



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING

HODNOCENÍ ZDRAVÍ POMOCÍ CHYTRÝCH ZAŘÍZENÍ

HEALTH ASSESSMENT USING SMART DEVICES

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. Enikő Vargová

VEDOUCÍ PRÁCE

SUPERVISOR

Ing. Andrea Němcová, Ph.D.

BRNO 2023

Diplomová práce

magisterský navazující studijní program **Biomedicínské inženýrství a bioinformatika**

Ústav biomedicínského inženýrství

Studentka: Bc. Enikő Vargová

ID: 211218

Ročník: 2

Akademický rok: 2022/23

NÁZEV TÉMATU:

Hodnocení zdraví pomocí chytrých zařízení

POKyny PRO VYPRACOVÁNÍ:

1) Proveďte literární rešerši v oblasti hodnocení zdraví pomocí chytrých zařízení. Vyberte si alespoň jeden konkrétní parametr (např. glykémii) a na ten se zaměřte. 2) Pomocí vybraného/ých chytrého/ých zařízení nasnímejte data potřebná pro stanovení vybraného parametru zdraví. Nasnímejte také referenční data. 3) Předzpracujte nasnímaná data, případně extrahujte příznaky pro stanovení vybraného parametru zdraví. 4) Vytvořte algoritmus pro stanovení vybraného parametru zdraví. 5) Porovnejte výsledky vašeho algoritmu s referencí a vyhodnotte úspěšnost vašeho algoritmu. 6) Diskutujte dosažené výsledky a porovnejte je s výsledky jiných autorů.

DOPORUČENÁ LITERATURA:

[1] JAT, Avnish Singh a Tor-Morten GRØNLI. Smart Watch for Smart Health Monitoring: A Literature Review. In: Bioinformatics and Biomedical Engineering [online]. Cham: Springer International Publishing, 2022, 2022-06-08, s. 256-268. Lecture Notes in Computer Science. ISBN 978-3-031-07703-6. Dostupné z: doi:10.1007/978-3-031-07704-3_21.

[2] ISLAM, Tanvir Tazul, Md Sajid AHMED, Md HASSANUZZAMAN, Syed Athar BIN AMIR a Tanzilur RAHMAN. Blood Glucose Level Regression for Smartphone PPG Signals Using Machine Learning. Applied Sciences. 2021, 11(2), 1-20. ISSN 2076-3417. Dostupné z: doi:10.3390/app11020618.

Termín zadání: 6.2.2023

Termín odevzdání: 22.5.2023

Vedoucí práce: Ing. Andrea Němcová, Ph.D.

prof. Ing. Valentine Provazník, 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

Tato práce se zabývá možnostmi neinvazivního stanovení glykémie z fotopletysmografických signálů. Zvýšená glykémie často souvisí s onemocněním zvaným diabetes mellitus, což je jedno z nejčastějších chronických onemocnění na světě. Neléčený diabetes bývá často příčinou úmrtí. Cílem práce je vytvořit algoritmus pro klasifikaci a predikci glykémie.

Nejdříve byly naměřeny dvě databáze PPG signálů s využitím dvou chytrých zařízení (chytrý náramek a chytrý telefon). Současně byla měřena také referenční hodnota glykémie invazivně. PPG signály byly předzpracovány a byly z nich extrahovány vhodné příznaky pro klasifikaci i predikci konkrétní hodnoty glykémie. Následně bylo vytvořeno několik různých modelů pro klasifikaci glykémie do dvou skupin (nízká a vysoká glykémie) a pro predikci konkrétní hodnoty glykémie.

Klíčová slova

PPG, glykémie, diabetes, chytrý telefon, chytrá zařízení, klasifikace, predikce

Abstract

This thesis deals with the possibilities of non-invasive determination of blood glucose from photoplethysmographic signals. Elevated blood sugar is often associated with disease called diabetes mellitus. Diabetes is one of the world's major chronic diseases. Untreated diabetes is often a cause of death. The aim of the work is to propose methods for glycemetic classification and prediction.

Two datasets have been created by recording the PPG signals using two smart devices (a smart wristband and a smartphone), along with their blood glucose levels measured in an invasive way. The PPG signals were preprocessed, and suitable features were extracted from them. Various machine-learning models for glycemetic classification and prediction were created.

Keywords

PPG, glycemia, diabetes, smartphone, smart devices, classification, prediction

Bibliografická citace

VARGOVÁ, Enikö. *Hodnocení zdraví pomocí chytrých zařízení*. Brno, 2023. Dostupné také z: <https://www.vut.cz/studenti/zav-prace/detail/150871>. Diplomová práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Andrea Němcová.

Prohlášení autora o původnosti díla

Jméno a příjmení studenta: *Enikö Vargová*

VUT ID studenta: *211 218*

Typ práce: *Diplomová práce*

Akademický rok: *2022/23*

Téma závěrečné práce: *Hodnocení zdraví pomocí chytrých zařízení*

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: 8. 5. 2023

podpis autora

Poděkování

Poděkování patří vedoucí diplomové práce Ing. Andree Němcové, PhD. za účinnou metodickou, pedagogickou a odbornou pomoc, další cenné rady při zpracování této práce a za propůjčení zařízení potřebných k nasnímání dat. Dále bych chtěla poděkovat MUDr. Zuzaně Novákové, Ph.D. za konzultaci.

V Brně dne: 8. května 2023

podpis autora

Obsah

SEZNAM OBRÁZKŮ	9
SEZNAM TABULEK.....	11
ÚVOD	12
1. HODNOCENÍ ZDRAVÍ POMOCÍ CHYTRÝCH ZAŘÍZENÍ.....	14
2. SLEDOVANÉ PARAMETRY	17
2.1 TEPOVÁ FREKVENCE	17
<i>Možnosti stanovení TF</i>	<i>17</i>
<i>Možnosti stanovení TF s využitím chytrých zařízení</i>	<i>17</i>
2.2 DECHOVÁ FREKVENCE	19
<i>Možnosti stanovení DF.....</i>	<i>19</i>
<i>Možnosti stanovení DF s využitím chytrých zařízení.....</i>	<i>20</i>
2.3 KREVŇÍ TLAK	20
<i>Možnosti stanovení KT.....</i>	<i>20</i>
<i>Možnosti stanovení KT s využitím chytrých zařízení</i>	<i>21</i>
2.4 SATURACE KRVE KYSLÍKEM	22
<i>Možnosti stanovení SpO₂.....</i>	<i>22</i>
<i>Možnosti stanovení SpO₂ s využitím chytrých zařízení.....</i>	<i>23</i>
2.5 ELEKTROKARDIOGRAFIE	24
<i>Snímání EKG.....</i>	<i>25</i>
<i>Snímání EKG s využitím chytrých zařízení.....</i>	<i>25</i>
3. GLYKÉMIE A DIABETES MELLITUS	26
3.1 HORMONY REGULUJÍCÍ GLYKÉMII	26
3.2 DIABETES MELLITUS.....	28
3.3 KOMPLIKACE DOPROVÁZEJÍCÍ DIABETES	29
3.4 INVAZIVNÍ MĚŘENÍ GLYKÉMIE.....	30
3.4.1 <i>Self-monitoring</i>	<i>30</i>
3.4.2 <i>Orální glukózový toleranční test (oGTT)</i>	<i>31</i>
3.4.3 <i>Glykovaný hemoglobin (HbA1c).....</i>	<i>31</i>
3.4.4 <i>Kontinuální monitorování glykémie</i>	<i>31</i>
3.5 NEINVAZIVNÍ MĚŘENÍ GLYKÉMIE	32
3.6 DETEKCE DIABETU Z EKG SIGNÁLU	33
4. FOTOPLETYSMOGRAFIE.....	34
4.1 FOTOPLETYSMOGRAFICKÁ KŘIVKA	35
5. METODY STANOVENÍ GLYKÉMIE Z PPG	36
5.1 KLASIFIKACE GLYKÉMIE PRO PREDIKCI DIABETU Z PPG SIGNÁLU	36
5.2 MOŽNOSTI STANOVENÍ KONKRÉTNÍ HODNOTY GLYKÉMIE Z PPG SIGNÁLU	40
6. SNÍMÁNÍ DAT.....	43
6.1 POSTUP MĚŘENÍ.....	43
6.2 SNÍMÁNÍ PPG POMOCÍ E4 WRISTBAND.....	44

6.3	SNÍMÁNÍ PPG POMOCÍ CHYTRÉHO TELEFONU	45
6.4	MĚŘENÍ GLYKÉMIE.....	46
7.	PŘÍPRAVA DATASETU A PŘEDZPRACOVÁNÍ SIGNÁLŮ.....	48
7.1	PŘÍPRAVA DATASETU	48
7.2	STANOVENÍ KVALITY PPG SIGNÁLŮ NASNÍMANÝCH CHYTRÝM TELEFONEM.....	50
7.3	PŘEDZPRACOVÁNÍ PPG SIGNÁLŮ	52
8.	PŘÍZNAKY	53
8.1	ODLEHLÉ HODNOTY	53
8.2	STATISTICKÁ ANALÝZA PŘÍZNAKŮ	55
8.3	VÝBĚR PŘÍZNAKŮ.....	55
8.4	POPIS VYBRANÝCH PŘÍZNAKŮ	56
9.	MODELY PRO KLASIFIKACI A PREDIKCI	66
9.1	ROZDĚLENÍ A STANDARDIZACE DAT.....	66
9.2	KLASIFIKAČNÍ MODELY	67
9.2.1	<i>K-nejbližších sousedů.....</i>	<i>67</i>
9.2.2	<i>Logistická regrese</i>	<i>69</i>
9.2.3	<i>Metoda podpůrných vektorů</i>	<i>70</i>
9.2.4	<i>Náhodný les.....</i>	<i>71</i>
9.3	REGRESNÍ MODELY	75
9.3.1	<i>K-nejbližších sousedů.....</i>	<i>75</i>
9.3.2	<i>Metoda podpůrných vektorů</i>	<i>76</i>
9.3.3	<i>Náhodný les.....</i>	<i>77</i>
10.	ÚSPĚŠNOST JEDNOTLIVÝCH MODELŮ.....	80
10.1	VÝSLEDKY KLASIFIKACE.....	80
10.2	VÝSLEDKY PREDIKCE	84
10.3	DISKUSE	86
	ZÁVĚR.....	91
	SEZNAM SYMBOLŮ A ZKRATEK	102

SEZNAM OBRÁZKŮ

Obrázek 1.1: Klasifikace chytrých zařízení na základě umístění na těle, převzato a upraveno z [4].....	14
Obrázek 2.1: PPG signál získaný zpracováním zelené složky videa a SWT frekvenční pásma pro nízkou TF, fyziologickou TF a vysokou TF. Frekvenční pásma používána pro odhad TF zvýrazněna zeleně. [23].....	18
Obrázek 2.2: Stanovení úhlu vzestupné hrany křivky PPG ($m\lambda$) a výšky PPG křivky ($vp\lambda$) [37].....	23
Obrázek 2.3: EKG křivka, převzato a upraveno z [42].....	24
Obrázek 3.1: Mechanismus udržování požadované glykémie, upraveno z [46].....	27
Obrázek 3.2: Komplikace diabetu, převzato a upraveno z [46].....	30
Obrázek 3.3: Senzor FreeStylelibre (vlevo) a potřebná čtečka (vpravo) [46].....	32
Obrázek 4.1: Transmisní (a) a reflexní (b) snímání, upraveno z [57].....	34
Obrázek 4.2: PPG křivka, upraveno z [58].....	35
Obrázek 5.1: Klasifikátory, které byly testovány, upraveno z [59].....	37
Obrázek 5.2: Schéma navrhovaného modelu využívajícího EBTA, převzato a upraveno z [59].....	38
Obrázek 5.3: Architektura CNN, převzato a upraveno z [59].....	39
Obrázek 5.4: Architektura CNN, převzato a upraveno z [59].....	40
Obrázek 5.6: Navrhovaný model pro stanovení glykémie z PPG signálu, převzato a upraveno z [6].....	41
Obrázek 6.1: Schéma měření.....	43
Obrázek 6.2: Specifikace zařízení Empatica E4 wristband, převzato a upraveno z [62].....	44
Obrázek 6.3: Ukázka 10 s úseku PPG signálu nasnímaného Empaticou.....	45
Obrázek 6.4: Ukázka 10 s úseku PPG signálu nasnímaného chytrým telefonem.....	46
Obrázek 6.5: Extrapolace změřených hodnot glykémie (červené kroužky).....	47
Obrázek 7.1: Schéma přípravy datasetu (chytrý náramek).....	49
Obrázek 7.2: Ukázka nekvalitního a kvalitního PPG signálu nasnímaného chytrým telefonem.....	50
Obrázek 7.3 : Ukázka GUI pro anotaci PPG signálů (invertované pro lepší vizualizaci).....	51
Obrázek 7.4: Ukázka PPG signálu č.8 před a po předzpracování.....	52
Obrázek 8.1: Skóre důležitosti příznaků.....	56
Obrázek 8.2: Krabicový graf šikmost.....	57
Obrázek 8.3 : Krabicový graf špičatost.....	58
Obrázek 8.4: Krabicový graf entropie.....	59
Obrázek 8.5: Krabicový graf SDNN.....	60
Obrázek 8.6 : Krabicový graf CV.....	61
Obrázek 8.7: Krabicový graf pNN50.....	62
Obrázek 8.8: Krabicový graf minimální vzdálenosti PP intervalů.....	62
Obrázek 8.9: Krabicový graf poměru minima a maxima 1. derivace PPG signálu.....	63
Obrázek 8.10 : Krabicový graf p2rms.....	64
Obrázek 8.11: Krabicový graf poměru minima a maxima TKEO obálky.....	64
Obrázek 8.12: Krabicový graf poměru minima a maxima 2. derivace PPG signálu.....	65
Obrázek 9.1: Optimální počet sousedů (10 příznaků).....	68
Obrázek 9.2: Optimální počet sousedů (23 příznaků).....	69
Obrázek 9.3: Data rozdělená pomocí různých jader SVM.....	70
Obrázek 9.4: Schéma klasifikačního stromu.....	71
Obrázek 9.5: Nalezení optimálního parametru <i>min sample leaf</i> (10 příznaků).....	73
Obrázek 9.6: Nalezení optimálního počtu stromů (10 příznaků).....	73
Obrázek 9.7: Nalezení optimálního parametru <i>min sample leaf</i> (23 příznaků).....	74
Obrázek 9.8: Nalezení optimálního počtu stromů (23 příznaků).....	74
Obrázek 9.9: Nalezení optimálního počtu sousedů pro regresní úlohu (23 příznaků).....	75

Obrázek 9.10: Nalezení optimálního počtu sousedů pro regresní úlohu (10 příznaků)	76
Obrázek 9.11: Nalezení optimálního počtu parametru min samples leaf pro regresí (23 příznaků)	77
Obrázek 9.12: Nalezení optimálního počtu stromů pro regresí (23 příznaků)	78
Obrázek 9.13: Nalezení optimálního počtu parametru min samples leaf pro regresí (10 příznaků)	78
Obrázek 9.14: Nalezení optimálního počtu stromů pro regresní úlohu (10 příznaků)	79
Obrázek 10.1: ROC křivky a hodnota AUC pro jednotlivé modely na testovací sadě dat z Empaticy.....	83
Obrázek 10.2: ROC křivky a hodnota AUC pro jednotlivé modely na testovací sadě dat z chytrého telefonu.....	84
Obrázek 10.3: Matice záměn RF klasifikátoru pro testovací sadu dat z chytrého náramku (vlevo) a testovací sadu dat z chytrého telefonu (vpravo)	87
Obrázek 10.4: Matice záměn SMV (rbf) klasifikátoru pro testovací sadu dat z chytrého náramku (vlevo) a testovací sadu dat z chytrého telefonu (vpravo)	87
Obrázek 10.5: MAE pro jednotlivé iterace křížové validace na testovací sadě dat z chytrého náramku....	89

SEZNAM TABULEK

Tabulka 1.1: Typy chytrých zařízení a jejich základní funkce.....	15
Tabulka 1.2: Výhody a nevýhody chytrých zařízení	16
Tabulka 6.1: Počet měřených zdravých osob a osob s diabetem / prediabetem.....	43
Tabulka 6.2: Celkový počet naměřených záznamů zdravých osob a osob s diabetem / prediabetem.....	44
Tabulka 7.1: Zastoupení hodnot glykémie dle zvolené prahové hodnoty (chytrý náramek)	48
Tabulka 7.2: Zastoupení počtu 10 s záznamů do skupin nízká/vysoká hodnota glykémie	49
Tabulka 7.3: Zastoupení počtu 10 s záznamů do skupin nízká/vysoká hodnota glykémie (chytrý telefon).....	51
Tabulka 8.1: Popis příznaků	54
Tabulka 9.1: Trénovací sada (chytrý náramek).....	66
Tabulka 9.2: Testovací sada (chytrý náramek)	66
Tabulka 10.1: Výsledky klasifikace na trénovací a testovací sadě dat z Empaticy.....	81
Tabulka 10.2: Výsledky klasifikace na testovací sadě dat z chytrého telefonu	82
Tabulka 10.3: Výsledky predikce glykémie na trénovací a testovací sadě dat z Empaticy	85
Tabulka 10.4: Výsledky predikce glykémie nejvýkonnějších modelů na testovací sadě dat z chytrého telefonu.....	85
Tabulka 10.5: Výsledky klasifikace glykémie nejvýkonnějších modelů na obou testovacích sadách.....	86
Tabulka 10.6: Porovnání výsledků s jinými autory (klasifikace).....	88
Tabulka 10.7: Průměrná MAE nejlepšího modelu pro obě testovací sady dat	88
Tabulka 10.8: Porovnání výsledků s jinými autory (predikce konkrétní hodnoty glykémie)	90

ÚVOD

V posledních letech se čím dál častěji setkáváme s konceptem chytrého zdravotnictví. Chytrá zařízení a nositelnou elektroniku lze využít ke každodennímu monitorování vitálních funkcí jako je například tepová frekvence, dechová frekvence, saturace krve kyslíkem, tělesná teplota aj. Pod pojmem chytrá zařízení si lze představit například chytrý telefon, chytré hodinky, chytré brýle, nebo třeba chytré textilie. Výhodou těchto zařízení je jejich postupná miniaturizace a přenosnost. Chytrá zařízení jsou schopná nám poskytnout informace o našem zdraví v podstatě kdekoliv a kdykoliv. Mimo jiné se můžeme setkat také s chytrými zařízeními, která například umožňují pořídit elektrokardiografický záznam a detekovat v něm případné arytmie, nebo sledovat a hodnotit jednotlivé spánkové fáze. V posledních letech je snahou predikovat, nebo alespoň klasifikovat hladinu glukózy v krvi právě pomocí chytrých zařízení. [1][2][3][4]

Glykémie je pojem, který označuje koncentraci glukózy v plazmě. Zvýšená glykémie se nazývá hyperglykémie a často souvisí s onemocněním zvaným diabetes mellitus (DM). DM patří do heterogenní skupiny chronických onemocnění. Příčinou tohoto onemocnění je relativní nebo absolutní nedostatek inzulínu. IDF (International Diabetes Federation) uvádí, že cukrovkou trpí přibližně 537 milionů lidí na světě ve věku 20-79 let. Neléčený DM bývá často příčinou úmrtí. [5][6][7][8]

V současné době se běžně pro stanovení glykémie používají převážně invazivní metody, které jsou pro pacienty značně nepohodlné. Chytrá zařízení, která by umožňovala stanovit glykémii neinvazivně zatím nejsou běžná. Právě z tohoto důvodu se v této práci věnuji možnosti klasifikace a predikce glykémie z fotopletysmografických (PPG) signálů nasnímaných pomocí chytrých zařízení.

