obsahuje výsledky statistických testů a následné analýzy. První sloupec obsahuje výsledky Shapiro-Wilkova testu, přesněji jsou zde zobrazeny p-hodnoty testu pro navržené parametry. Zeleně jsou vyznačeny p-hodnoty menší než 0.05. Ve druhém sloupci jsou zobrazeny p-hodnoty Mann-Whitneyho testu. Další sloupce obsahují porovnání průměrných hodnot, mediánu a směrodatných odchylek jednotlivých parametrů s porovnáním zdravých kontrol (ZK) a pacientů s Parkinsonovou nemocí (PN), sloupec se šipkou znázorňuje, zdali hodnota parametru pacientů s PN je větší či menší v porovnání se zdravými kontrolami. Poslední dva sloupce tabulky obsahují procentuální zastoupení pacientů s PN mimo normy dané parametry zdravých kontrol.

Tabulka 13 obsahuje porovnání výsledků Shapiro-Wilkova testu normality pro parametry, které byly získány z nahrávek na základě ruční anotace nádechů (vlevo) a pro parametry, jež byli získány na základě časů predikovaných pomocí algoritmu pro automatizovanou detekci nádechů (vpravo).

Na obrázku 10 jsou vyobrazeny dva histogramy znázorňující rozložení pacientů s PN a zdravých kontrol pro parametry *E_breath_norm_mean* a *zcr_mean*. Tyto parametry vykazovaly největší míru subjektů ležící mimo horní a dolní mez.

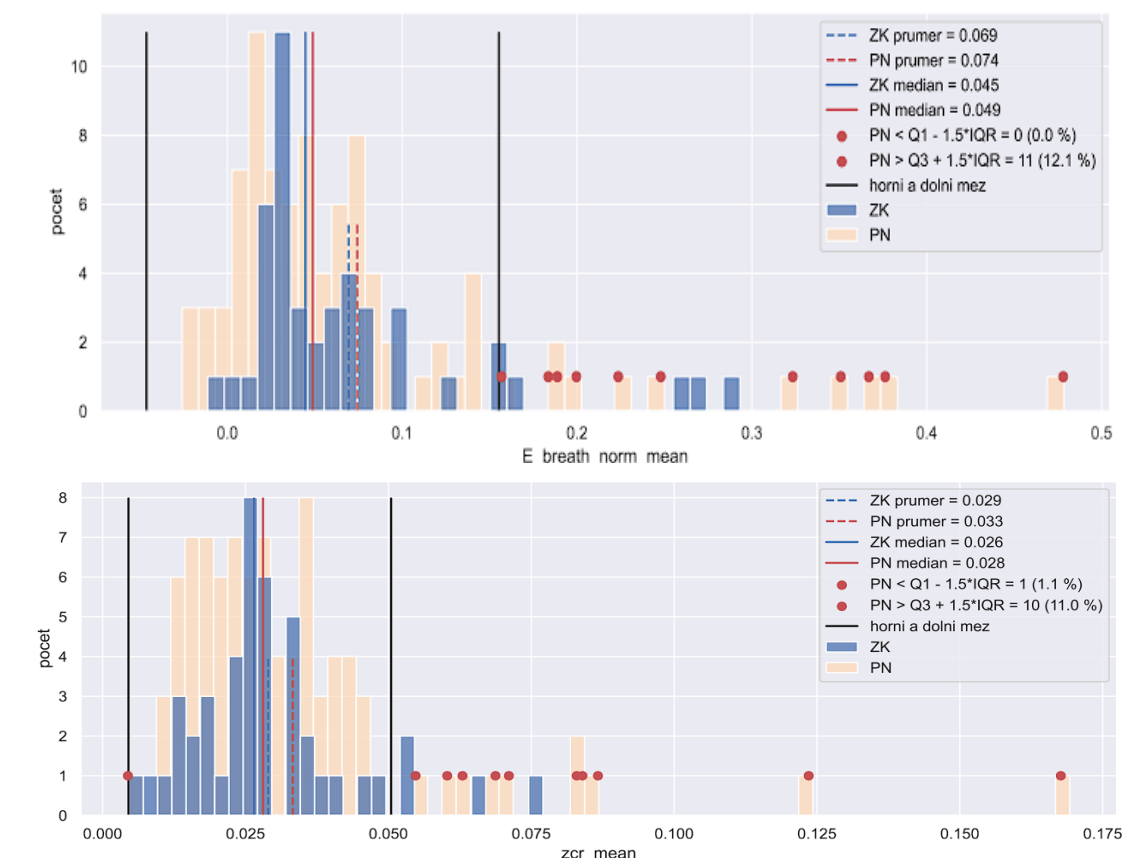
	Shapiro-Wilk p-hodnota	Mann-Whitney p-hodnota	průměr ZK	průměr PN	průměr
n_breath	0,405	0,901	13,089	13,247	↑
breath_rate	0,475	0,479	0,211	0,222	↑
speech_length_mean	<0,001	0,621	4,883	4,559	↓
speech_length_std	0,021	0,857	0,505	0,535	↑
breath_length_mean	0,005	0,416	0,464	0,484	↑
breath_length_std	0,928	0,298	0,668	0,699	↑
pause_before_mean	0,010	0,664	0,144	0,150	↑
pause_before_std	<0,001	0,868	0,896	0,872	↓
pause_after_mean	0,014	0,457	0,100	0,107	↑
pause_after_std	0,667	0,222	0,920	1,008	↑
zcr_mean	0,003	0,582	0,029	0,033	↑
E_speech_std	<0,001	0,511	0,934	0,940	↑
E_breath_std	0,019	0,718	1,488	1,574	↑
E_breath_norm_mean	<0,001	0,573	0,069	0,074	↑