Prvních pět kapitol práce se týká literární rešerše v oblasti dané problematiky. Následující kapitoly se již věnují praktické části. Nejdříve bylo třeba nasnímat databázi PPG signálů. Byly vytvořeny dvě databáze PPG signálů. Jedna databáze obsahuje data nasnímaná pomocí chytrého náramku. Jelikož pro klasifikaci i predikci glykémie budou použity metody strojového učení, je první databáze využívána pro trénování a testování modelu. Druhá databáze pak obsahuje data nasnímaná z chytrého telefonu a je použita k ověření funkčnosti navržených algoritmů.

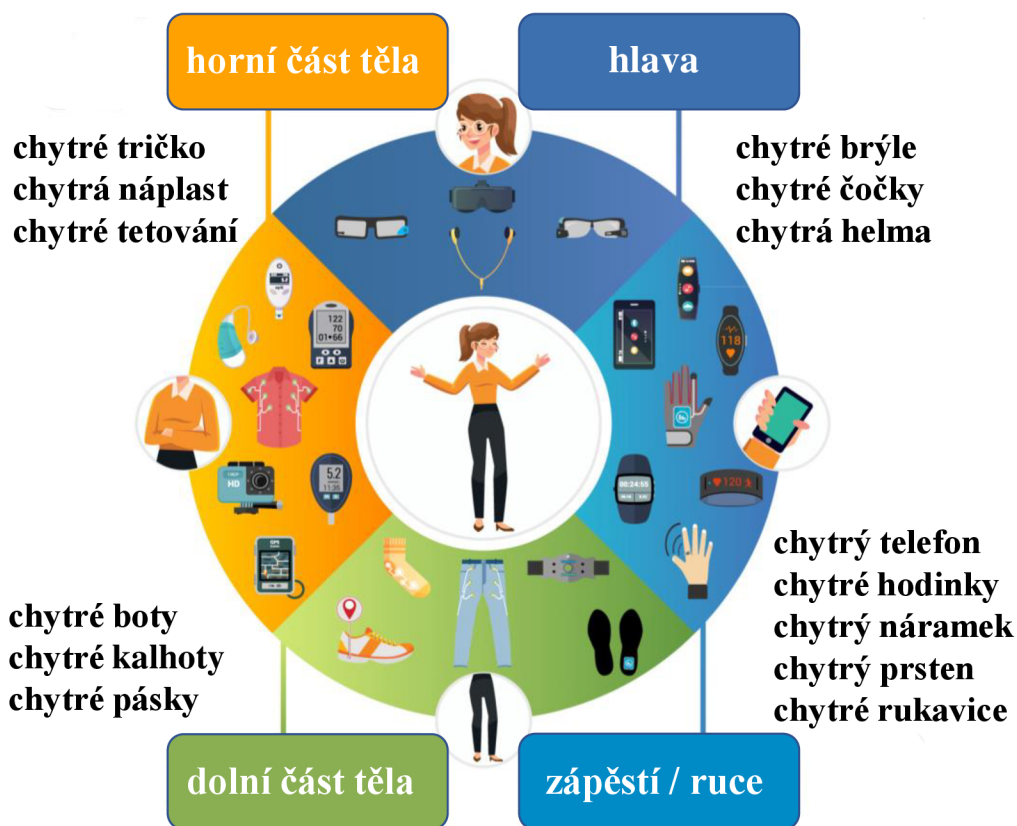
Následuje příprava datasetu a předzpracování PPG signálů. Z dat nasnímaných pomocí chytrého náramku bylo vypočítáno několik příznaků. Nejdříve byl proveden výběr příznaků s využitím statistických testů. Dodatečně byla použita také selekční metoda, kterou bylo vybráno omezené množství příznaků, které se jeví jako vhodné pro klasifikaci/predikci glykémie z PPG signálů. Po normalizaci a odstranění odlehklých hodnot jsou tato data použita pro učení a poté také testování.

Další části práce se věnují vytvoření několika klasifikačních a regresních modelů. Poslední kapitola diplomové práce se věnuje vyhodnocení výsledků klasifikačních i regresních modelů, výběru nejlepších modelů pro obě zmíněné úlohy a ověření jejich funkčnosti na datech nasnímaných chytrým telefonem. Kapitola také obsahuje diskusi a srovnání dosažených výsledků s jinými autory.

1. HODNOCENÍ ZDRAVÍ POMOCÍ CHYTRÝCH ZAŘÍZENÍ

S rozvojem informačních technologií se v posledních letech dostává do popředí koncept chytrého zdravotnictví (smart healthcare). Možnosti bezdrátové komunikace, pokrok v umělé inteligenci, vývoj pokročilých zobrazovacích technologií a trend miniaturizace společně vedly ke vzniku špičkových chytrých zařízení, která lze využít ke každodennímu sledování vitálních funkcí jedince. Pod pojmem chytrá zařízení ve zdravotnictví si můžeme představit chytrý telefon nebo nositelnou elektroniku, anglicky Wearable Health Devices (WHD). [1][2][3][4]

Chytrá zařízení napomáhají získávat data o zdravotním stavu jedince i mimo klinická prostředí a zlepšují tak prevenci ve zdravotnictví. Umožňují dlouhodobé sledování zdraví, což napomáhá lékařům stanovit diagnózu, sledovat stav pacienta po lékařském zákroku, zranění či během rekonvalescence. V dnešní době existuje velké množství různých typů chytrých zařízení, které se mohou lišit svým vzhledem i funkcí. WHD můžeme v zásadě klasifikovat na základě umístění zařízení na těle během používání (Obrázek 1.1). [1][2][3][4]



Obrázek 1.1: Klasifikace chytrých zařízení na základě umístění na těle, převzato a upraveno z [4]

K nositelným zařízením, která jsou umístěná na hlavě patří například chytrá helma, kontaktní čočky nebo chytré brýle. Dále rozlišujeme WHD umístěná na těle, tzn. chytré náramky, chytré hodinky, chytré náplasti, řetízky a chytrá tetování (invazivní). Tato zařízení slouží především ke sledování aktivity a vitálních funkcí. Do této skupiny lze také zařadit chytré textilie, chytré rukavice nebo chytré boty. Zařízení, která jsou umístěná kdekoli na těle lze dále klasifikovat na zařízení umístěná na horní nebo dolní části těla. Jednotlivé typy zařízení a funkce, které vykonávají shrnuje Tabulka 1.1. K nejčastěji monitorovaným parametrům patří tepová frekvence (TF), dechová frekvence (DF), krevní tlak (KT) a saturace krve kyslíkem (SpO₂). Více v kapitole 2. [4]

Tabulka 1.1: Typy chytrých zařízení a jejich základní funkce

Typ zařízení	Aplikace v oblasti zdraví a fitness
chytrá helma	detekce ospalosti, detekce mozkových vln, monitorování pohybu hlavy, sledování tepové frekvence, snímání EEG a EKG [4][9][10]
chytré kontaktní čočky	monitorování hladiny glukózy v krvi, tělních tekutin, měření nitroočního tlaku [4]
chytré brýle	vizualizace žil a tělních struktur, mentoring, přístup k medicínským datům a obrazům [4][11]
chytrý řetízek	sledování aktivity a vitálních funkcí, korekce držení těla [4]
chytrá náplast	sledování aktivity a vitálních funkcí, sledování tělních tekutin a hydratace [4]
chytré rukavice	rozpoznávání gest, rehabilitace [4]
chytré náramky	sledování aktivity a vitálních funkcí, detekce stresu, měření EKG, PPG a EDA [4][12]
chytré hodinky	sledování aktivity a vitálních funkcí, detekce třesu a záchvatů, detekce stresu, detekce pádu, sledování stravy a užívání léků, monitorování EKG, PPG a EDA, detekce arytmií, sledování spánku, glykémie aj. [4]
chytré tetování	měření glykémie, sledování tělních tekutin a hydratace, vystavení UV světlu [13]
chytrá textilie	sledování aktivity a vitálních funkcí, snímání EKG, EMG, EDA, měření tělesného tuku a spálených kalorií, sledování hydratace [4][14]
chytré boty	sledování aktivity, držení těla, trénink profesionálních sportovců [4]

Dále lze WHD klasifikovat dle těchto tří hledisek – použití (domácí/ klinické prostředí), typ uživatele (zdravý/ pacient) a typ monitorování (online/ offline). Monitorování lze dále rozdělit na dvě oblasti – monitorování během aktivity (fitness, wellness, rehabilitace, nelékařské aplikace a sebe-monitorace) a monitorování pro účely medicíny, které lze rozdělit do tří skupin - predikce, detekce a diagnostická podpora. Účelem predikce je předpovídat události, které zatím nenastaly. Využívá se především v oblasti prevence a může napomoci při stanovení diagnózy. Z nasbíraných dat o pacientovi lze predikovat například hladinu glukózy v krvi nebo vytvořit systém prediktivního rozhodování pro dialyzované pacienty. Detekce má za úkol identifikovat anomálie, které neodpovídají normě. Jedná se například o detekci různých arytmií v elektrokardiografických (EKG) záznamech nebo detekci anomálií ve spánku. Třetí oblastí je diagnostická podpora, jejímž výsledkem je klinické rozhodnutí na základě získaných dat a záznamů. Tato oblast je úzce spjata s detekcí anomálií, jelikož anomálie mají velký význam v procesu rozhodování. Rozdíl mezi detekcí anomálií a diagnostickou podporou je v tom, že diagnostická podpora nevyužívá pouze detekce anomálií, ale při rozhodování uplatňuje globálnější informace než jen samotné abnormální vzorce. Diagnostika nachází využití například v odhadu závažnosti různých chronických onemocnění a zdravotních stavů, rozpoznávání emocí apod. [15]

Chytrá zařízení mohou tedy pomoci zlepšit zdravotní péči i prevenci, a snižovat náklady díky nepřetržitému monitorování zdraví. Mezi hlavní přednosti těchto zařízení patří monitorování zdravotního stavu v reálném čase, snadná přenosnost, relativně dostupná cena a možnost jednoduše ukládat data a sdílet je s lékařem. Chytrá zařízení tak mají potenciál zvyšovat kvalitu života pacientů a poskytují značné ekonomické výhody. Na druhou stranu používání chytrých zařízení má i svá úskalí, například možný únik citlivých dat, potřeba neustálého dobíjení, poměrně nízká životnost baterií a s tím související problémy s jejich likvidací aj. Možné výhody a nevýhody shrnuje Tabulka 1.2. [2][3][16]

Tabulka 1.2: Výhody a nevýhody chytrých zařízení

Výhody	Nevýhody
Nepřetržité monitorování zdravotního stavu	Možný únik citlivých dat
Možnost sledování zdravotního stavu odkudkoliv	Poměrně časté dobíjení
Automatické vyhodnocení a zpětná vazba uživateli	Životnost baterie
Jednoduché používání	Aktualizace softwaru
Možnost ukládání a sdílení dat	Plně nenahrazuje návštěvu lékaře
Cenově dostupná zařízení	
Neinvasivní (většinou)	

2. SLEDOVANÉ PARAMETRY

Chytrá zařízení jako chytré telefony, chytré hodinky, fitness náramky aj. disponují řadou různých senzorů, které napomáhají sledovat životní funkce. K nejčastěji monitorovaným parametrům patří TF, SpO₂, DF, KT a tělesná teplota. WHD lze také využít ve spirometrii, pro monitoraci EKG, ke sledování spánku, při vyšetření poddajnosti cév, detekci pádu, držení těla, snímání pocení nebo pro zaznamenávání informací o hladině glukózy v krvi. [15]

2.1 Tepová frekvence

TF udává počet tepů za minutu. Je základním ukazatelem zdraví a poskytuje okamžitou informaci o fyziologickém stavu monitorované osoby. Za klidových podmínek by TF měla odpovídat 60-90 tepů/minutu. Nízká srdeční frekvence (pod 60 tepů/minutu) je označována jako bradykardie. Fyziologicky se pomalá srdeční frekvence objevuje ve spánku a u některých sportovců. Naopak vysoká TF (nad 90 tepů/minutu) se nazývá tachykardie a je fyziologická při stresu či námaze. Při měření tepu se sleduje jeho pravidelnost, frekvence a kvalita. [5]

Možnosti stanovení TF

TF lze určit nejčastěji pohmatem na a. radialis nebo a. carotis (palpační metoda), poslechem (fonokardiografie) nebo ze snímání pulzové křivky (sfygmografie) či z EKG křivky, což je považováno za zlatý standard. Elektrokardiografie je nejpoužívanější technikou pro vyšetření srdce. Snímá jeho elektrickou aktivitu. Výstupem snímání je EKG signál, z něhož lze určit hodnotu TF z délky trvání RR intervalů. [5][17]

Možnosti stanovení TF s využitím chytrých zařízení

TF lze získat i s využitím chytrých zařízení. Například EKG signál je možné nasnímat i pomocí některých chytrých hodinek (FitBit, Apple Watch, Samsung Galaxy Watch) a následně z něj lze opět určit hodnotu TF z délky trvání RR intervalů. Více o snímání EKG pomocí chytrých zařízení v kapitole 2.5.

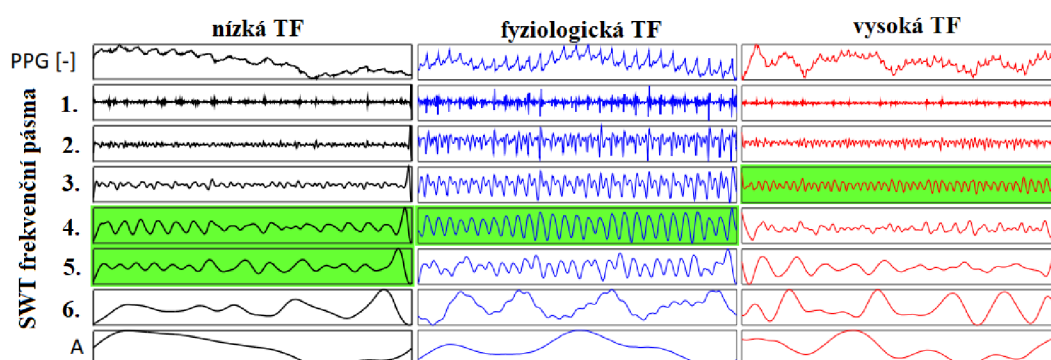
Další z možností stanovení TF s využitím chytrého zařízení využívá akcelerometru, senzoru, který umožňuje sledovat pohyb či detekovat pád. Srdeční aktivita způsobuje vibrace hrudního koše, které jsou způsobené přesunem krve z aorty do cév v periférii během každého srdečního cyklu. Při měření je mobilní zařízení upevněno na hrudi měřené osoby a akcelerometr je tak schopen zachytit vibrace hrudního koše. Tato metoda se nazývá seismokardiografie. Získaný záznam má tři osy (x, y a z). Na jedné z os v závislosti na jejich orientaci (většinou na ose z) lze identifikovat systolický komplex odpovídající R vlnám v EKG signálu, díky nimž je možné určit TF. [18][19]

Fonokardiografie (FKG) je metoda, při níž jsou zaznamenávány srdeční ozvy. Při této metodě se využívá mikrofonu chytrého telefonu, který je přiložen k hrudi. V záznamu jsou slyšitelné dvě srdeční ozvy S1 a S2. Po filtraci získaného signálu a vytvoření jeho obálky, je možné detekovat peaky. Jelikož S1 ozvy odpovídají R vlnám v EKG záznamu, lze touto metodou stanovit TF. [20]

Většina chytrých zařízení (chytré telefony, chytré hodinky atd.) využívají pro stanovení TF PPG signál. Fotopletysmografie je optická metoda, která detekuje změny objemu krve. PPG signál reflektuje pohyb krve v cévách. Nejčastěji se snímá na prstu ruky (chytrý telefon, pulzní oxymetr) nebo na zápěstí (chytré hodinky). Více o PPG v kapitole 4. [21][22]

Pro stanovení TF z PPG pomocí chytrého telefonu se využívá fotoaparátu. Měřená osoba přiloží svůj prst na čočku fotoaparátu a snaží se prstem překrýt čočku i zdroj světla (LED) emitující světlo v průběhu měření. Tímto způsobem je nasnímán videozáznam přiloženého prstu. Nasnímaný videozáznam je poté zpracován v programovém prostředí. Ve většině případů je RGB video rozděleno na tři základní barevné složky – červená, zelená a modrá. PPG lze získat i z videozáznamu obličeje.

Jedním z možných přístupů zisku TF z PPG je použití vlnkové transformace, jak popisují autoři článku [23]. Nasnímané video bylo rozděleno na jednotlivé snímky. Každý snímek byl rozložen na tři barevné složky (červená, zelená a modrá). Pro každý snímek a každou RGB složku byla spočtena průměrná hodnota snímku, což odpovídá jednotlivým PPG signálům. Následně byla pro odhad TF použita stacionární vlnková transformace (SWT). Autoři po rozkladu vlnkovou transformací zvolili pro fyziologické pásmo TF (55 – 111 tepů/minutu) 4. frekvenční pásmo. Pro TF nižší či vyšší je vhodné volit pásma jiná (Obrázek 2.1). Ve vhodně zvoleném pásmu byly detekovány peaky a z nich stanovená TF. [23]



Obrázek 2.1: PPG signál získaný zpracováním zelené složky videa a SWT frekvenční pásma pro nízkou TF, fyziologickou TF a vysokou TF. Frekvenční pásma používána pro odhad TF zvýrazněna zeleně. [23]

Například autoři článku [24] po extrakci tří barevných složek zvolili odlišný postup. Pro každý snímek a kanál stanovili prahovou hodnotu dle vzorce (1.1). Poté byla pro každý snímek spočtena suma pixelů, jejichž intenzita jasu převyšuje práh. [24]

$$Práh = 0,99 \times (I_{max} - I_{min}), \quad (1.1)$$

kde $(I_{max} - I_{min})$ značí rozdíl maximální a minimální intenzity.

Autoři také uvádějí, že červená složka videa je pro stanovení TF nejvhodnější. Dále jsou detekována minima, protože počet snímků na jeden cyklus srdeční je dán rozdílem vzdálenosti po sobě jdoucích minim. Dosazením do vzorce (1.2) se spočítá TF. [24]

$$TF = \frac{\text{snímková frekvence}}{\text{počet snímků na 1 srdeční cyklus}} \times 60 \quad (1.2)$$

2.2 Dechová frekvence

Dýchání je jednou ze základních životních funkcí, která je do jisté míry ovlivnitelná vůlí. Dechová frekvence (DF), tedy počet dechů za minutu, je důležitým fyziologickým parametrem, který indikuje aktuální stav pacienta. Fyziologická hodnota DF u zdravého a dospělého člověka je v rozmezí 12-16 dechů/minutu. Změna DF může být prediktorem srdeční zástavy a různých poruch dýchání, nebo může v jistých případech poukazovat na plicní embolii, zápal plic, sepsi či jiná onemocnění. Změny DF lze také pozorovat při úzkostech nebo užití narkotik. [25][26]

Možnosti stanovení DF

Zjistit DF je možné pohledem či poslechem, pomocí indukční pletysmografie, impedanční pneumografie nebo jako odhad z EKG či PPG signálu. Indukční pletysmografie je založená na měření změn indukčnosti, které jsou způsobené dýcháním. Při měření se používá hrudní a břišní pás s integrovanými cívkami. Součástí je také řídicí modul a sada vodičů. Jelikož cívkou v pásu prochází střídavý proud, generuje se magnetické pole. Vlivem dýchání se mění průřez těla, což má za následek i změnu magnetického pole a také změny vlastní indukčnosti cívek a frekvence. Impedanční pneumografie je metoda založená na měření změny impedance tkáně v hrudním koši během dýchání. Tato metoda využívá vysokofrekvenčního zdroje proudu pro napájecí elektrody. Na zesilovač a demodulátor je přivedena dvojice měřících elektrod. Výstupem demodulátoru je poté nízkofrekvenční složka, která odpovídá dýchání. [25][26]

Možnosti stanovení DF s využitím chytrých zařízení

Pomocí chytrých zařízení lze nasnímat EKG i PPG signál. Odhad dechové křivky lze provést ze zmíněných signálů metodou extrakce driftu. Drift, tedy kolísání nulové izoliny, je rušení, které postihuje EKG i PPG signály a je zapříčiněno především dýcháním. Filtrací signálu FIR filtrem typu dolní propust s vhodnou mezní frekvencí (například 0,44 Hz, což je přibližně 26 dechů/minutu) lze ze signálů získat dechovou křivku. Z dechové křivky je pak možné různými metodami detekce peaků zjistit dechovou frekvenci. [27][28]

DF lze získat i za pomoci vlnkové transformace [29], empirického rozkladu [30], homomorfní filtrace [31] aj. V některých případech je vhodné použít robustnější metodu, která je založena na fúzi více odhadů DF získaných různými metodami. [32]

2.3 Krevní tlak

Krevní tlak je nejdůležitějším kardiopulmonálním parametrem a souvisí s vypuzováním krve ze srdce v průběhu systoly. Fyziologicky dosahuje hodnot 120/80 mmHg. Během systoly krev přitéká do arterií a tlak dosahuje svých maximálních hodnot – systolický tlak. Ten je ovlivněn silou srdečního stahu, poddajností arterií a objemem vypuzené krve. Snížená poddajnost cév a zvýšený objem vypuzené krve systolický tlak zvyšují. Po skončení systoly následuje na tlakové křivce dikrotický zářez. V průběhu diastoly tlak klesá, protože už nepřitéká žádná krev do arterií. Tlak v tuto chvíli dosahuje svých nejnižších hodnot – diastolický tlak, který závisí především na srdeční frekvenci a na periferním odporu. Pokles TF diastolický tlak snižuje, naopak zvýšení periferního odporu jej zvyšuje. [3][33]

Možnosti stanovení KT

Metody měření KT lze rozdělit na přímé (invazivní) a nepřímé (neinvazivní). Neinvazivní měření krevního tlaku je častější a pro pacienta značně komfortnější. Snímač se nachází vně těla pacienta bez přímého spojení s krevním řečištěm. [17]

Neinvazivní měření KT je možné například pomocí auskultační metody. Jedná se o diagnostický standard, který využívá manžetu omezující průtok krve v měřeném místě. Měření je nejčastěji prováděno na paži horní končetiny. Princip metody spočívá v úplné zástavě proudění krve v arterii, což je způsobeno vyšším tlakem v manžetě, než je systolický tlak v arterii. Tlak v manžetě postupně klesá a jakmile klesne pod hodnotu systolického tlaku, vznikají tzv. Korotkovovy ozvy způsobené turbulentním prouděním krve. Korotkovovy ozvy vymizí ve chvíli, kdy tlak v manžetě klesne pod hodnotu diastolického tlaku. Systolický tlak tedy odpovídá hodnotě na tonometru při vzniku Korotkovových fenoménů a diastolický tlak pak hodnotě na tonometru při jejich vymizení. Dalšími možnými metodami pro neinvazivní měření KT jsou například oscilometrická metoda, ultrazvuková metoda nebo Peňázova (fotopletysmografická) metoda, která je na rozdíl od ostatních výše zmíněných metod

spojitá. Manžeta je v tomto případě umístěna na prstu a je vybavena zdrojem světla - infračervenou diodou v oblasti kolem 940 nm, jelikož pohlcení světla při této vlnové délce není závislé na nasycení krve kyslíkem. Tlak v manžetě je řízen tak, aby se intenzita světla procházející prstem neměnila a byl zachován konstantní objem krve. [17]

Oscilometrická metoda je konvenční a neinvazivní metoda, která snímá tlakové změny uvnitř manžety způsobené pulzací brachiální artérie. Manžeta umístěná na paži se nafoukne (natlakuje) na předem definovanou hodnotu a dojde k uzavření artérie. Postupem času se tlak v manžetě snižuje a arterií opět začíná proudit krev, což má za následek vznik oscilací, jejichž amplituda narůstá. Maximální hodnota amplitudy odpovídá střednímu arteriálnímu tlaku. [17]

Možnosti stanovení KT s využitím chytrých zařízení

Neinvazivně lze odhadnout KT také na základě parametru VTT (vascular transit time), což je čas, za který urazí pulzní vlna vzdálenost mezi dvěma místy v arteriálním řečišti. Autoři článku [34] používají ke stanovení VTT dva senzory na mobilním telefonu – kameru pro snímání PPG a mikrofon pro snímání I. a II. srdeční ozvy.

V rámci předzpracování dat je třeba provést synchronizaci PPG a FKG signálů získaných pomocí dvou výše zmíněných senzorů. Po řádném předzpracování je nutné detekovat peaky v PPG signálech, srdeční ozvy v FKG signálech a provést matematickou analýzu.

Dosazením do následujících vzorců lze tedy odhadnout krevní tlak. [34]

$$STK = -0,425 \times VTT + 214, \quad (1.3)$$

kde STK je systolický krevní tlak.

$$BSA = 0,007184 \times m^{0,425} \times v^{0,725}, \quad (1.4)$$

BSA značí plochu povrchu těla, m odpovídá hmotnosti a v výšce.

$$SV = -6,6 + 0,25 \times (ET - 35) - 0,62 \times TF + 40,4 \times BSA - 0,51 \times věk, \quad (1.5)$$

kde ET odpovídá ejekčnímu času, což je v podstatě interval mezi otevřením a uzavřením chlopní a SV je tepový objem stanovený v ml.

$$ET = S2 - S1 \quad (1.6)$$

$$PTK = \frac{SV}{(0,013 \times m - 0,007 \times věk - 0,004 \times TF) + 1,307} \quad (1.7)$$

kde PTK je pulzní krevní tlak.

$$DTK = STK - PTK, \quad (1.8)$$

přičemž DTK je označení pro diastolický krevní tlak.

Měření KT pomocí chytrých hodinek v podstatě funguje na stejném principu. Hodinky využívají senzoru na jejich zadní straně pro získání PPG signálu v kombinaci s EKG signálem. Následně je určen PTT (Pulse Transit Time, doba VTT). Velice zjednodušeně lze tedy konstatovat, že pokud je KT vysoký, tak doba průchodu pulzu ze srdce k zápěstí je rychlejší a v případě nízkého KT naopak pomalejší. Z parametru PTT lze následně dle výrobcem daných postupů odhadnout KT . [35]

Dalším možným přístupem je odhad KT přímo z PPG signálu. Z PPG křivky jsou extrahovány různé příznaky související s jejím tvarem, například šířka vlny, výška vlny, doba mezi dvěma pulzy aj. Tyto příznaky se poté stávají vstupem do různých modelů strojového učení. Na rozdíl od metody popsané výše, je využíván k odhadu KT z PPG signálu pouze jeden senzor. [36]

2.4 Saturace krve kyslíkem

Saturace krve kyslíkem je veličina, která odpovídá procentuálnímu nasycení krve kyslíkem. Hodnota SpO_2 by se za normálních podmínek u zdravého jedince měla pohybovat v rozmezí 95-100 %. V případě nižších hodnot SpO_2 může poukazovat na respirační či srdeční onemocnění. [3][33]

Možnosti stanovení SpO_2

Běžně je kyslíková saturace periferní arteriální krve měřena pomocí pulzního oxymetru přiloženého na konci prstu. Metoda je založena na faktu, že absorpce světla na odlišných vlnových délkách je různá pro krev dostatečně nasycenou kyslíkem (jasně červenou) a málo nasycenou kyslíkem (tmavě červenou). Měření tedy probíhá na dvou vlnových délkách, 660 nm (červené světlo) a 940 nm (infračervené světlo). Červené světlo absorbuje deoxyhemoglobin a infračervené světlo oxyhemoglobin. Kyslíková saturace pak odpovídá procentuálnímu poměru koncentrace oxyhemoglobinu vůči součtu oxyhemoglobinu a deoxyhemoglobinu, přičemž je koncentrace karboxyhemoglobinu a methemoglobinu zanedbána. [17]

Výpočet kyslíkové saturace pomocí pulzního oxymetru je vyjádřen vztahem:

$$SpO_2 = 100 \times \frac{HbO_2}{RHb + HbO_2}, \quad (2.9)$$

kde HbO_2 je koncentrace oxyhemoglobinu a RHb je koncentrace deoxyhemoglobinu.

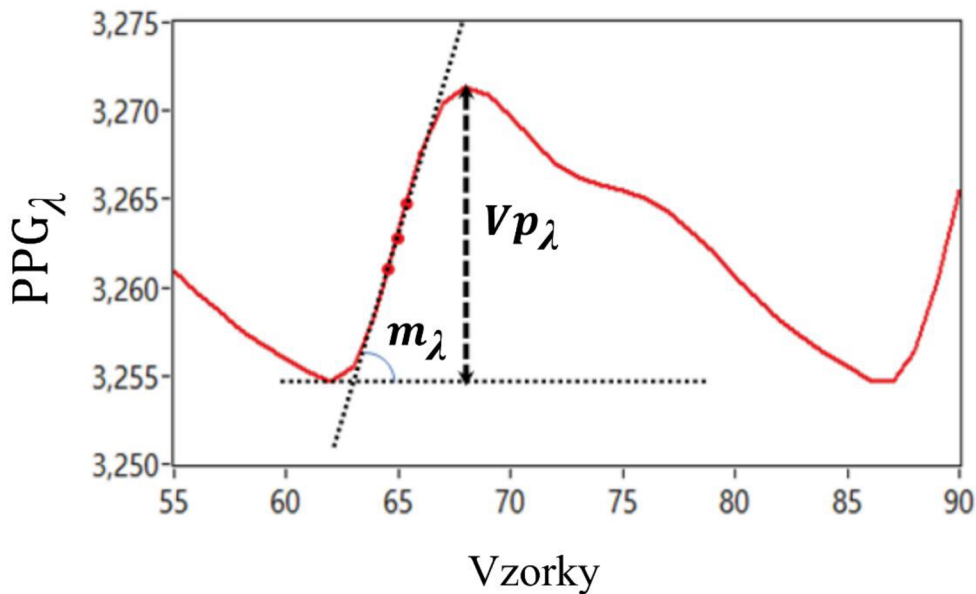
Možnosti stanovení SpO₂ s využitím chytrých zařízení

Možností stanovit hodnotu SpO₂ disponuje většina chytrých hodinek i náramků. Jednu z možných metod, jak stanovit SpO₂ z PPG signálů popisují autoři [37]. Opět se jedná o videozáznam prstu získaný smartphonem s diodou emitující světlo. Video je následně rozloženo na jednotlivé RGB složky, přičemž je využito dvou PPG signálů (zelená složka a červená složka). Po řádném předzpracování a filtraci následuje zjištění parametrů m (úhel vzestupné hrany křivky PPG) a výšky PPG křivky (Obrázek 2.2). Pro dosažení do vzorce (2.10) pro výpočet SpO₂ je třeba také stanovit absorpční koeficienty. [37]

$$SpO_2 = 100 \times \frac{\epsilon_{Hb,600} \sqrt{m_{940} \ln Vp_{940}}}{\sqrt{m_{940} \ln Vp_{940}} (\epsilon_{Hb,600} - \epsilon_{HbO,600})} \dots \quad (2.10)$$

$$\dots \frac{-\epsilon_{Hb,940} \sqrt{m_{600} \ln Vp_{600}}}{\sqrt{m_{600} \ln Vp_{600}} (\epsilon_{Hb,940} - \epsilon_{HbO,940})} ,$$

kde $\epsilon_{Hb,600}$, $\epsilon_{Hb,940}$ jsou absorpční koeficienty deoxyhemoglobinu na vlnových délkách 600 nm a 940 nm, $\epsilon_{HbO,600}$, $\epsilon_{HbO,940}$ jsou absorpční koeficienty oxyhemoglobinu na vlnových délkách 600 nm a 940 nm, m_{940} a m_{600} odpovídají úhlům vzestupných hran PPG křivky (červené a zelené PPG) a Vp značí výšku vrcholů PPG křivky zmíněných vlnových délek.