	medián ZK	medián PN	medián	σ ZK	σ PN	σ	odlehlá hodnota - pod [%]	odlehlá hodnota - nad [%]
n_breath	12,573	13,034	↑	4,664	5,105	↑	0,000	0,000
breath_rate	0,219	0,230	↑	0,074	0,074	↓	1,099	1,099
speech_length_mean	3,858	3,714	↓	3,656	4,226	↑	0,000	4,396
speech_length_std	0,484	0,486	↑	0,163	0,231	↑	2,198	5,495
breath_length_mean	0,457	0,478	↑	0,160	0,182	↑	0,000	1,099
breath_length_std	0,660	0,724	↑	0,172	0,215	↑	0,000	0,000
pause_before_mean	0,135	0,138	↑	0,071	0,078	↑	0,000	5,495
pause_before_std	0,842	0,851	↑	0,365	0,264	↓	0,000	0,000
pause_after_mean	0,091	0,095	↑	0,051	0,059	↑	1,099	5,495
pause_after_std	0,911	0,968	↑	0,372	0,342	↓	0,000	1,111
zcr_mean	0,026	0,028	↑	0,014	0,024	↑	1,099	10,989
E_speech_std	0,766	0,861	↑	0,580	0,504	↓	0,000	5,495
E_breath_std	1,422	1,416	↓	0,540	0,700	↑	0,000	6,593
E_breath_norm_mean	0,045	0,049	↑	0,067	0,092	↑	0,000	12,088

Tabulka 12: Výsledky statistické analýzy

	Shapiro-Wilk – anotace	Shapiro-Wilk – predikce
n_breath	0,053	0,405
breath_rate	0,175	0,475
speech_length_mean	0,662	<0,001
speech_length_std	0,001	0,021
breath_length_mean	0,160	0,005
breath_length_std	0,009	0,928
pause_before_mean	0,622	0,010
pause_before_std	<0,001	<0,001
pause_after_mean	0,001	0,014
pause_after_std	0,021	0,667
zcr_mean	<0,001	0,003
E_speech_std	<0,001	<0,001
E_breath_std	<0,001	0,019
E_breath_norm_mean	<0,001	<0,001

Tabulka 13: Porovnání výsledků Shapiro-Wilkova testu, p-hodnoty



Obrázek 10: Rozložení parametrů s nejvyšším počtem pacientů s PN ležící mimo normy zdravých kontrol

V tabulce 14 jsou výsledky výpočtu Pearsonova korelačního koeficientu pro jednotlivé parametry pro zjištění míry korelace s klinickými daty pacientů. U každého parametru je vypočítána hodnota korelačního koeficientu a p-hodnota pro oblasti hodnocení klinického skóre. Zeleně jsou vyznačeny p-hodnoty menší než 0,05.

	UPDRS III		faciokineze		fonorespirace	
	koeficient	p-hodnota	koeficient	p-hodnota	koeficient	p-hodnota
n_breath	0,077	0,466	-0,121	0,161	-0,083	0,339
breath_rate	0,156	0,14	-0,092	0,286	-0,072	0,405
speech_length_mean	-0,121	0,254	0,108	0,213	-0,023	0,791
speech_length_std	-0,04	0,709	-0,099	0,254	-0,125	0,148
breath_length_mean	0,162	0,124	-0,195	0,023	-0,218	0,011
breath_length_std	-0,118	0,265	-0,014	0,867	-0,054	0,529
pause_before_mean	0,082	0,44	-0,067	0,438	-0,098	0,259
pause_before_std	-0,208	0,047	-0,16	0,063	-0,023	0,789
pause_after_mean	0,047	0,659	-0,187	0,029	-0,177	0,039
pause_after_std	0,02	0,852	-0,18	0,037	-0,124	0,152
zcr_mean	0,291	0,005	-0,145	0,093	-0,173	0,043
E_speech_std	0,039	0,712	-0,099	0,25	-0,066	0,448
E_breath_std	0,128	0,228	-0,14	0,103	-0,033	0,7
E_breath_norm_mean	0,248	0,018	-0,139	0,106	-0,113	0,19
	fonetika		celkový DX index		MMSE	
	koeficient	p-hodnota	koeficient	p-hodnota	koeficient	p-hodnota
n_breath	-0,052	0,547	-0,096	0,268	0,003	0,973
breath_rate	-0,029	0,737	-0,072	0,402	0,12	0,241
speech_length_mean	0,024	0,784	0,038	0,658	-0,018	0,864
speech_length_std	-0,133	0,123	-0,135	0,117	-0,224	0,027
breath_length_mean	-0,168	0,051	-0,22	0,01	-0,007	0,947
breath_length_std	-0,039	0,649	-0,042	0,63	0,176	0,083
pause_before_mean	-0,164	0,056	-0,125	0,147	-0,131	0,199
pause_before_std	-0,04	0,641	-0,082	0,345	-0,024	0,814
pause_after_mean	-0,268	0,002	-0,239	0,005	-0,232	0,022
pause_after_std	0,037	0,673	-0,099	0,253	0,027	0,791
zcr_mean	-0,108	0,211	-0,161	0,061	0,164	0,106
E_speech_std	-0,015	0,862	-0,067	0,437	0,034	0,736
E_breath_std	-0,088	0,309	-0,097	0,262	0,143	0,161
E_breath_norm_mean	-0,032	0,713	-0,106	0,217	0,061	0,55

Tabulka 14: Pearsonův korelační koeficient pro navržené parametry

4.3 Strojové učení

V této sekci jsou prezentovány výsledky modelu strojového učení pro predikci klinického skóre poruch řeči v jednotlivých oblastech na základě navržených parametrů zkoumaných subjektů.

	MAE	EER [%]
faciokineze	2,83 ± 0,65	13,47 ± 3,11
fonorespirace	2,85 ± 0,67	13,58 ± 3,21
fonetika	3,25 ± 0,74	18,04 ± 4,13
celkový DX index	7,82 ± 1,87	13,49 ± 3,23
UPDRS III	9,82 ± 2,36	18,89 ± 4,54
MMSE	1,62 ± 0,60	11,60 ± 4,32

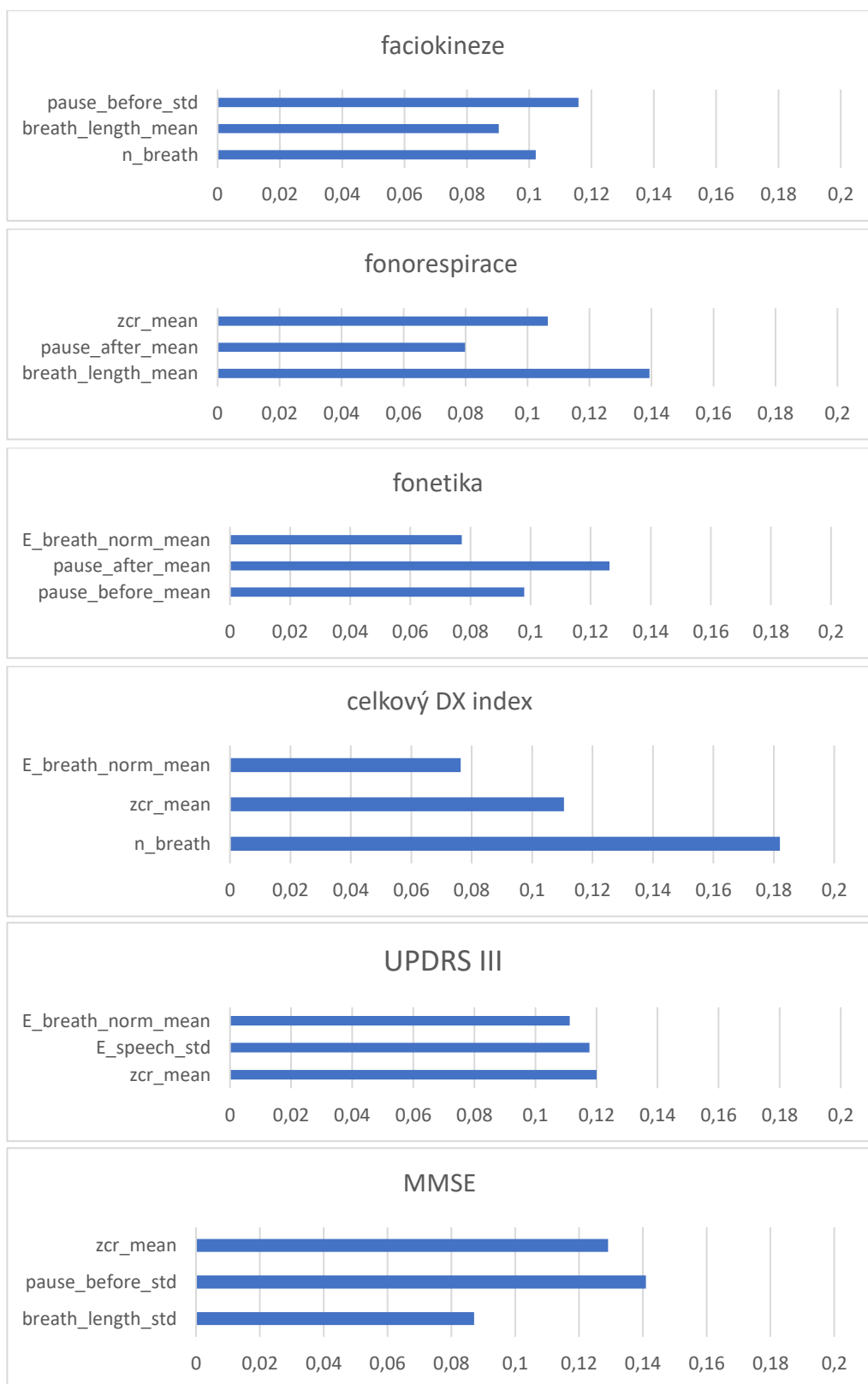
Tabulka 15: Hodnocení modelu strojového učení pro predikci klinických dat (průměr + směrodatná odchylka)

Tabulka 15 obsahuje výsledky hodnocení predikčního modelu pro predikci klinického skóre poruch hypokinetické dysartrie v různých oblastech, kterými jsou oblast faciokineze, fonorespirace, fonetiky, dále celkový DX index, UPDRS III a hodnocení MMSE. První sloupec obsahuje hodnotu průměrné absolutní chyby modelu a její směrodatnou odchylku. Ve druhém sloupci je zobrazena procentuální míra chybovosti predikčního modelu. V tabulce 16 jsou tyto výsledky porovnávány s výsledky, které byly získané na základě ručně anotovaných časů oproti časům, které byly vygenerovány pomocí predikčního modelu pro automatizovanou detekci nádechů.

Na obrázku číslo 9 je zobrazena série grafů, která zobrazuje vždy tři nejdůležitější parametry pro predikci klinického skóre v jednotlivých oblastech.

	MAE – predikované časy	MAE – anotované časy	EER [%] – predikované časy	EER [%] – anotované časy
faciokineze	2,83 ± 0,65	2,60 ± 0,61	13,47 ± 3,11	12,39 ± 2,89
fonorespirace	2,85 ± 0,67	2,78 ± 0,74	13,58 ± 3,21	13,25 ± 3,52
fonetika	3,25 ± 0,74	3,07 ± 0,78	18,04 ± 4,13	17,05 ± 4,33
celkový DX index	7,82 ± 1,87	6,88 ± 1,92	13,49 ± 3,23	11,86 ± 3,31
UPDRS III	9,82 ± 2,36	9,58 ± 2,29	18,89 ± 4,54	18,42 ± 4,41
MMSE	1,62 ± 0,60	1,67 ± 0,60	11,60 ± 4,32	11,91 ± 4,29

Tabulka 16: Porovnání výsledků strojového učení



Obrázek 11: Tři nejvýznamnější parametry pro predikci klinického skóre pro jednotlivé oblasti

5. VYHODNOCENÍ VÝSLEDKŮ, DISKUSE

5.1 Automatizovaná detekce nádechů

Pro vyřazení řečových segmentů byly testovány dva detektory řečové aktivity. První VAD (viz příloha *VAD_praat.m*) však vykazoval horší výsledky detekce řeči, míra překrytí vzorků detekované řeči a vzorků ručně označených nádechů byla v některých případech až desetkrát větší než u druhého VAD popsaného v kapitole 3.4.1.