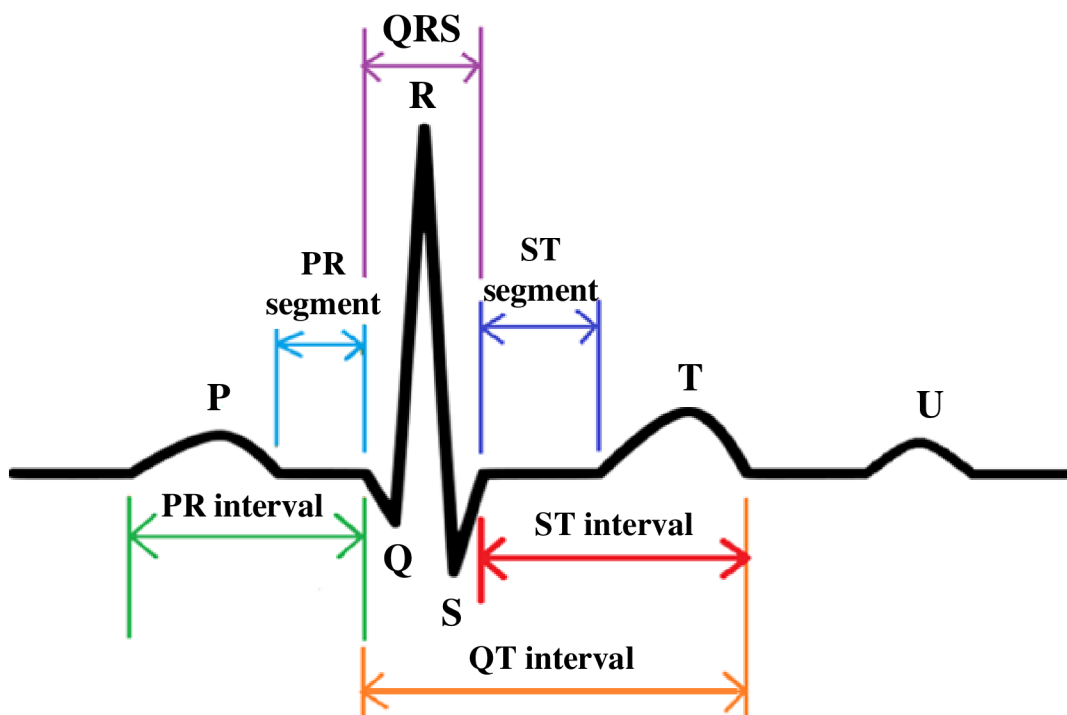
Obrázek 2.2: Stanovení úhlu vzestupné hrany křivky PPG (m_λ) a výšky PPG křivky (v_{p_λ}) [37]

Kromě chytrých hodinek či telefonu lze použít ke stanovení SpO₂ například prstýnek O₂ Ring [38], nebo optická vlákna vsívaná do textilií (chytré textilie). [39]

2.5 Elektrokardiografie

Elektrokardiografie snímá elektrickou aktivitu srdce. Přístroj snímající akční potenciály srdeční se nazývá elektrokardiograf. Elektrokardiogram označuje získaný grafický záznam zobrazující rozdíly těchto potenciálů. V klinické praxi je analýza EKG velice důležitá pro diagnostiku různých kardiovaskulárních onemocnění (fibrilace síní, bradyarytmie, tachyarytmie aj.), srdečního selhání a onemocnění cév. V dnešní době je snaha analyzovat EKG pomocí WHD. [17][33]

Běžně na EKG rozeznáváme 6 základních bodů: P, Q, R, S, T, U (Obrázek 2.3). Vlna P je dána depolarizací síní, za ní následuje QRS komplex odpovídající depolarizaci komor. Vlna T souvisí s repolarizací komor a vlna U je pravděpodobně způsobena repolarizací papilárních svalů. Jako PQ (PR) interval se označuje interval, během kterého je vzruch převeden ze síní na komory. V QRS komplexu je velmi důležitý pozitivní R kmit, jelikož se při analýze EKG používají R-R intervaly například pro měření srdečních cyklů. Jako QT interval se označuje vzdálenost mezi kmitem Q a koncem vlny T. Jelikož se délka QT intervalu mění s TF, často se místo klasického QT hodnotí tzv. normalizovaná délka QT intervalu (QTc). Dále se při analýze EKG sleduje ST segment odpovídající fázi plató. Segment odpovídá úseku mezi QRS komplexem a vlnou T. Měl by být izoelektrický. [40][41]



Obrázek 2.3: EKG křivka, převzata a upraveno z [42]

Snímání EKG

Zlatým standardem pro hodnocení EKG signálů je 12ti svodový systém měřící potenciály na různých místech končetin i hrudníku. Hrudní svody hledí na srdce v rovině horizontální, končetinové svody hledí na srdce v rovině frontální. Celý 12ti svodový systém je tvořen šesti hrudními svody V1-V6. Tyto hrudní svody jsou unipolární. Dále 12ti svodový systém tvoří Eithovenovy bipolární svody I-III, které vychází z tzv. Eithovenova trojúhelníku a unipolární Goldbergovy svody aVR, aVF a aVL. Celkově je během měření využito deseti elektrod – šest hrudních, tři končetinové a nulová elektroda (pravá noha).

Snímání EKG s využitím chytrých zařízení

Chytré hodinky od různých výrobců jsou schopné provést základní analýzu EKG. Výhodou je možnost nasnímání EKG záznamu kdykoliv a odkudkoliv. Dále je výhodou také absence lepících elektrod, na rozdíl od klasického EKG nebo Holterovského měření. Elektrody jsou pacientům často přítěží, jsou nepohodlné a mohou se odlepit (například pocením), což může vést ke vzniku různých chyb a artefaktů v záznamu.

Princip snímání EKG signálu pomocí chytrých hodinek je založen na existenci dvou elektrod. Uživatel má většinou chytré hodinky nasazené na levé ruce, přičemž se jedna elektroda nachází právě na spodní straně hodinek. Prstem druhé ruky se pro úspěšné snímání musí uživatel dotknout druhé elektrody. Druhá elektroda se může nacházet na korunce (Apple Watch), může být součástí jednoho tlačítka (Samsung Watch), nebo obklopovat celý ciferník (Withings). Tento způsob měření však někdy vede ke vzniku nekvalitních záznamů, takže chytré hodinky vyzvou uživatele k opakovanému měření.

Chytré hodinky jsou schopné znázornit elektrokardiogram, vyhodnocovat TF a kontrolovat, zda síně a komory pracují synchronně. Jsou schopné vyhodnotit sinusový rytmus, zvýšenou/sníženou TF a fibrilaci síní. Velkou výhodou je možnost sdílet naměřené záznamy pomocí WHD s doktorem. Je třeba zmínit, že tento způsob měření EKG je pro uživatele spíše informativní a zcela nenahrazuje klasické či holterovské měření. [3][43]

Na závěr této kapitoly by bylo vhodné zmínit, že existují volně dostupné aplikace pro chytré telefony, které se zabývají detekcí fibrilací. Například Preventicus Heartbeats je certifikovaná aplikace, která umožňuje uživateli zaznamenat PPG signál a získat kompletní zprávu o variabilitě rytmu, a tedy detekovat fibrilaci v záznamech. Dále existuje aplikace zvaná FibrriCheck, která je momentálně považována za nejspolehlivější aplikaci v této oblasti. Jde o jedinou aplikaci pro monitorování srdečního rytmu založenou na snímání PPG signálu, která je schválena FDA i CE. [44]

Dále je možné použít WHD k zaznamenávání tělesné teploty, fázi spánku, nebo třeba hladiny glukózy v krvi, čemuž se věnují kapitoly 5.1 a 5.2.

3. GLYKÉMIE A DIABETES MELLITUS

Systémy pro monitoraci hladiny glukózy v krvi (glykémie) jsou nezbytné pro časnou detekci diabetu. Bohužel v současné době se běžně používají pro stanovení glykémie invazivní metody, jejichž nevýhodou je nutnost odběru krve. Chytrá zařízení, která by umožňovala stanovit glykémii neinvazivně zatím nejsou běžná. Z výše zmíněných důvodů se v této práci věnuji možnosti neinvazivního stanovení glykémie z PPG signálu pomocí chytrých zařízení. Následující kapitola tedy pojednává o glykémii a diabetu mellitu.

Glykémie je pojem označující koncentraci glukózy v plazmě. Glukóza je sacharid patřící k hexózám a je nejdůležitějším cukrem v přírodě. Je pohotovým zdrojem energie pro člověka. Představuje hlavní zdroj energie pro mozek i erythrocyty. Fyziologická hodnota glykémie by se měla pohybovat v rozmezí 3,5 – 5,5 mmol/l. Zvýšená hladina glukózy v krvi se nazývá hyperglykémie a je spojována s diabetem. Snížená glykémie je označována jako hypoglykémie a může vést k bezvědomí a náhlé smrti. [33][45]

3.1 Hormony regulující glykémii

Glukagon a inzulín jsou hormony podílející se na regulaci koncentrace glukózy v plazmě. Právě tyto hormony označujeme jako hormony slinivky břišní. Pankreas (slinivka břišní) je žláza s endokrinní i exokrinní sekrecí. Endokrinní částí této žlázy jsou Langerhansovy ostrůvky, které jsou tvořeny α -buňkami produkujícími glukagon, β -buňkami produkujícími inzulín a δ -buňkami, které produkují gastrin a somatostatin. [33]

Inzulín je hormonem bílkovinné povahy, jehož hlavním úkolem je snížit hladinu glukózy v krvi a zvýšit její využití (utilizaci). Zvýšená koncentrace glukózy v plazmě tedy způsobuje zvýšenou sekreci inzulínu. Sekreci inzulínu zvyšuje také stimulace vagu, což považujeme za neurogenní stimulaci, dále sekretin, gastrin a somatotropní hormon. Inzulín snižuje glykémii zvýšenou propustností membrán a napomáhá vstupu glukózy do buněk. Uvnitř buněk způsobí zvýšenou aktivitu enzymů, které se podílí na zpracování glukózy. Díky tomu zvyšuje tvorbu tuků v tukových buňkách a zvyšuje tvorbu glykogenu. Kromě toho také umožňuje potřebným aminokyselinám vstup do buněk, čímž napomáhá růstu. Snižuje katabolismus (rozpad) tuků a bílkovin. Aktivuje sodíko-draslíkovou pumpu a napomáhá vstupu draslíku do buněk, čímž ovlivňuje klidový membránový potenciál. [33]

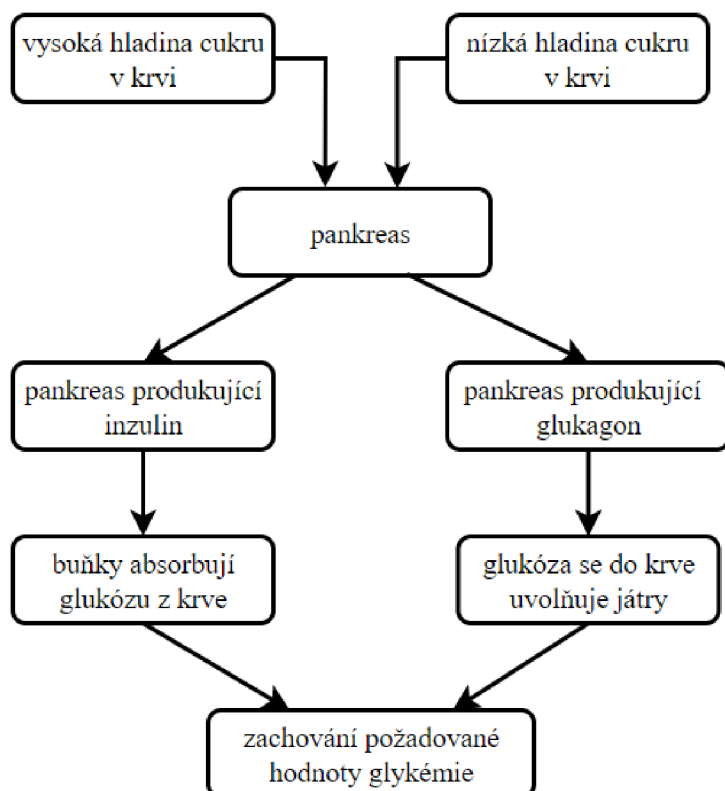
Zvýšená sekrece inzulínu se projevuje tzv. hypoglykemií (sníženou hladinou cukru v krvi). Může být projevem například přítomnosti nádoru, který produkuje inzulín. K příznakům hypoglykémie patří hlad, slabost, ospalost nebo i bezvědomí. Projev některých zmíněných příznaků je způsoben nedostatkem energie v CNS (centrální nervové soustavě). Glukóza je totiž hlavním zdrojem energie pro mozek. Jiné příznaky

(třes, bledost, pocení) mohou být způsobeny aktivací sympatiku. Hypoglykémii tělo chápe jako život ohrožující stav, proto se začnou do těla vyplavovat některé stresové hormony (kortizol, adrenalin aj.). [33]

Dlouhodobě snížená sekrece inzulínu má za následek onemocnění zvané úplavice cukrová (diabetes mellitus), které je nejčastěji doprovázené hyperglykemií. Při zvýšené hladině cukru v krvi se přebytečná glukóza začíná vylučovat močí (glykosurie). Právě glykosurie je příčinou polyurie (časté močení) a polydipsie (žíznivost). Zvýšená glykémie může vést k odvodnění organismu, jelikož je glukóza osmoticky aktivní a při vylučování se sebou strhává i molekuly vody. Nemožnost „prostupu“ glukózy do buněk způsobuje poruchy v metabolismu tuků i bílkovin. Hyperglykémie ztěžuje navázání glukózy na bílkoviny v těle, což může mít za následek poškození stěn velkých cév nebo kapilár. [33]

Antagonistou inzulínu je glukagon, který se tvoří v α -buňkách slinivky břišní a vyplavuje se při snížené koncentraci glukózy v krvi a při fyzické zátěži. Glukagon zvyšuje glukoneogenezi (tvorbu glukózy z glycerolu a aminokyselin), zvyšuje glykémii zvýšením glykogenolýzy v játrech (štěpení glykogenu na glukózu) a zvyšuje vyplavování inzulínu. [33][45]

Schéma mechanismu udržování požadované hladiny glykémie je na Obrázek 3.1.



Obrázek 3.1: Mechanismus udržování požadované glykémie, upraveno z [46]

3.2 Diabetes Mellitus

Diabetes Mellitus (DM) se řadí do heterogenní skupiny chronických onemocnění, která se vyznačují hyperglykemií, tedy zvýšenou hladinou glukózy v krvi. Nastává v případě, že slinivka břišní není schopna produkovat inzulin, nebo tělo není schopno inzulin správně využít. Příčinou onemocnění je tedy relativní nebo absolutní nedostatek inzulinu. Nedostatek inzulinu může vést i k dalším zdravotním komplikacím, které diabetes doprovází (kardiovaskulární onemocnění, nefropatie, neuropatie, oční onemocnění postihující především sítnici aj.). [5][6][7]

IDF (International Diabetes Federation) uvádí, že cukrovkou trpí přibližně 537 milionů lidí na světě ve věku 20-79 let. Předpokládá se, že do roku 2045 by počet diabetiků mohl přesáhnout 783 milionů. Diabetes lze klasifikovat do čtyř základních typů: DM 1. typu, DM 2. typu, gestační diabetes a diabetes související se specifickým druhem patologie. U diabetu 1. typu se vyskytuje absolutní nedostatek inzulinu, který se vyvíjí jako autoimunitní onemocnění. Relativní nedostatek se projevuje u DM 2. typu, kterým trpí přibližně 90 % diabetiků. Prediabetes souvisí s horší glukózovou tolerancí a indikuje vyšší riziko rozvoje diabetu 2. typu. [5][6][7][8]

DM 1. typu, označován také jako inzulin-dependentní diabetes, je způsoben absolutním nedostatkem inzulinu, který je důsledkem destrukce β -buněk slinivky břišní. Často se jedná o autoimunitně podmíněné onemocnění, které se na rozdíl od DM 2. typu objevuje už v dětství. [5][7]

DM 2. typu, také známý jako non-inzulin dependentní diabetes, je nejčastějším typem diabetu. Příčinou tohoto onemocnění je tzv. inzulinová rezistence, tedy neschopnost buněk reagovat na inzulin, což vede k hyperglykémii. Je známo, že výskyt DM 2. typu souvisí se životním stylem a jistý vliv hraje i genetické predispozice. K rozvoji DM 2. typu může přispět fyzická nečinnost, kouření, vysoká konzumace alkoholu, sedavý životní styl a především obezita, která přispívá až přibližně k 55 % DM 2. typu. [7][47][46]

Gestační diabetes je typ diabetu, který se může objevit kdykoliv během prenatálního období a odezní po porodu. U všech těhotných žen je prováděn test na gestační diabetes, jelikož během těhotenství může být neléčený gestační diabetes rizikový pro miminko i matku. [7]