Použitím VAD a následně vyřazením segmentů obsahujících řeč byla značně redukována velikost datasetu, to vedlo ke snížení nároků na výpočet a zvýšená rychlost procesu. Stejného výsledku zrychlení procesu bylo použitím VAD dosaženo v práci od K. M. I. Y. Arafath a A. Routray [12]. Použití VAD však není dokonalé. Část vzorků, které byly označeny při anotaci jako nádech, VAD i přes snahu optimalizovat vstupní parametry označil jako řeč. Naopak vzorky, které výstup VAD označil jako neřečové obsahují řečový signál. Pro účely práce je však toto použití dostačující.

Výpočet a výběr parametrů řeči byl proveden na základě rešerší prací z kapitoly 2 a testování datasetu na klasifikačním modelu pomocí 10násobné křížové validace. Dataset byl rozdělen na 10 stejně velkých objektů, kdy vždy jeden objekt sloužil jako testovací a zbytek datasetu sloužil k trénování modelu.

Z tabulky 4 je patrné, že nejlepších výsledků algoritmu pro detekci nádechů bylo dosaženo za použití těchto parametrů řeči: 14MFCC, energie signálu, parametrů zero-crossing rate/count a 9LFS, ze kterých byl vytvořen dataset pro klasifikační model. Pro lepší detekci nádechů a například vynechání fáze ruční anotace by mohl být při čtení textu použit kontaktní monitor nádechu v podobě hrudního pásu či jiný druh detektoru dechu.

Pouze u 32 subjektů z celkových 136 bylo zajištěno nulové překrytí vzorků výstupu VAD a vzorků označených jako nádech pomocí anotovaných časů. To znamená, že VAD vyhodnotil vzorky nádechů jako neřečové vzorky. Průměrný počet překrývajících se vzorků je 722. Průměrná doba nádechů zdravých kontrol je 0,3901 sekund, u pacientů s PN je průměrná doba nádechu rovna 0,3958 sekund.

Tabulka 6 obsahuje minimální a maximální hodnoty jednotlivých metrik pro vyhodnocení predikčního modelu společně s průměrnými hodnotami, mediánem a směrodatnou odchylkou těchto metrik získaných na základě dat všech zkoumaných subjektů.

Průměrná hodnota přesnosti vykazuje dobrý výsledek klasifikačního modelu, nejvyšší dosažená přesnost byla 0,9481 u subjektu K2046. Také průměrná hodnota oblasti pod ROC křivkou ($AUC = 0,9071$) výsledného modelu značí dobrý výsledek (obr.8). Specificita modelu s průměrnou hodnotou 0,9342 je vysoká oproti průměrné sensitivitě. Tento výsledek je zapříčiněn nevyvážeností datasetu, kdy se operuje s mnohem větším počtem segmentů mimo respiraci, v tomto důsledku jev true výrazně převyšuje nad

druhým jevem false. Model tak dokáže úspěšně a přesně predikovat, kdy se v signálu nádech nevyskytuje.

Nejhorších výsledků klasifikace nádechů bylo dosaženo u subjektu P2005, kdy hodnota AUC je rovna 0,5801 a sensitivita dosahovala hodnoty pouhých 0,0054. Po zpětné kontrole nahrávky tohoto subjektu bylo zjištěno, že nádechy jsou velice krátké a mělké, téměř neslyšitelné. Naopak v nahrávce subjektu K2046, u kterého byla přesnost modelu nejvyšší jsou nádechy silné a zřetelné.

Z tohoto faktu plyne, že pro správnou detekci a predikci nádechů z nahrávek řeči je důležitý nejenom kvalitně vytvořený dataset a algoritmus klasifikačního modelu, ale hlavně kvalitní nahrávka samotná.

Výraznou limitací práce byla délka trvání celkového výpočtu predikce nádechů, přesněji fáze trénování modelu a následné vyhodnocení pomocí původní metody upravené křížové validace, výpočetní doba skriptu Log_Reg.m trvala téměř 14 hodin. Dlouhá doba výpočtu, i v případě vyřazení řečových segmentů signálu je pravděpodobně zapříčiněna velkým množstvím dat ve vstupním datasetu.

V pozdější fázi práce byl model znovu přeučten tak, že data pro trénování modelu byla vybrána pouze z řečových parametrů zdravých kontrol oproti původnímu modelu, kdy jako trénovací data sloužili parametry řeči všech zkoumaných subjektů vyjma dat testovaného subjektu. Kromě zlepšení výsledků predikčního modelu vedl tento krok ke zkrácení délky trvání skriptu o více jak 80 %.

Porovnáním predikovaných časů nádechů s časy ručně anotovanými lze pozorovat, že model kvalitně detekuje nádechy odpovídající nádechům, které byly zaznačeny ručně (tab.9). I v případě subjektu, pro který vykazoval model nejhorší hodnocení (tab.10).

U řady subjektů se objevovaly predikované nádechy, které nebyly zaznamenány pomocí ruční anotace. V některých případech byly predikovány nádechy, které byly dlouhé pouhých 12,5 – 50 ms nebo byly jako nádech označeny úseky řeči. K tomuto jevu docházelo nejvíce v případě vyslovení hlásek „ch“ nebo „s“, případně u dalších slov obsahující výrazné frikativní hlásky. Další jev, který při detekci nádechů nastával, bylo správné označení začátku nádechu, ale označení konce bylo chybné a společně s nádechem byl označen i kus řečového signálu nebo naopak. Částečně byly výše uvedené jevy eliminovány korekcí, při které byly odstraněny nádechy kratší než 100 ms, maximální délka nádechu byla zvolena 1,5 sekundy a rozhodovací úroveň predikčního modelu byla nastavena na hodnotu 0,7. Ne ve všech případech však došlo k úplné eliminaci a objevují se detekované nádechy, které jsou ve skutečnosti řečový signál.

Po korekci model pro automatizovanou detekci nádechů detekoval průměrně o 2 nádechy méně, než bylo zaznamenáno při ruční anotaci. V některých případech však bylo automaticky detekováno mnohem méně či více nádechů. V případě subjektu K2047 byly detekovány pouze dva nádechy, což je o 21 nádechů méně než v případě ruční anotace. U subjektu P1109 bylo naopak detekováno o 10 nádechů více než v případě ruční anotace.

V řádku 5, v levé části tabulky 10 lze pozorovat nádech, který nebyl zaznamenán během ruční anotace. Po kontrolním poslechu v programu wavesurfer bylo zjištěno, že na těchto místech skutečně došlo s největší pravděpodobností k nádechu, ovšem nádech nebyl označen příslušnými labely, jelikož nebyl dostatečně zřetelný.

Na základě uvedených výsledků lze říct, že byl vytvořen funkční algoritmus pro detekci nádechů z řečového signálu s dostatečně kvalitními výsledky. Průměrná přesnost modelu je rovna 0,8497 a průměrná hodnota oblasti pod ROC křivkou je rovna 0,907. Chybu algoritmu v podobě vyhodnocení příliš krátkých, falešných nádechů byly odstraněny a pomocí úpravy algoritmu došlo ke zlepšení celkové detekce nádechů.

Algoritmus také detekuje nádechy, které byly v řeči provedeny, avšak z určitých důvodů nebyly anotovány.

Nevýhodou vytvořeného algoritmu je občasná detekce nádechů, které obsahují řečový signál. Další nevýhodou je v některých případech malý počet predikovaných nádechů oproti ručním anotacím. V těchto případech však bylo po kontrole zjištěno, že i přes malé množství predikovaných nádechů je predikce velmi přesná.

Algoritmus pro automatizovanou detekci nádechů z nahrávek subjektů byl úspěšně použit a na základě predikovaných časů mohlo dojít k extrakci navržených parametrů pro kvantifikaci respiračních poruch.

5.2 Statistická analýza

Na základě získaných znalostí v oblasti respiračních poruch a dysfunkcí, ve vztahu k hypokinetické dysartrii a parkinsonovy nemoci, bylo navrženo 14 parametrů které by mohly být užitečné pro kvantifikaci respiračních obtíží. Tyto parametry byly následně statisticky analyzovány pomocí statistických testů.

Statistické testy (tab.12) odhalily, že pouze 4 parametry vykazují normální rozložení v rámci skupiny zdravých kontrol. Těmito parametry jsou: *n_breath*, *breath_rate*, *breath_length_std* a *pause_after_std*. Z tohoto důvodu je u pacientů s Parkinsonovou nemocí u těchto parametrů menší pravděpodobnost, že se vymykají mezím z důvodů jiných než hypokinetická dysartrie. Průměrný počet nádechů u zdravých kontrol (*n_breath*) byl roven 13,089 a u pacientů s PN 13,247. Tento výsledek je téměř stejný i když pacienti s PN vykazují průměrně pouze o 0,158 nádechů více během čtení textu. Podle předpokladů na základě zkoumané literatury [3][5] by se dal předpokládat větší počet nádechů u pacientů s PN než u zdravých kontrol. Tento jev může být zapříčiněn nedokonalostmi algoritmu pro automatizovanou detekci nádechů, kdy u pacientů s PN byla detekce horší než u zdravých kontrol. Dalším z důvodů může být zvýšená pozornost zkoumaného subjektu na správnost čtení textu, tedy i dodržování správných pozic pauzy.

Při porovnání parametrů zdravých kontrol a pacientů s PN nebyly zjištěny žádné významné rozdíly či artefakty, což odpovídá výsledkům studie Hlavničky [20]. Je potřeba zmínit, že celkový DX index (tab. 3) pacientů s PN dosahuje v průměru hodnoty 74,4

bodů, což je podle tabulkových hodnot hodnoceno jako pomezí mezi lehkou poruchou a velmi mírnou dysartrií, která je případně hodnocena jen jako prosté snížení motorických schopností bez patologické příčiny. Průměrné hodnoty DX indexu pro zdravé kontroly rovny 85.3 jsou zase na pomezí lehké dysartrie a řeči bez poruchy [23].

Přesto existuje několik znaků, s nimiž se pacienti s PN odchylojí od norem zdravých kontrol. V tomto smyslu dosahují nejlepších výsledků parametr *zcr_mean* tedy průměrná míra průchodu signálu nulovou hodnotou a parametr *E_breath_norm_mean*, který odpovídá průměrné normalizované energii nádechu. V případě parametru *zcr_mean* se téměř 11 % pacientů s PN potýká s vyšším kolísáním signálu, což může značit přítomnost stridoru, který je definován jako drsný, chraptavý či pískavý rušivý zvuk během respirace. Až 12 % pacientů s PN dosahuje větších hodnot průměrné energie nádechu, což může odpovídat hlubokým nádechům, při kterých se snaží pacienti dostat do plic více vzduchu. Dále to může značit přítomnost dalších rušivých zvuků během nádechu.

Pearsonův korelační test (tab. 14) ukazuje, že existuje významná lineární závislost mezi parametry kvantifikující míru průchodu signálu nulovou hodnotou (*zcr_mean*), průměrnou délku nádechu (*breath_length_mean*) a průměrnou délku pauzy mezi nádechem a začátkem řeči (*pause_after_mean*) a klinickým skóre hodnotící oblast fonorespirace. Parametry kvantifikující průměrnou délku nádechu a průměrnou délku pauzy mezi nádechem a začátkem řeči korelují s klinickým hodnocením testu fonetiky. Hodnocení UPDRS III koreluje s parametry *zcr_mean*, *E_breath_norm_mean* a parametrem *pause_before_std*. Pro hodnocení délky trvání nemoci nevykazuje korelaci žádný z navržených parametrů.

Pomocí modelu strojového učení bylo možno do určité míry predikovat klinické skóre na základě extrahovaných parametrů. Pro predikci klinického skóre fonorespirace bylo dosaženo EER 13,58 %. V rámci predikce skóre fonorespirace byly nejdůležitější parametry *zcr_mean*, *pause_after_mean* a *breath_length_mean* (obr. 11). V rámci predikce celkového dysartrického indexu (DX index) bylo dosaženo EER 13,49 % a mezi nejdůležitějšími parametry se opět objevuje *zcr_mean*, tedy průměrná hodnota průchodu signálu nulovou hodnotou, dále parametry *n_breath* udávající celkový počet nádechů a parametr *E_breath_norm_mean*. Pro predikci UPDRS III skóre byly opět nejdůležitější parametry *zcr_mean* a *E_breath_norm_mean*. Predikce skóre MMSE dosahovalo v rámci hodnocení modelu nejlepších výsledků. Celkově byl pro predikci klinického skóre nejdůležitější parametr *zcr_mean*.