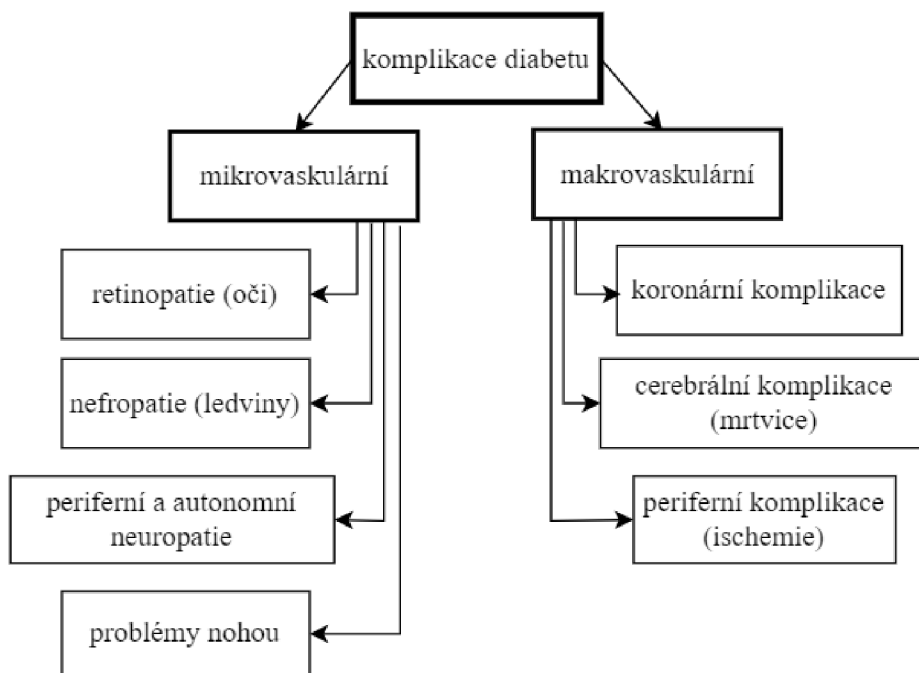
Příkladem diabetu souvisejícím se specifickou patologií může být diabetes, který byl spuštěn virovou infekcí, jejímž následkem je destrukce β -buněk. Dále diabetes vyvolaný léky a chemikáliemi, které narušují sekreci nebo působení inzulinu. Spouštěčem diabetu mohou být také endokrinní poruchy způsobující nadměrnou sekreci hormonů, které jsou antagonisty inzulinu (například u lidí s Cushingovým syndromem). [7]

3.3 Komplikace doprovázející diabetes

Dlouhodobě nekontrolovaný diabetes může vést k mnoha dalším zdravotním komplikacím. Diabetes Mellitus je svázán s kardiovaskulárními chorobami. Uvádí se, že přibližně 70 % úmrtí diabetiků je způsobeno právě některou z doprovázejících kardiovaskulárních chorob. Pravděpodobnost výskytu cévní mozkové příhody, nebo infarktu myokardu je u diabetiků téměř dvojnásobná. [48]

Komplikace při diabetu mohou být různé povahy (Obrázek 3.2). Například makrovaskulární onemocnění způsobené diabetem může vést právě k mozkové mrtvici, srdečnímu selhání, infarktu, angině pectoris, křečovitým bolestem nohou apod. Diabetici častěji trpí aterosklerózou, která se u nich projevuje mnohem dříve a závažněji než u nediabetické populace. V důsledku možného výskytu ischemické choroby dolních končetin je riziko amputace dolních končetin u diabetiků až 30krát vyšší. [46]

Diabetes je spjat také s mikrovaskulárními chorobami (tzv. mikroangiopatické komplikace), které vedou ke zvýšení permeability v celém těle. K takovýmto komplikacím řadíme diabetickou neuropatii, retinopatii a nefropatii. Ve většině případů se jedná spíše o pozdní projevy diabetu, než-li o komplikace. V klinické praxi je vždy nutné myslet na možný vznik diabetické nefropatie, jelikož v České republice patří mezi hlavní indikace k zařazení do dialyzační terapie. Toto onemocnění je charakterizováno tzv. mikroalbuminurií (ztráta albuminu do moči). Ve vyspělých zemích je diabetická retinopatie jednou z nejčastějších příčin slepoty. Včasně zachycení rozvoje retinopatie ještě umožňuje zahájit laserovou terapii. V rámci prevence je důležitá návštěva oftalmologa minimálně jednou ročně. Při prohlídce by měla být provedena kontrola zrakové ostrosti, a především očního pozadí. Další možnou komplikací při diabetu je neuropatie. Při periferní neuropatii bývají postiženy periferní cévy, což způsobuje značné oslabení motorických a sensorických funkcí. Projevem může být oslabená chůze, bolesti chodidel a končetin a také oslabení proximálních svalů. Autonómni neuropatie souvisí s postižením sympatických a parasympatických nervů a může ovlivnit funkci řady systémů v těle. Postihuje například kardiovaskulární systém (tachykardie), pupilární systémy (snížení reflexu na světlo), gastrointestinální systémy (průjmy, zácpy aj.). [46][48]



Obrázek 3.2: Komplikace diabetu, převzato a upraveno z [46]

3.4 Invazivní měření glykémie

Metody stanovení glykémie lze rozdělit na invazivní (zásah do organismu – odběr krve) a neinvazivní. Téměř všechny testy pro stanovení hladiny cukru v krvi běžně používané v praxi jsou invazivní. [46]

3.4.1 Self-monitoring

Self-monitoring je jeden z nejjednodušších způsobů zjištění aktuální hodnoty glykémie. Kromě informace o aktuální hodnotě glykémie poskytuje self-monitoring možnost dlouhodobého sledování glykémie a kontrolu stavu diabetiků. Výsledkem pravidelné kontroly poté může být informace vedoucí k rozhodnutí o úpravě léčby. Ke zjištění hodnoty hladiny cukru v kapilární krvi slouží přístroje zvané glukometry. [49]

K získání vzorku krve je používána lanceta, kterou se vyšetřovaná osoba píchne do konečku prstu. Kapku kapilární krve je poté třeba nanést na testovací proužek, který se zasune do glukometru. Glukometry v zásadě fungují na jednom ze dvou základních principů. Fotometrický princip využívá vlnovou délku odraženého světla od barevného produktu, který vznikne reakcí glukózy v kapce krve s enzymem. Druhý přístup je

elektrochemický a je založen na měření množství elektronů, které vznikají opět při enzymatické reakci. Elektrony jsou poté převedeny na elektrický signál. [49]

3.4.2 Orální glukózový toleranční test (oGTT)

Orální glukózový toleranční test se nejčastěji provádí u těhotných žen ke kontrole gestačního diabetu. Testovaná osoba musí vypít nápoj obsahující 75 gramů cukru za standardních podmínek. V daných časových intervalech jsou odebírány a testovány vzorky krve. Po dvou hodinách od oGTT musí hladina cukru v krvi klesnout pod 11 mmol/l. V opačném případě je přítomnost diabetu potvrzena. [5][46]

3.4.3 Glykovaný hemoglobin (HbA1c)

Glykovaný hemoglobin (HbA1c) je protein, který lze využít k diagnostice diabetu a průběžné kontrole glykémie. Test s využitím glykovaného hemoglobinu reflektuje průměrnou hladinu cukru v krvi za posledních cca 90 dní. [50]

Hemoglobin se nachází v červených krvinkách a podílí se na přenosu kyslíku celým tělem. Hemoglobin se z krevního řečiště glykuje (potahuje se glukózou). Zvýšená hladina glukózy se proto projeví společně se zvýšenou hladinou glykovaného hemoglobinu. Tímto způsobem lze sledovat hladinu cukru v krvi právě za poslední tři měsíce, jelikož životnost erytrocytu je v průměru 90 dní. HbA1c vyšší než 6,5 % poukazuje na přítomnost diabetu. Procento HbA1c pak odpovídá průměrné hladině glukózy v těle za poslední 90 dní. [50]

3.4.4 Kontinuální monitorování glykémie

Kontinuální monitorování hladiny glukózy umožňuje měřit hladinu glukózy v intersticiální tekutině. Součástí poměrně nově vyvinuté technologie je senzor, který je umístěn pod kůži. Data z tohoto senzoru mohou být získána ručně, nebo bezdrátově přenesena do přijímače pro rychlé zobrazení naměřených hodnot. Senzor může měřit glykémii každých pět minut v reálném čase, což umožňuje uživateli sledovat denní i noční profil glukózy. Pacient má dostatek dat, které může sdílet se svým lékařem. Získané informace mohou vést ke změně medikace, výživy apod. Součástí některých senzorů mohou být i alarmy upozorňující na hypoglykémii. Sensory mohou být integrované také s inzulínovou pumpou, díky čemuž lze upravit či pozastavit podávání inzulínu. [51]

Příkladem chytrého zařízení, které kontinuálně monitoruje glykémii je senzor FreeStyle Libre (Obrázek 3.3). Jedná se o malý kulatý senzor o průměru 35 mm, který se aplikuje na zadní stranu horní části paže. Během aplikace je zavedeno tenké ohebné vlákno do podkoží. FreeStyle Libre měří glykémii v intersticiální tekutině. Jde tedy o chytré zařízení, které je sice přenosné, na druhou stranu je invazivní. K používání senzoru FreeStyle Libre je potřebná čtečka skenující senzor. Čtečka je po naskenování schopna ukázat aktuální hodnotu ze skenu a také hodnoty glykémie za posledních 8 hodin. Do její paměti se vejde 90denní historie měření. Součástí systému FreeStyle

Libre je také software generující řadu různých reportů, které pomáhají lékařům s analýzou naměřených hodnot glykémie. [52]

Výhodou senzoru je určitě možnost kontinuálního měření glykémie bez nutnosti píchání do prstu. Senzor je malý, přenosný a pohodlný. Jeho životnost je přibližně 14 dní a je voděodolný. Nevýhodou může být nutnost používání druhého zařízení pro získání dat ze senzoru a invazivní způsob měření. [52]



Obrázek 3.3: Senzor FreeStylelibre (vlevo) a potřebná čtečka (vpravo) [46]

3.5 Neinvazivní měření glykémie

V posledních letech lze sledovat pokusy o měření glykémie neinvazivně. Možnost měření bez jakéhokoli zásahu do organismu (odběr krve) pacientům nabízí daleko větší komfort. Existuje řada odlišných metod a postupů, jak stanovit hladinu cukru v krvi neinvazivním způsobem. Bohužel málokteré metody jsou schopné poskytnout výsledky srovnatelné s výsledky získanými pomocí invazivních metod. K metodám, které se ukázaly jako vhodné pro komerční trh patří například blízká červená spektroskopie (Near Infrared Spectroscopy, NIR), metoda založená na fluorescenci a mikrovlnná metoda. [53]

Vzhledem k rychlému rozvoji umělé inteligence, strojového učení a hlubokého učení existuje řada studií, které se zabývají klasifikací diabetu, nebo predikcí glykémie z EKG a PPG signálu. Velké společnosti jako je například Apple nebo Samsung slibují možnost sledování glykémie pomocí chytrých hodinek. V následujícím textu je blíže popsáno, jak lze predikovat diabetes z EKG signálu. Možnost stanovení glykémie nebo klasifikace diabetu z PPG signálu je popsána v kapitole 5.1 a 5.2.

3.6 Detekce diabetu z EKG signálu

Autoři článku [46] se zabývali detekcí diabetu z EKG signálu. Zvýšená hladina glukózy v krvi způsobuje tzv. prekapilární poškození, které ovlivní normální funkci endoteliálních buněk. Narušená funkce endoteliálních buněk způsobuje zablokování průchodu oxidu dusnatého (NO), který je velice důležitý pro vazodilataci. Kromě toho hyperglykémie snižuje aktivaci fosforylační kaskády, čímž se snižuje množství endoteliální NO syntázy potřebné k syntéze NO. Oba tyto faktory značně snižují dostupnost NO. Dochází k produkci volných kyslíkových radikálů, které aktivují proteinkinázu C zvyšující vazokonstrikci. V důsledku poškozeného endotelu se zvyšuje aktivita krevních destiček. Leukocyty, monocyty i krevní destičky jsou silně adherovány k endotelu, snižuje se fibrinolytická aktivita a zvyšuje se srážlivost krve. Tukový materiál se může nakupit na stěně krevní cévy a vést k tzv. ateroskleróze, což je často přidružená komplikace diabetu. S diabetem je spjata i funkce srdce. V průběhu diabetu se může objevit i kardiovaskulární autonomní neuropatie (CAN). Je dokázáno, že diabetes se projevuje i změnami v EKG záznamech (ST-T změny, změna variability srdeční frekvence, dlouhé QTc aj.). Bylo potvrzeno, že k prediktorům úmrtí u pacientů trpících diabetem patří QT, QTc a ST disperze. U pacientů s DM je patrné prodloužení QTc intervalu. Z výše zmíněných důvodů se autoři rozhodli predikovat diabetes z EKG signálu s využitím HRV analýzy. Změny v autonomní nervové soustavě (ANS), které mohou být dány hyperglykemií, je možné dobře popsat právě pomocí variací intervalů RR, tedy HRV analýzou. [46]

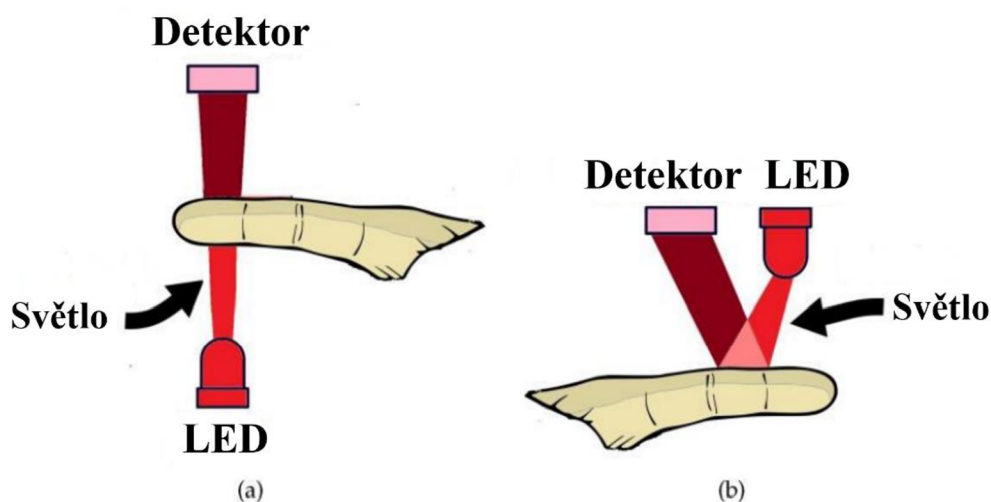
Pro HRV analýzu je třeba spočítat různé časové i frekvenční příznaky (průměr RR intervalů, rozptyl RR intervalů, hustota výkonového spektra aj.), lze využít vlnkovou transformaci nebo například metodu analýzy fluktuací (DFA). Různé studie prokázaly, že časové parametry jako průměr RR intervalu, RMSSD, SDNN, NN50 aj., dosahují u diabetických pacientů nižších hodnot. Po extrakci zmíněných příznaků bylo použito pro klasifikaci diabetu hluboké učení, jiní autoři vyzkoušeli metodu podpurných vektorů – SVM. [46][54][55][56]

HRV analýzu lze provést i na PPG signálech, o kterých pojednává následující kapitola 4.

4. FOTOPLETYSMOGRAFIE

Fotopletysmografie (též fotoelektrická pletysmografie) je neinvazivní optická metoda, která umožňuje detekovat změny objemu krve v mikrovaskulární tkáni, ke kterým dochází v důsledku pulzace. Tato technika je známá již řadu let a používá se například v pulzní oxymetrii. Velký „boom“ však zaznamenala v posledních 20ti letech s vývojem nositelných zařízení. Z PPG signálu lze stanovit například TF, SpO₂, DF, KT, lze provést vyšetření poddajnosti cév aj. Možnosti stanovení zmíněných parametrů z PPG signálu jsou diskutovány v kapitole 2. PPG signál lze získat z oblastí, které jsou dostatečně prokrvené. Většinou se měří z prstu (smartphone), ušního lalůčku, zápěstí (chytré hodinky), nebo nohou. Za výhodu lze považovat jednoduchost této metody, dostupnost a nízkou cenu nositelné elektroniky snímající PPG. [24][57]

PPG signál lze získat s využitím světelného zdroje (například LED diody) osvětlujícího tkáň v oblasti, kterou chceme snímat. Zdrojem světla jsou většinou infračervené diody, jelikož je u nich malý rozdíl v absorpci světla mezi krví okysličenou a neokysličenou. Druhou základní komponentou při tomto měření je fotodetektor, který může snímat světlo prošlé (průsvitový, též transmisní režim). V tomto případě se snímaná oblast, například prst, nachází mezi světelným zdrojem a fotodetektozem. Detektor může snímat i světlo odražené (reflexní režim), kdy se zdroj světla a fotodetektor nachází vedle sebe. Reflexní přístup snímání je konstrukčně jednodušší. Dva odlišné principy získání PPG signálu demonstruje obrázek 4.1. [21][24][57]

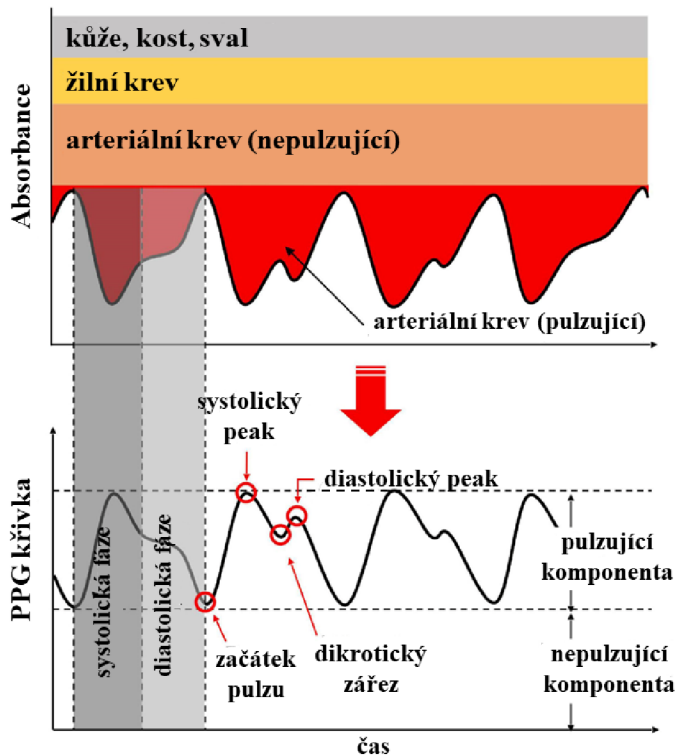


Obrázek 4.1: Transmisní (a) a reflexní (b) snímání, upraveno z [57]