Z výsledků statistické analýzy a strojového učení lze říct, že nebyly nalezeny významné odlišnosti mezi parametry zdravých kontrol a pacientů s Parkinsonovou nemocí, podobně jako v práci Dr. Hlavničky [20]. Parametry vykazující největší odchylku od norem byla průměrná hodnota průchodu signálu nulovou hodnotou (*zcr_mean*) a průměrná normalizovaná energie nádechu (*E_breath_norm_mean*). Navzdory očekávání, parametry kvantifikující počet nádechů (*n_breath*), míru nádechů během čteného textu (*breath_rate*) a průměrná délka nádechu (*breath_length_mean*),

nevykazovaly výrazné odchylky u pacientů s PN oproti zdravým kontrolám. V rámci statistické analýzy v práci Dr. Hlavničky [20] byly objeveny mírné výchylky pacientů s PN od norem parametrů zdravých kontrol. Těmito parametry byla relativní energie nádechu, míra respirace udávající počet nádechů v určitém intervalu a délka pauzy mezi nádechem a začátkem či koncem řeči.

Na základě získaných parametrů bylo možné do určité míry predikovat hodnoty klinického skóre zkoumaných subjektů. V rámci této predikce byl parametr *n_breath* druhým nejdůležitějším parametrem modelu. Stěžejním parametrem pro predikci klinického skóre byla průměrná hodnota průchodu signálu nulovou hodnotou. Tento parametr také vykazoval významnou lineární závislost s klinickým skóre hodnotící oblast fonorespirace a hodnocení UPDRS III.

Na základě výsledků lze usoudit, že parametry *zcr_mean*, *E_breath_norm_mean*, *n_breath*, *pause_after_mean* a *breath_length_mean* by mohly být použity jako parametry pro objektivní hodnocení a kvantifikaci respiračních obtíží.

5.3 Porovnání detekovaných nádechů s ruční anotací

V rámci diplomové práce byly navržené parametry extrahovány z nahrávek i na základě ručně anotovaných časů nádechů a následně u nich byly provedeny statistické testy a predikce klinického skóre. Výsledky těchto testů pak byly porovnány s výsledky statistických testů parametrů získaných pomocí algoritmu pro automatizovanou detekci nádechů.

Statistické testy ukázaly, že normálního rozdělení ve skupině zdravých kontrol bylo dosaženo také pouze u čtyřech parametrů (tab. 13) a to u parametrů *breath_rate*, *speech_length_mean*, *breath_length_mean* a *pause_before_mean*. Ani u parametrů získaných na základě ručně anotovaných nádechů nebyly při porovnání zjištěny významnější odchylky či rozdíly mezi pacienty trpící PN a zdravými kontrolami.

U parametru *zcr_mean* bylo pozorováno, že 10 % pacientů s PN vykazovalo větší odchylky průchodu signálu nulovou hodnotou, což odpovídá výsledkům získaných ze statistických testů parametrů, které byly získány na základě detekovaných nádechů. U parametru *breath_rate* vykazovalo 8 % pacientů vyšší hodnoty počtu nádechů během čteného textu, což by mohlo odpovídat znakům zvýšené frekvence dýchání či znakům hyperventilace. Tento parametr vykazuje normální rozdělení ve skupině zdravých kontrol, má tak proto potenciál pro objektivní hodnocení hypokinetické dysartrie. U obou parametrů je u osob s PN vyšší průměrná hodnota, medián i směrodatná odchylka.

Pearsonův korelační test ukazuje významnou lineární závislost parametru *n_breath*, *breath_rate* a *zcr_mean* a klinického testu hodnotící oblast fonorespirace. Pro oblast fonetiky vykazují lineární závislost parametr *breath_length_mean* kvantifikující průměrnou délku nádechu a parametr *pause_after_mean*, tedy pauza mezi nádechem a začátkem řeči. Zajímavostí je silná negativní korelace mezi parametry *pause_after_mean*,

speech_length_mean a výsledky MMSE, což může značit, že lidem s kognitivními poruchami trvá déle, než po nádechu začnou mluvit a intervaly mezi nádechy jsou delší.

Porovnáním výsledků strojového učení pro predikci klinického skóre na základě parametrů extrahovaných pomocí detekovaných nádechů a výsledků, které byly získány pomocí parametrů extrahovaných z ručně anotovaných nádechů lze říct, že vyhodnocení predikčního modelu má velmi podobné výsledky (tab. 16). V rámci predikce fonorespirace bylo dosaženo EER 13,25 %, což je o pouhých 0,33 % méně než v případě predikce na základě algoritmem detekovaných nádechů. Téměř totožných výsledků bylo dosaženo také v predikci klinického skóre UPDRS III. Celkově byly pro predikci klinických dat nejdůležitější parametry *pause_after_mean*, *zcr_mean* a *breath_rate*. Parametr kvantifikující průměrnou hodnotu průchodu signálu nulovou hodnotou *zcr_mean* byl důležitý v rámci predikce klinického skóre na základě parametrů extrahovaných z detekovaných nádechů.

Shrnutím výsledků strojového učení a statistických testů lze odvodit, že pro objektivní hodnocení a kvantifikaci respiračních poruch u pacientů s PN by mohl být užitečný parametr *zcr_mean* a parametr *breath_rate*.

5.4 Limitace práce

Limitací diplomové práce je poměrně malý a nevyvážený počet zkoumaných subjektů, v případě pacientů s Parkinsonovou nemocí je průměrná hodnota celkového dysartrického indexu rovna 74,4 což znamená pouze velmi lehkou dysartrii. Další limitací je fakt, že všichni pacienti s PN, až na subjekt P1079, byly pod vlivem léku snižující dopady nemoci. To mohlo přímo ovlivnit parametry, které byly z nahrávek řeči extrahovány. Poslední limitací byla časově náročná ruční anotace nádechů, která může být do jisté míry subjektivně ovlivněna.