4.1 Fotopletysmografická křivka

Obecně má PPG křivka složku pulzující (AC) a nepulzující (DC). Pulzující složka kopíruje srdeční cyklus a souvisí s objemovými změnami krve. Její frekvence se typicky pohybuje kolem 1 Hz. Může být ovlivněna vazodilatací a cévním tonem. Naopak složka nepulzující, tzv. stejnosměrná složka, se mění velice pomalu, obsahuje nízkofrekvenční složky a je ovlivněna dýcháním, termoregulací a případnou vazomotorickou aktivitou. DC složka může podléhat také vlivu složení tkáně v místě měření a vnějším faktorům (okolní světlo). [57][58]

Na PPG křivce znázorněné na Obrázku 4.2 lze vidět vzestupnou (anakrotickou) a sestupnou (katakrotickou) část. Anakrotická fáze souvisí se zvýšeným krevním objemem během srdeční kontrakce a katakrotická fáze souvisí se sníženým objemem kapilární krve při srdeční dilataci. To v podstatě znamená, že vzestupná část PPG křivky odpovídá fázi systolické a sestupná část pak fázi diastolické. Systolický vrchol tedy odpovídá bodu, ve kterém je objem krve maximální. Následuje pokles během diastoly a opětovný vzestup, ke kterému dochází v důsledku tlakového gradientu v opačném směru, než je směr toku krve, před uzavřením aortální chlopně. U lidí se zdravými tepnami lze na PPG křivce pozorovat dikrotický zářez, který se mění s poddajností cév i vaskulárním tonusem. Dikrotický zářez se u starších osob a osob s kardiovaskulárním onemocněním nemusí vůbec vyskytovat. Stejně tak se s věkem snižuje vzdálenost mezi systolickým a diastolickým peakem. [24][57][58]



Obrázek 4.2: PPG křivka, upraveno z [58]

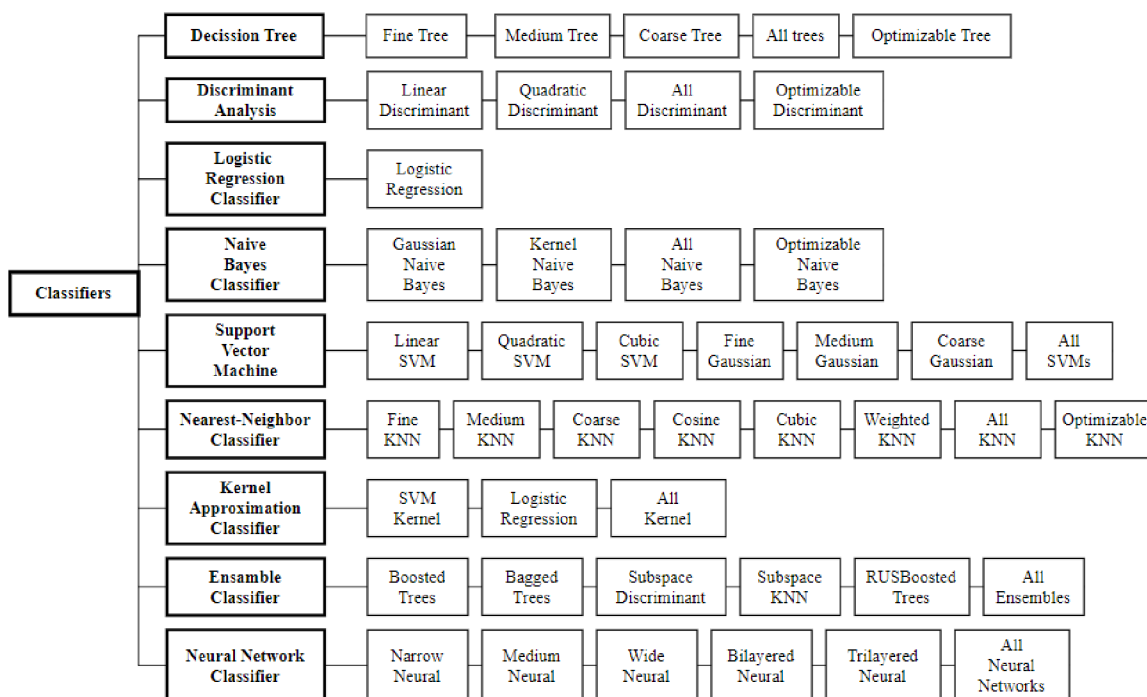
5. METODY STANOVENÍ GLYKÉMIE Z PPG

Tato kapitola se zabývá možnostmi stanovení glykémie z PPG signálu. V současné době existují metody klasifikační i regresní, ale zatím jich není mnoho. Cílem klasifikačních metod je většinou klasifikace PPG záznamů do skupin diabetes/ non-diabetes. Regresní metody se snaží odhadovat konkrétní hodnotu glykémie.

5.1 Klasifikace glykémie pro predikci diabetu z PPG signálu

Autoři článku [59] se zabývali klasifikací hladiny glukózy v krvi pro časnou detekci diabetu. Pro svůj výzkum používali PPG signály získané prostřednictvím vytvořeného elektronického obvodu umístěného ve snímači prstu. Ke klasifikaci využili metod strojového učení (Obrázek 5.1). Pro získání referenčních hodnot využívali glukometr. Dataset je tvořen 400 PPG signály, které byly změřeny na 80 subjektech. Kromě samotné PPG vlny pro každý subjekt zaznamenali také srdeční frekvenci, krevní tlak a saturaci krve kyslíkem. Jedno měření trvalo 15 minut, přičemž každých 5 minut byla zaznamenána referenční hodnota glykémie. Měřené osoby s hodnotou glykémie vyšší než 200 mg/dl (11,1 mmol/l) byly klasifikovány jako diabetici. Signály z datasetu byly rozděleny na základě kvality do tří skupin: kvalitní, akceptovatelné a nekvalitní. Kvalita signálu byla stanovena manuálně. Do kategorie „kvalitní“ spadaly pouze PPG signály, z nichž lze určit TF a mají patrný dikrotický zářez. Kategorie „akceptovatelné“ označuje záznamy, které sice nemají patrný dikrotický zářez, ale lze z nich určit TF. Dále byly předzpracovány a využity pouze signály označené jako kvalitní nebo akceptovatelné. V rámci této metody nejsou počítány žádné specifické příznaky z PPG signálu. Všechny PPG signály jsou rozděleny na úseky o délce 2100 vzorků a extrakce příznaků je prováděna vzorek po vzorku. Znamená to tedy, že pro každý úsek PPG signálu je označeno 2100 vzorků jako 2100 příznaků, které následně vstupují do učení. [59]

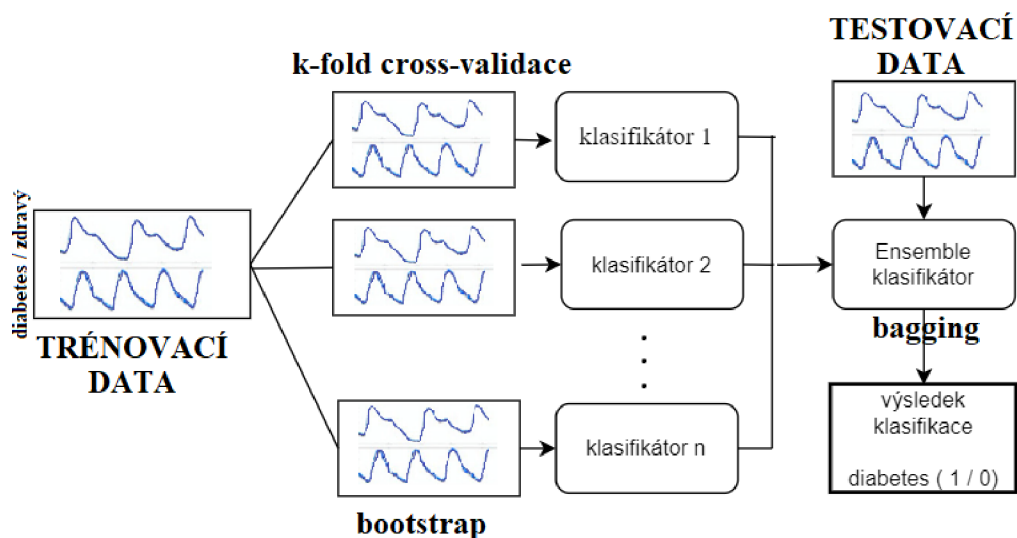
Autoři článku využili několik různých metod strojového učení pro klasifikaci PPG signálů. Pro trénování i testování byla sice použita data ze stejného datasetu, ale s využitím cross-validace. [59]



Obrázek 5.1: Klasifikátory, které byly testovány, upraveno z [59]

Autoři článku následně využili i tzv. Ensemble metody, které kombinují výsledky jednotlivých klasifikátorů tak, aby došlo ke zlepšení výsledné klasifikace a stability modelu. Ensemble metody se opírají o “bagging” a “boosting”. Studie využívá pro neinvazivní klasifikaci glykémie konkrétně Ensemble bagged trees algoritmus (EBTA) na základě tvaru PPG, pro které je určeno výše zmíněných 2100 vzorků daného PPG (odpovídá 2100 příznakům). Na Obrázek 5.2 lze vidět schéma modelu. Použitá metoda je ideální pro problémy s malým trénovacím setem. Pro klasifikaci využívá tzv. většinového hlasování. [59]

Autoři vyhodnotili několik různých přístupů pro klasifikaci (Obrázek 5.1). Nejlepším přístupem se jevila právě zmiňovaná metoda EBTA, která dosahovala přesnosti 98 %. [59]



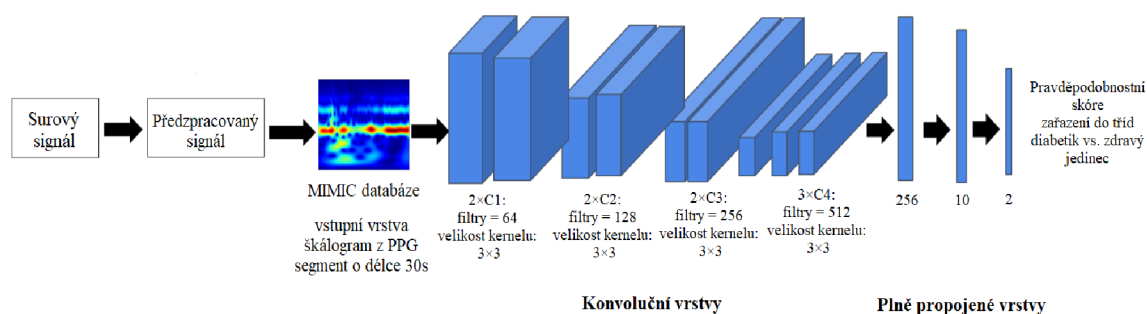
Obrázek 5.2: Schéma navrhovaného modelu využívajícího EBTA, převzato a upraveno z [59]

Jiný přístup zvolili autoři [60], kteří klasifikovali pacienty do dvou skupin (diabetici/zdraví jedinci) s využitím hlubokého učení. Pro klasifikaci využili databázi PhysioNet MIMIC-III, která obsahuje PPG signály, EKG signály, tlakovou křivku, dechovou křivku, ale také informace o TF, SO₂ aj. Záznamy z databáze označené kódem ICD-9 pochází od pacientů s diabetem. Byla použita data od 808 pacientů. PPG signály byly nasnímané z prstu a použita vzorkovací frekvence je 125 Hz. Databáze MIMIC-III obsahuje již normalizované signály.

Autoři provedli filtraci s využitím Butterworthovy horní propusti (mezní frekvence 0,5 Hz) a signály rozdělili na segmenty o délce 30 s. Segmenty prošly algoritmem kontroly kvality signálu na základě porovnávání šablon. Poškozené segmenty byly rekonstruovány pomocí extrahovaných peaků. Dále byla provedena lineární regrese rekonstruovaného signálu s využitím metody nejmenších čtverců s původním signálem pro zjištění korelace. Segmenty s korelačním koeficientem vyšším než 0,8 jsou považovány za kvalitní a jsou použity v procesu učení. Vstupem do vytvořené sítě jsou však RGB škálogramy (anglicky scalograms) kvalitních PPG segmentů, které lze definovat jako absolutní hodnotu kontinuální vlnkové transformace signálu (CWT). Mohou identifikovat nízkofrekvenční a rychle se měnící frekvenční složky signálu. [60]

Architektura navržené konvoluční neuronové sítě (CNN) je na Obrázek 5.3. Každý škálogram má dimenzi 640×480×3. Nejdříve dojde ke změně dimenze na 320×240×3, aby se snížila výpočetní náročnost během učení. Jak lze vidět z obrázku CNN obsahuje

9 konvolučních vrstev. Součástí jsou také „pooling“ vrstvy. „Pooling“ vrstvy slouží ke zmenšení mapy dle zvolené velikosti masky a kroku. Dále je také „pooling“ užitečný pro extrakci příznaků s vyšší důležitostí. Princip výpočtu je takový, že maska o zvolené velikosti je posouvána po vstupní mapě a z daných hodnot, které maska obsahuje je vybrána buď maximální hodnota, nebo průměrná hodnota na odpovídající pixel ve výstupní mapě. Podílí se tak i na snížení výpočetní náročnosti. Ve všech vrstvách je použit kernel o velikosti 3×3 . Následují 2 plně propojené vrstvy a poslední klasifikační vrstva (Softmax) obsahující 2 neurony pro 2 klasifikační třídy (diabetici / zdraví jedinci). [60]

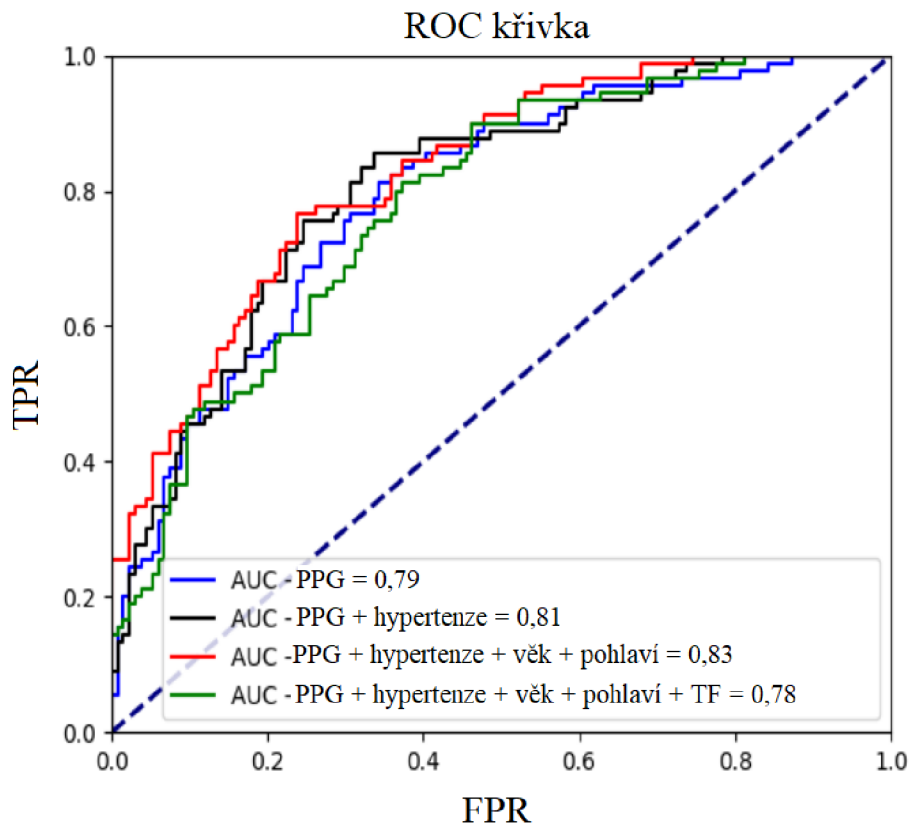


Obrázek 5.3: Architektura CNN, převzato a upraveno z [59]

Trénovací sada obsahuje data celkem od 584 pacientů, testovací sada pak od 224 pacientů. Pro klasifikaci byly vyzkoušeny 4 odlišné vstupy:

- PPG
- PPG + klasifikace hypertenze
- PPG + klasifikace hypertenze + věk + pohlaví
- PPG + klasifikace hypertenze + věk + pohlaví + TF

Jak lze vidět na Obrázek 5.4, kombinace PPG, klasifikace hypertenze, věku a pohlaví poskytuje nejlepší výsledek pro klasifikaci. [60]

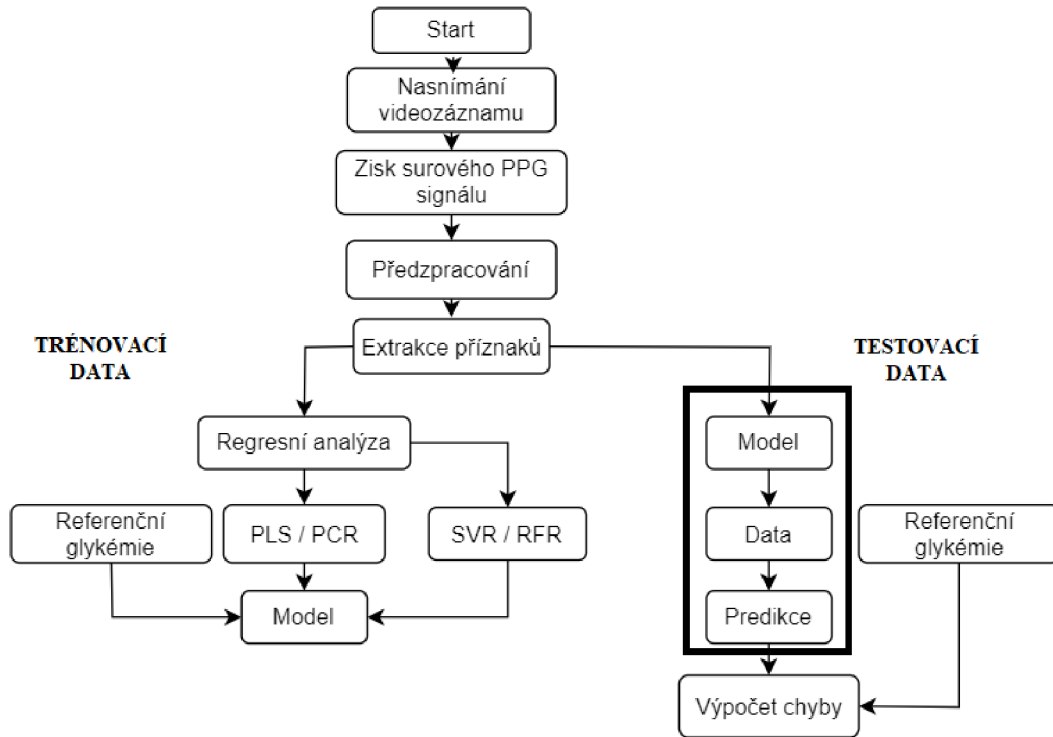


Obrázek 5.4: Architektura CNN, převzato a upraveno z [59]

5.2 Možnosti stanovení konkrétní hodnoty glykémie z PPG signálu

Autoři článku [6] se zabývají predikcí konkrétní hodnoty glykémie z PPG signálu. V rámci této studie bylo měřeno 52 subjektů. Měření bylo prováděno z prstu s využitím chytrého telefonu. Nasnímané videozáznamy byly dlouhé 60 sekund. Následně se videozáznam získaný chytrým telefonem převedl do programového prostředí, kde se pro extrakci PPG signálu využila červená složka videa. Za účelem vytvoření přesnějších modelů pro predikci glykémie, byla provedena analýza kvality signálů a v dalším zpracování se používaly pouze kvalitní signály. Ze všech PPG signálů byl odstraněn vysokofrekvenční šum použitím Gaussovského filtru. Součástí předzpracování všech signálů byla také korekce základní linie (baseline) pomocí asymetrické metody nejmenších čtverců. Z takto upravených signálů byly extrahovány příznaky. Autoři této metody využívají k predikci příznaky vycházející z morfologie samotného signálu: pozice systolického peaku, pozice diastolického peaku, vzdálenost mezi systolickým a diastolickým peakem a také peaky první a druhé derivace signálu. Pro predikci glykémie bylo využito několika metod strojového učení (Obrázek 5.5): PCR (Principal component regression, regrese hlavních komponent), PLS (Partial least

square model, částečná regrese nejmenších čtverců), SVR (Support Vector Regression, využití podpůrných vektorů pro regresi) a RFR (Random forest regression, regrese s využitím náhodného lesu). PLS model dosahoval nejnižší standardní chyby, konkrétně 17,02 mg/dL, což je 0,9446 mmol/l. [6]



Obrázek 5.6: Navrhovaný model pro stanovení glykémie z PPG signálu, převzato a upraveno z [6]

Autoři [61] vyvinuli svůj vlastní hardwarový systém snímající PPG signál z prstu pro odhad hladiny glukózy v krvi. Systém je schopen snímat PPG signály v transmisním i reflexním režimu. Pro predikci glykémie autoři využívají červenou složku PPG. Signály naměřili na 26 dobrovolnících. Měření probíhalo po dobu 4 minut v obou režimech (reflexní i transmisní). Referenční hodnota byla získána pomocí glukometru. [61]

V rámci předzpracování signálů autoři nejdříve provedli korekci základní linie (baseline) pomocí metody klouzajícího okna (fitting-based sliding window, FSW). Následně byly signály filtrovány s využitím Butterworthova filtru 6. řádu a mezní frekvencí 8 Hz. Signály byly rozděleny na segmenty o délce 3 s. [61]

Z těchto segmentů autoři spočítali 17 různých příznaků: Zero crossing rate (počet průchodů nulou), příznaky vycházející z výkonové spektrální hustoty (špičatost,

průměr, rozptyl), příznaky vycházející z výpočtu Teager-Kaiserovy energie (špičatost, šikmost, rozptyl, průměr), autoregresní koeficienty, TF, DF, SpO₂, index tělesné hmotnosti aj. Příznaky jako TF a DF byly využity pouze pro rozhodnutí, zda je signál kvalitní. [61]

Následně byly vytvořeny dva odlišné algoritmy – náhodný les (RF) a XGBoost (XGB). Oba algoritmy jsou algoritmy využívající rozhodovacích stromů. Pro učení se využívá leave-one-out metody, kdy je během každé iterace jeden vzorek zvolen jako testovací a všechny ostatní vzorky jsou použity pro trénování. Střední absolutní chyba pro RF model dosahovala na signálech snímaných transmisním režimem 12,29 mg/dL (0,7 mmol/l) a 8,73 mg/dL (0,5 mmol/l) na signálech získaných během reflexního snímání. XGB dosahoval obecně lepších výsledků – 11,63 mg/dL, což je přibližně 0,6 mmol/l (transmisní režim) a 8,31 mg/dL, což je 0,46 mmol/l (reflexní režim). Nejlepších výsledků tedy dosahoval XGB model na datech nasnímaných v reflexním režimu. [61]

6. SNÍMÁNÍ DAT

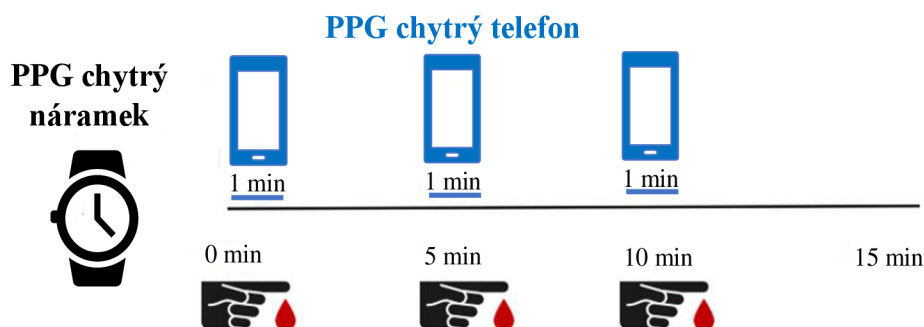
Cílem diplomové práce je klasifikovat hladinu glukózy v krvi do skupin nízká/vysoká glykémie z PPG signálů a poté také predikovat konkrétní hodnotu glykémie. Za tímto účelem byla nasnímana data na 16 subjektech (11 žen a 5 mužů) ve věkovém rozmezí 23 – 80 let. Všechny měřené osoby podstoupily měření dobrovolně a podepsaly informovaný souhlas. PPG signály byly nasnímány pomocí chytrého náramku Empatica wristband. V průběhu měření byl také použit chytrý telefon značky Samsung Galaxy Note 20. Signály nasnímané pomocí chytrého telefonu budou následně použité k ověření funkčnosti vytvořených algoritmů na datech snímaných pomocí chytrého náramku. Referenční hodnota glykémie byla stanovena invazivně s využitím glukometru značky Diamond Mini. Měření se účastnily, jak zdravé osoby, tak osoby s prediabetem nebo diabetem (Tabulka 6.1).

Tabulka 6.1: Počet měřených zdravých osob a osob s diabetem / prediabetem

Počet osob	Zdraví	Diabetes / prediabetes
16	11	5

6.1 Postup měření

Na začátku každého měření byl měřený subjekt seznámen s průběhem měření a musel podepsat informovaný souhlas poté, co byl seznámen s případnými riziky. Měřená osoba je požádána, aby se posadila a v průběhu celého měření byla v klidu. Celé měření trvá 15 minut. Na začátku každého jednoho měření byla odebrána kapka krve z prstu na pravé ruce, aby bylo možné stanovit hodnotu glykémie. Ihned poté byl zapnut chytrý náramek umístěný na levé ruce, který snímal data po celých 15 minut. Dále byla hodnota glykémie stanovena v průběhu měření ještě 2×, vždy po 5 minutách. Po každém odebrání kapky krve a zjištění hladiny glykémie za pomoci glukometru, byl naměřený také minutový záznam s využitím chytrého telefonu. Tento záznam byl získán také z pravé ruky, ale z jiného prstu, než je stanovována glykémie, aby v žádném případě nedošlo k zašpinění čočky fotoaparátu krví, a tím ovlivnění celého měření. Schéma průběhu měření lze vidět na Obrázek 6.1.



Obrázek 6.1: Schéma měření

Jelikož měřených subjektů s diabetem/prediabetem je méně než osob zdravých, byly subjekty s diabetem/prediabetem měřeny dvakrát. Tento krok vede k navýšení počtu záznamů s vysokou glykemií a snaze vybalancovat tak dataset. Celkový počet záznamů a počet záznamů v obou třídách shrnuje Tabulka 6.2. Bohužel jednu osobu s prediabetem se nepodařilo změřit i po druhé.

Tabulka 6.2: Celkový počet naměřených záznamů zdravých osob a osob s diabetem / prediabetem

Počet záznamů	Zdraví	Diabetes / prediabetes
20	11	9

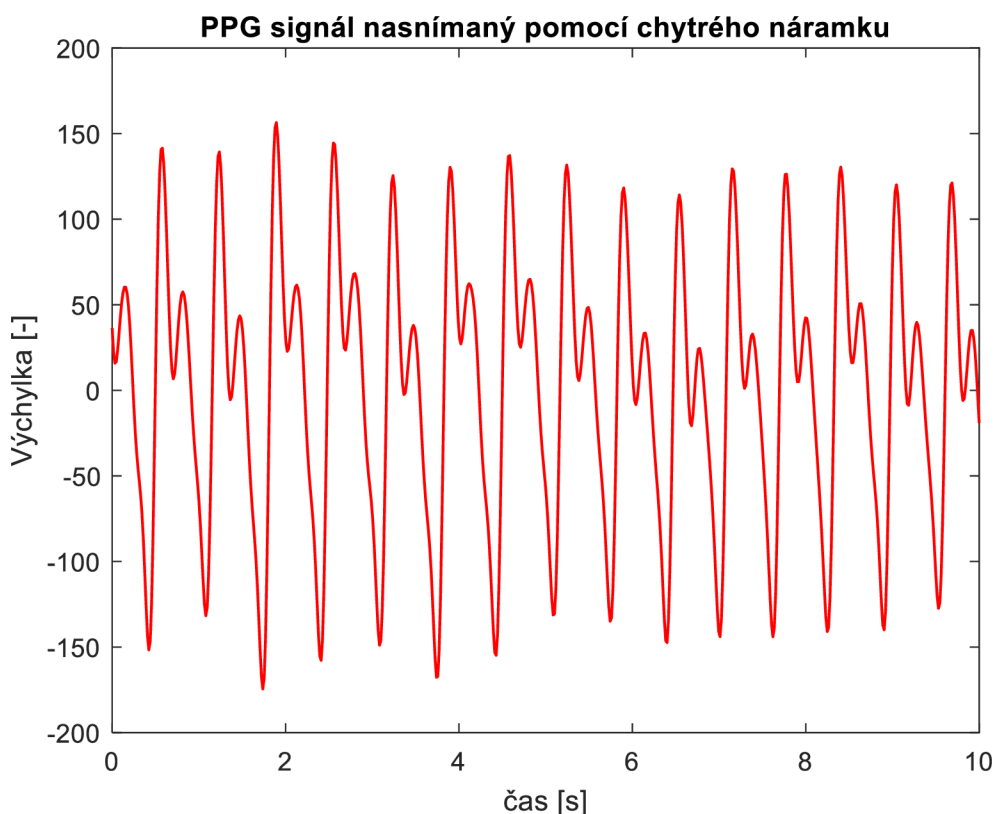
6.2 Snímání PPG pomocí E4 wristband

Chytrý náramek E4 wristband je WHD určené pro výzkumníky a lékaře. Disponuje řadou různých senzorů a funkcí, jak lze vidět na Obrázek 6.2. Součástí je i Cloud pro ukládání dat a aplikace Emaptica Realtime App, kde lze sledovat snímání v reálném čase. [62]



Obrázek 6.2: Specifikace zařízení Empatica E4 wristband, převzato a upraveno z [62]

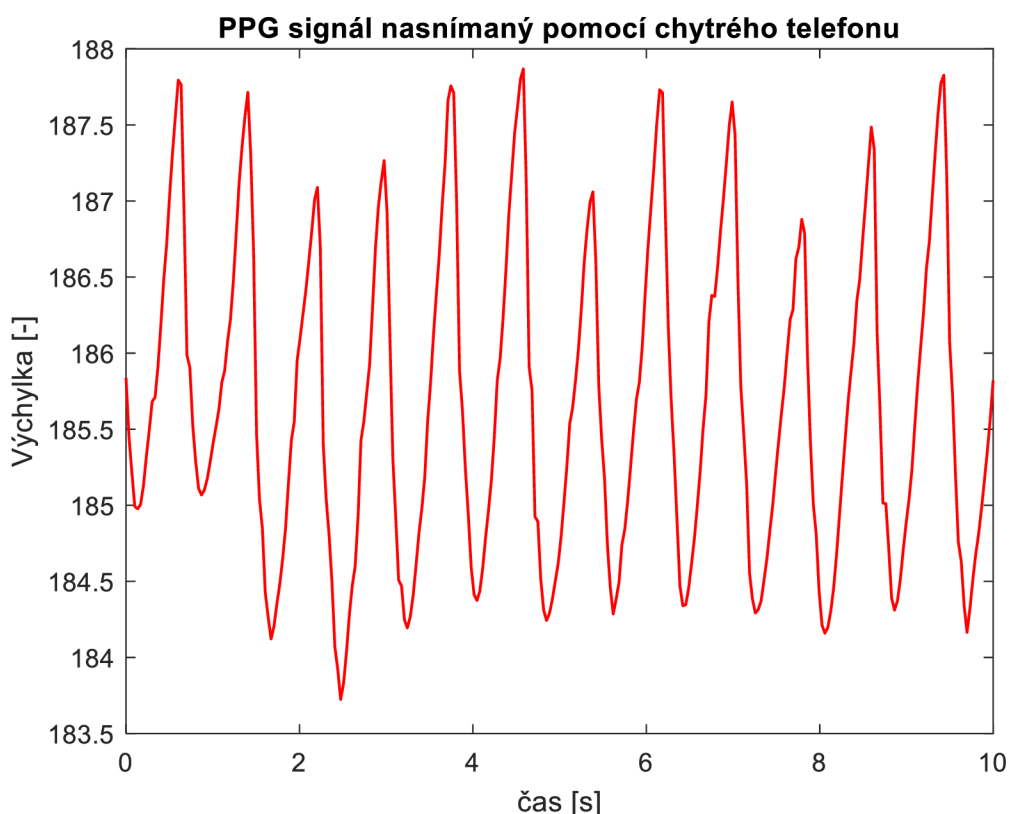
Chytrý náramek snímal PPG signály celých 15 minut. Po každém měření byla data nahrána na cloud a následně stažena ve formátu .csv. Data jsou označena jako BVP (Blood Volume Pulse). Důležité je zmínit, že výsledný BVP signál vzniká dle patentovaného algoritmu a kombinuje světlo zelené i červené. Signály jsou vzorkovány s frekvencí 64 Hz. Získaný signál je již invertovaný. PPG signály byly následně načteny do programového prostředí Matlab (Obrázek 6.3). [63]



Obrázek 6.3: Ukázka 10 s úseku PPG signálu nasnímaného Empaticou

6.3 Snímání PPG pomocí chytrého telefonu

Během jednoho měření byl také 3krát nasnímán PPG signál pomocí chytrého telefonu. Subjekt přiložil svůj prst na objektiv nacházející se na zadní straně chytrého telefonu tak, aby zároveň překryl LED emitující světlo i samotnou čočku fotoaparátu. Rozlišení videa bylo nastaveno na 1080×1920 px a snímkovácí frekvence byla 30 snímků/sekundu. Tímto způsobem byly v rámci jednoho měření získány tři minutové videozáznamy, které byly následně načteny do programového prostředí Matlab pomocí příkazu *VideoReader*. Surový PPG signál pak odpovídá průměru červené složky videa, která se ze všech tří složek jeví jako nejlepší. Signál nasnímaný pomocí chytrého telefonu je na Obrázek 6.4. Takto získaný PPG signál je třeba invertovat, jelikož fotoaparát snímá světlo odražené.