V návaznosti na tuto práci může být zdokonalen algoritmus pro automatizovanou detekci nádechů, kdy bude algoritmus detekovat pouze čisté nádechy bez úseků řeči, což by vedlo k přesnějším výsledkům a eliminaci zdlouhavé ruční anotace. Dále by bylo vhodné rozšíření počtu respondentů o pacienty, jejichž dysartrie je více rozvinutá, případně zkoumat vliv medikace na respirační obtíže.

6. ZÁVĚR

Výsledkem této práce je funkční algoritmus pro automatickou detekci nádechů v nahrávkách řeči zkoumaných subjektů. Vytvořený klasifikační model pracuje na základě logistické regrese, kvalita modelu je vyhodnocena metrikami získaných pomocí křížové validace. Průměrná přesnost modelu byla 0,85, průměrné hodnoty AUC a SPEC byly větší než 0.9. Průměrná sensitivita modelu byla menší než 0,4 což bylo způsobeno nevyvážeností tříd true a false ve vstupním datasetu.

Časy predikovaných nádechů odpovídají časům získaným při ruční anotaci, a to i v případě subjektu, pro který model vykazoval nejnižší přesnost. Algoritmus detekuje i nádechy, které nebyly při anotaci zaznamenány z důvodu nežádoucích zvuků v okolí nebo nedostatečné zřetelnosti nádechů. V některých případech však algoritmus detekuje jako nádech úsek, který obsahuje řeč, případně chybně detekuje začátek či konec nádechu. Na základě predikovaných časů nádechů bylo z řečových nahrávek respondentů extrahováno 14 parametrů. Tyto parametry byly následně statisticky analyzovány pomocí statistických testů. Pomocí strojového učení bylo možno z extrahovaných parametrů do jisté míry predikovat klinické skóre zkoumaných subjektů. V rámci práce byly navržené parametry z nahrávek extrahovány také na základě ručně anotovaných nádechů, výsledky pak byly porovnány s výsledky, po automatické detekci nádechů

Z výsledků bylo možné odvodit parametry, které by mohly být užitečné pro objektivní kvantifikaci respiračních poruch u pacientů s Parkinsonovou nemocí. Těmito parametry jsou parametry *zcr_mean*, *E_breath_norm_mean*, *n_breath*, *pause_after_mean*, *breath_length_mean* a parametr *breath_rate*. Parametr *zcr_mean* značí míru průchodu signálu nulovou hodnotou, jenž může být ovlivněna tzv. stridorem, tedy pískavým či rušivým zvukem, který způsobuje větší fluktuaci signálu. Parametry *n_breath* a *breath_rate* reprezentují počet nádechů a počet nádechů za dobu čtení textu, kdy pacienti trpící HD mají vlivem nemoci tendenci k častějším nádechům. Průměrná délka nádechu může detekovat krátké, mělké nádechy, kdy pacienti vlivem třesu a ztuhlosti svalstva nemohou dostatečně zaplnit plíce kyslíkem nebo naopak detekovat dlouhé nádechy, kterými se snaží pacienti tento deficit vyrovnat. Průměrná délka pauzy mezi nádechem a začátkem řeči lze využít pro hodnocení problémů kognitivních funkcí spojených s PN.

Závěrem lze říct, že byl sestaven funkční algoritmus pro automatizovanou detekci nádechů z řečového signálu a byly navrženy parametry pro kvantifikaci respiračních potíží u pacientů s Parkinsonovu nemocí, trpící hypokinetickou dysartrií.

LITERATURA

- [1] PARKINSON, James. An Essay on the Shaking Palsy. *The Journal of Neuropsychiatry and Clinical Neurosciences* [online]. 2002, 14(2), 223-236 [cit. 2022-11-28]. ISSN 0895-0172. Dostupné z: doi:10.1176/jnp.14.2.223
- [2] BAREŠ, PH.D., MUDr. Martin. Diagnostika a klinické příznaky Parkinsonovy nemoci. *Neurologie pro praxi* [online]. 2001, 2001, 2001(1), 22-24 [cit. 2022-11-28]. Dostupné z: https://www.internimedica.cz/artkey/neu-200101-0008_diagnostika_a_klinicke_priznaky_parkinsonovy_nemoci.php.
- [3] ZAMIŠKOVÁ, G., RESSNER, P., DLOUHÁ, J, ŠIGUTOVÁ, D. Poruchy řeči u Parkinsonovy nemoci. *Neurologie pro praxi* [online]. Konice: Solen, 2010, roč. 11, č. 2, s. 112–116 [cit. 2015–04–20]. Dostupné
- [4] KLÍŠŤOVÁ, Jana. Poruchy dechových funkcí u osob s Parkinsonovou chorobou a jejich ovlivnění pomocí respirační fyzioterapie. Olomouc, 2008, 92 s. Diplomové práce. Univerzita Palackého, Katedra fyzioterapie. Vedoucí práce Dagmar Bezděková.
- [5] HORÁKOVÁ, Michaela. Narušená komunikační schopnost při Parkinsonově chorobě [online]. Brno, 2017 [cit. 2022-11-28]. Available from: <https://is.muni.cz/th/mibdi/>. Master's thesis. Masaryk University, Faculty of Education. Thesis supervisor Petr KOPEČNÝ.
- [6] MEKYSKA, Jiří, Zdeněk SMĚKAL, Milena KOŠŤÁLOVÁ, Martina MRAČKOVÁ, Světlana SKUTILOVÁ a Irena REKTOROVÁ. Motorické aspekty poruch řeči u Parkinsonovy nemoci a jejich hodnocení. *Česká a slovenská neurologie a neurochirurgie*. 2011, roč. 74, č. 6, s. 662-668. ISSN 1210-7859.
- [7] RUSZ, Jan, Tereza TYKALOVA, Michal NOVOTNY, David ZOGALA, Karel SONKA, Evzen RUZICKA a Petr DUSEK. Defining Speech Subtypes in De Novo Parkinson Disease. *Neurology* [online]. 2021, **97**(21), e2124-e2135 [cit. 2022-12-05]. ISSN 0028-3878. Dostupné z: doi:10.1212/WNL.00000000000012878
- [8] D'ARRIGO, Andrea, Stefano FLORO, Francesca BARTESAGHI, Chiara CASELLATO, Giuseppe Francesco SFERRAZZA PAPA, Stefano CENTANNI, Alberto PRIORI a Tommaso BOCCI. Respiratory dysfunction in Parkinson's disease: a narrative review. *ERJ Open Research* [online]. 2020, 6(4), /erjor/6/4/00165-2020.atom [cit. 2023-04-24]. ISSN 2312-0541. Dostupné z: doi:10.1183/23120541.00165-2020

- [9] MCGOVERN, Katherine. Acoustic Intensity and Speech Breathing Kinematics in a Patient with Parkinson's Disease. 2020. PhD Thesis. University of Pittsburgh
- [10] DOCU AXELERAD, Any, Alina Zorina STROE, Oana Cristina ARGHIR, Daniel DOCU AXELERAD a Anca Elena GOGU. Respiratory Dysfunctions in Parkinson's Disease Patients. *Brain Sciences* [online]. 2021, 11(5) [cit. 2023-04-24]. ISSN 2076-3425. Dostupné z: doi:10.3390/brainsci11050595
- [11] FURUI, Sadaoki. Digital speech processing: synthesis, and recognition. CRC Press, 2018.
- [12] Yasar Arafath K., Mohamed & Routray, Aurobinda. (2019). Automatic Detection of Breath Using Voice Activity Detection and SVM Classifier with Application on News Reports. 609-613. 10.21437/Interspeech.2019-2434.
- [13] NALLANTHIGHAL, Venkata Srikanth; STRIK, H. Deep sensing of breathing signal during conversational speech. 2019.
- [14] PONOMARCHUK, Alexander, Ilya BURENKO, Elian MALKIN, Ivan NAZAROV, Vladimir KOKH, Manvel AVETISIAN a Leonid ZHUKOV. Project Achoo: A Practical Model and Application for COVID-19 Detection From Recordings of Breath, Voice, and Cough. *IEEE Journal of Selected Topics in Signal Processing* [online]. 2022, 16(2), 175-187 [cit. 2022-11-29]. ISSN 1932-4553. Dostupné z: doi:10.1109/JSTSP.2022.3142514
- [15] DESHPANDE, Gauri; SCHULLER, Björn. An overview on audio, signal, speech, & language processing for COVID-19. arXiv preprint arXiv:2005.08579, 2020.
- [16] A. Routray et al., "Automatic measurement of speech breathing rate," in 2019 27th European Signal Processing Conference (EU-SIPCO). IEEE, 2019, pp. 1–5
- [17] V. S. Nallanthighal, A. H`arm`a, and H. Strik, "Deep sensing of breathing signal during conversational speech," in Proceedings of the 20th Annual Conference of the International Speech Communication Association, Interspeech. Graz, Austria:[Sn], 2019, pp. 4110–4114
- [18] A. Imran, I. Posokhova, H. N. Qureshi, U. Masood, S. Riaz, K. Ali, C. N. John, and M. Nabeel, "Ai4covid-19: Ai enabled preliminary diagnosis for covid-19 from cough samples via an app," arXiv preprint arXiv:2004.01275, 2020
- [19] A. Martin, J. Nateqi, S. Gruarin, N. Munsch, I. Abdarahmane, and B. Knapp, "An artificial intelligence-based first-line defence against covid-19: digitally screening citizens for risks via a chatbot," bioRxiv, 2020.
- [20] Hlavnička, J., Čmejla, R., Tykalová, T. et al. Automated analysis of connected speech reveals early biomarkers of Parkinson's disease in patients with rapid eye movement sleep behaviour disorder. *Sci Rep* 7, 12 (2017). <https://doi.org/10.1038/s41598-017-00047-5>
- [21] SMĚKAL, Z.: Číslicové zpracování řeči (MZPR). Elektronická skripta pro magisterská studia, VUT Brno, 2010.

- [22] Zero-crossing rate - Introduction to Speech Processing. Zero-crossing rate - Introduction to Speech Processing - Aalto University Wiki [online]. Finsko, 2019 [cit. 2022-12-05]. Dostupné z: <https://wiki.aalto.fi/display/ITSP/Zero-crossing+rate>
- [23] M. Kostalova, M. Mrackova, R. Marecek, D. Berankova, I. Eliasova, E. Janousova, J. Roubickova, J. Bednarik, and I. Rektorova, “Test 3f dysartricky profil–normativni hodnoty reci v cestine,” Ceska a Slovenska Neurologie a Neurochirurgie, vol. 76, no. 109, p. 5, 2013
- [24] <https://sourceforge.net/projects/wavesurfer/>
- [25] STEIN, Jonathan Y. Digital signal processing: a computer science perspective [online]. In: . 2000 [cit. 2022-12-02]. Dostupné z: doi:ISBN 0-471-29546-9
- [26] Yi-Wen Chen (2022). Voice Activity Detection by Spectral Energy (<https://github.com/JarvusChen/MATLAB-Voice-Activity-Detection-by-Spectral-Energy>), GitHub. Retrieved December 2, 2022.

SEZNAM PŘÍLOH

- Příloha 1 – Obsah disku

PŘÍLOHA 1 – OBSAH DISKU

Na disku jsou soubory přílohy uloženy ve složce DIPLOMOVÁ PRÁCE:

- Soubory s anotací nádechů ve formátu .lab ve složce **./Labels**
- Skripty programu matlab ve složce **./Matlab**
- Výsledné tabulky ve formátu .mat, tabulka s klinickými daty a tabulky výsledků statistické analýzy a strojového učení ve složce **./Tabulky**
- Python skripty ve složce **./Python**
- Soubory s detekovanými nádechy ve formátu .lab ve složce **./Labels_predict**