Obrázek 6.4: Ukázka 10 s úseku PPG signálu nasnímaného chytrým telefonem

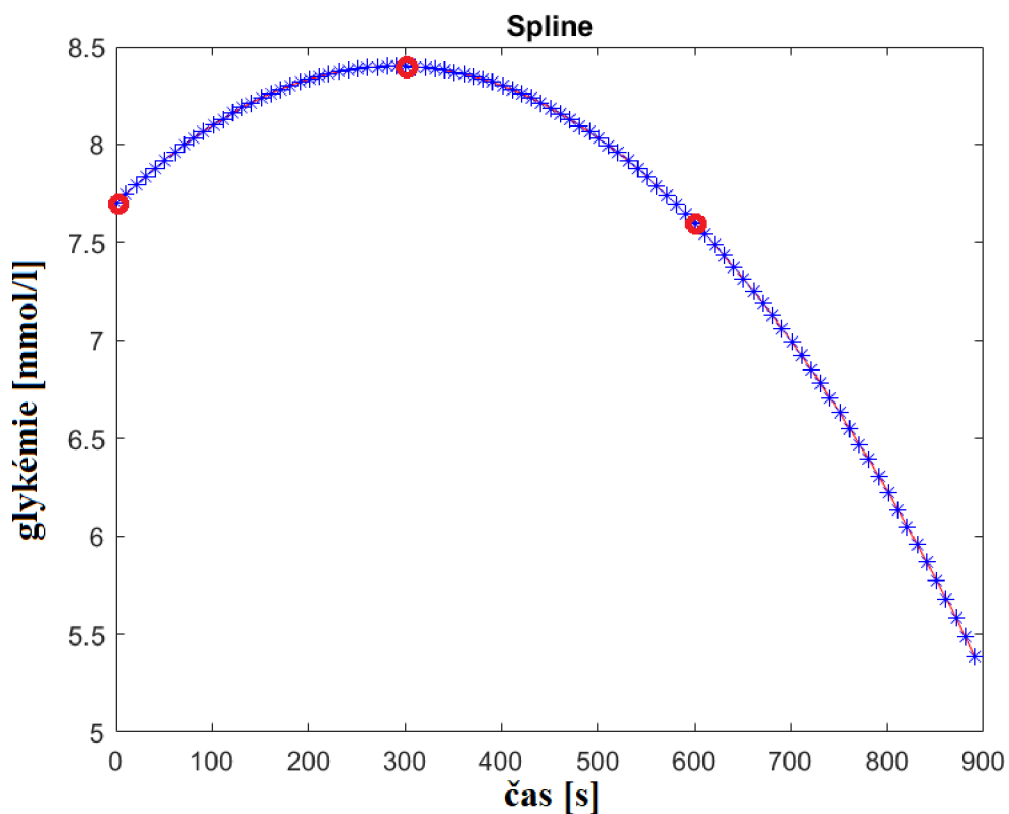
6.4 Měření glykémie

Hladina glukózy v krvi byla měřena invazivně s využitím glukometru Diamond Mini taktěž 3krát během jednoho měření (na začátku měření a následně vždy po 5 minutách). Hodnoty glykémie byly zaznamenány.

Pro klasifikační úlohu je práh pro rozdělení do tříd nízká/vysoká glykémie nastaven na 7,2 mmol/l. Tato hodnota byla určena dle manuálu glukometru, který zobrazuje konkrétní hodnotu glykémie a také danou hodnotu klasifikuje dle prahu 7,2 mmol/l na glykémii v normě (veselý emotikon na displeji glukometru) a zvýšenou glykémii (smutný emotikon na displeji glukometru).

Pro následnou klasifikaci, která je popsána v kapitole 9.2 by bylo dobré získat co nejvíce hodnot glykémie během jednoho měření. V rámci přípravy protokolu měření bylo vyzkoušeno na jednom ze subjektů změřit hodnotu glukózy v krvi během jednoho měření 5×. Tento způsob byl časově velice náročný, a především nekomfortní pro měřenou osobu. Hodnoty se navíc během měření téměř nezměnily, tudíž jsem se rozhodla pro variantu tří měření během 15 minut. Další z možných přístupů, jak získat více hodnot glykémie pro celé měření je pokusit se hodnoty interpolovat. Jelikož po posledním měření glykémie následuje ještě 5 minut snímání, byla zde

zvolena extrapolace (lineární, spline a makima). Žádná z těchto metod se nejevila jako vhodná, jelikož v posledních 5 minutách extrapolované hodnoty glykémie klesaly například k hodnotě 5 mmol/l, přestože změřené 3 hodnoty byly vyšší, než práh 7,2 mmol/l (Obrázek 6.5).



Obrázek 6.5: Extrapolace změřených hodnot glykémie (červené kroužky)

7. PŘÍPRAVA DATASETU A PŘEDZPRACOVÁNÍ SIGNÁLŮ

V následující kapitole je popsána příprava datasetu a dále předzpracování PPG signálů získaných z chytrého náramku i chytrého telefonu. V rámci DP budou využity dva datasety. Jeden dataset obsahuje PPG signály z chytrého náramku a slouží pro trénování i testování. Druhý dataset obsahuje pouze PPG signály nasnímané chytrým telefonem a slouží k ověření funkčnosti algoritmů i na sadě dat získané pomocí jiného zařízení.

7.1 Příprava datasetu

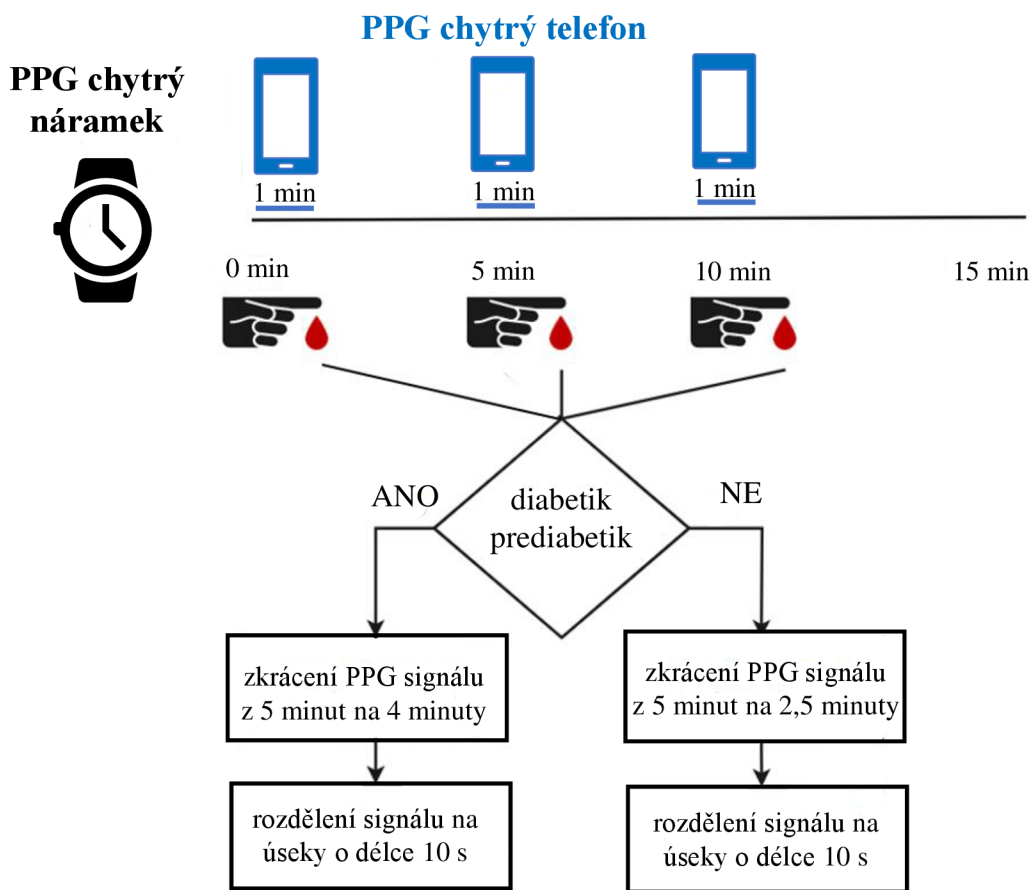
Jak již bylo zmíněno, každé jedno měření obsahuje 15minutový PPG signál z chytrého náramku, tři minutové PPG signály z chytrého telefonu a tomu odpovídající 3 hodnoty glykémie (Obrázek 6.1).

První dataset se tedy týká pouze PPG signálů snímaných chytrým náramkem. Celé 15minutové měření z chytrého náramku lze rozdělit na tři pětiminutové úseky, kterým odpovídají tři různé hodnoty glykémie. Pro 20 měření tedy obsahuje dataset 60 různých hodnot glykémie. Zastoupení hodnot glykémie do jednotlivých tříd zobrazuje Tabulka 7.1. Jelikož dochází k měření hodnot glykémie po každých 5 minutách, rozhodla jsem se pracovat s úseky o délce maximálně 4 minut, aby došlo k eliminaci úseků, které by mohly být nekvalitní v důsledku případného pohybu měřené osoby při přípravě na odběr kapky krve.

Tabulka 7.1: Zastoupení hodnot glykémie dle zvolené prahové hodnoty (chytrý náramek)

Počet měření glykémie	Nízká glykémie ($\leq 7,2$ mmol/l)	Vysoká glykémie ($>7,2$ mmol/l)
$20 \times 3 = 60$	36	24

Pro učení je ideální mít dostatek dat. Z toho důvodu jsem se rozhodla 4minutové úseky rozdělit dále na 10sekundové úseky. V tuto chvíli však nejsou třídy v datasetu rovnoměrně zastoupeny. Přestože diabetici podstoupili měření 2krát, dataset obsahuje více dat s nízkou glykémií. Právě z tohoto důvodu byly 4minutové signály u všech zdravých osob zkráceny na 2,5minutové úseky a poté opět rozděleny na 10 s. Schéma shrnující přípravu datasetu je zobrazeno na Obrázek 7.1.



Obrázek 7.1: Schéma přípravy datasetu (chytrý náramek)

Celkový počet 10s PPG záznamů nasnímaných chytrým náramkem a jejich zastoupení do tříd nízká/vysoká hodnota glykémie je v Tabulka 7.2. Takto vytvořený dataset bude následně využit pro trénování a testování.

Tabulka 7.2: Zastoupení počtu 10 s záznamů do skupin nízká/vysoká hodnota glykémie

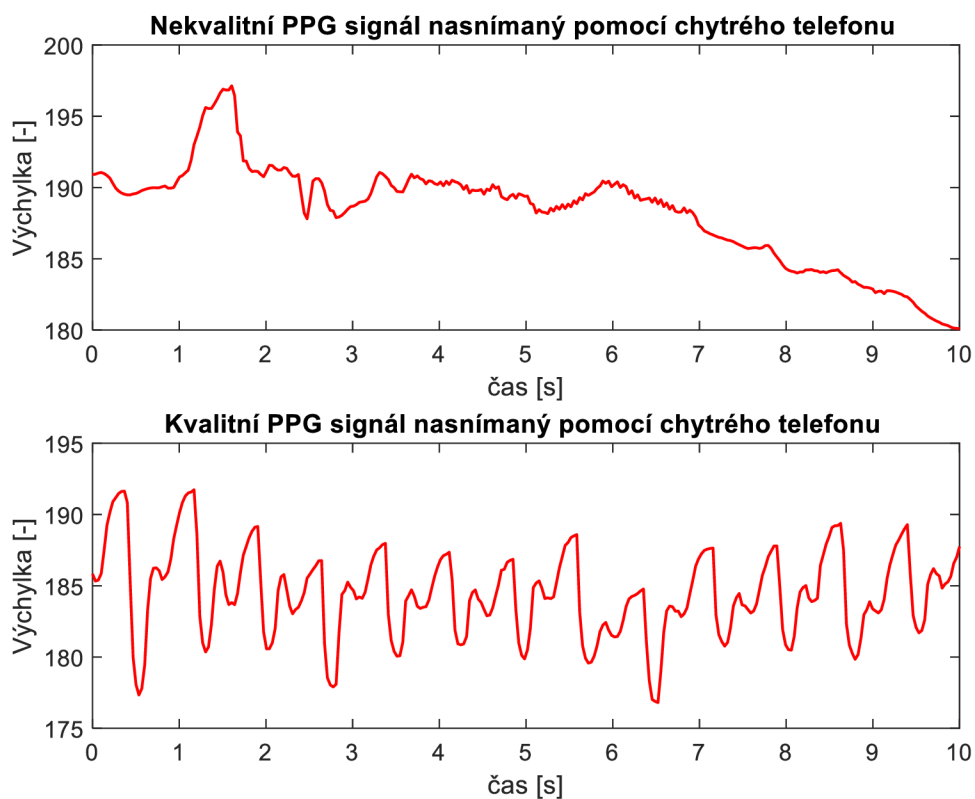
Počet 10 s záznamů	Nízká glykémie ($\leq 7,2$ mmol/l)	Vysoká glykémie ($>7,2$ mmol/l)
1143	537	606

Druhý dataset tvoří PPG signály nasnímané pomocí chytrého telefonu. Během jednoho měření byly nasnímané tři minutové záznamy. Tyto minutové úseky byly také rozděleny na 10sekundové úseky, jelikož budou následně použity k ověření funkčnosti algoritmů trénovaných na datech, které byly pořízeny s využitím chytrých hodinek. PPG signály nasnímané chytrým telefonem bývají často nekvalitní, což je dáno pohybem měřené osoby nebo pohyby prstu po čočce fotoaparátu. Proto je třeba určit kvalitu PPG signálů nasnímaných chytrým telefonem, o čemž pojednává následující kapitola 7.2.

7.2 Stanovení kvality PPG signálů nasnímaných chytrým telefonem

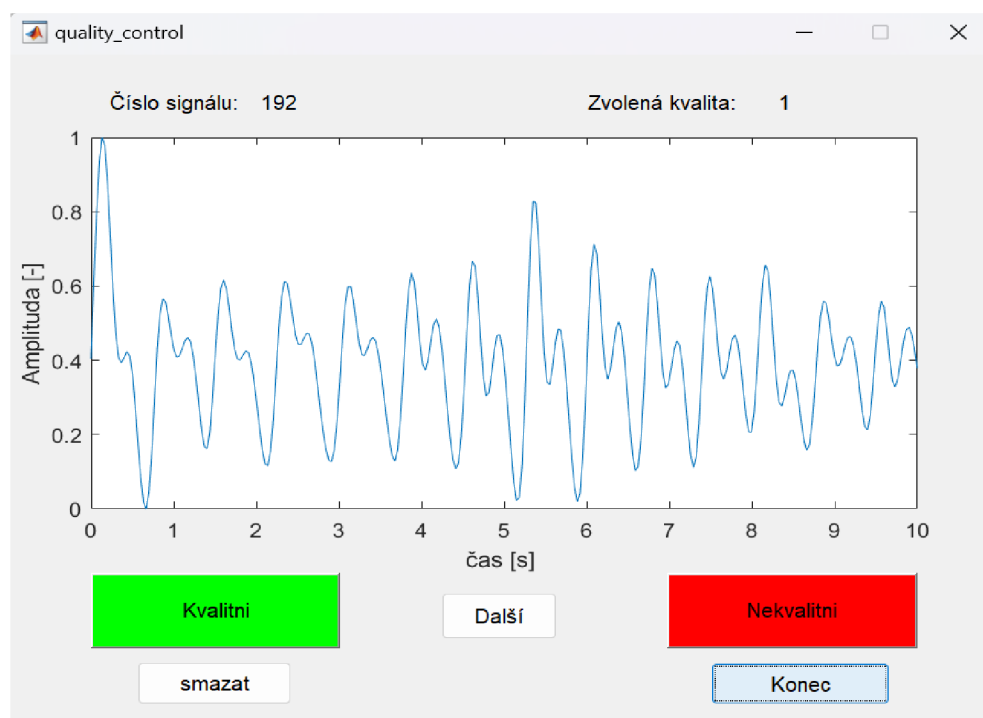
Důležitým krokem před zpracováním PPG signálů je stanovení jejich kvality. Právě signály nasnímané chytrým telefonem bývají často nekvalitní, což může být způsobeno nechtěnými pohyby prstu po čočce fotoaparátu, vyvíjením tlaku na čočku při držení chytrého telefonu apod. Takové signály je třeba z datasetu vyřadit, aby se předešlo zkreslení výsledků následující analýzy.

Pro stanovení kvality PPG signálů byl použit algoritmus z mé bakalářské práce [64], která se zabývala stanovením kvality PPG signálů a následným odhadnutím TF z kvalitních signálů. Algoritmus je založen na výpočtu dvou příznaků: perfuze a Shannonovy entropie, a rozřazení dat do skupin kvalitní/nekvalitní nehierarchickou metodou shlukování k-means. Algoritmus k-means shlukuje data do předem definovaného počtu skupin na základě podobnosti. Jelikož tento algoritmus dosahuje poměrně dobrých výsledků (přesnost na trénovací sadě: 97,5 %, přesnost na testovací sadě: 87,5 %), rozhodla jsem se jej použít i pro účely této DP. Do algoritmu vstupovalo celkem 324 PPG signálů nasnímaných chytrým telefonem o délce 10 s. Použitím algoritmu bylo vyhodnoceno 131 PPG signálů jako nekvalitních. Ukázka kvalitního a nekvalitního PPG signálu nasnímaného chytrým telefonem je na Obrázek 7.2.



Obrázek 7.2: Ukázka nekvalitního a kvalitního PPG signálu nasnímaného chytrým telefonem

Vizuální kontrola (Obrázek 7.3) zbylých 193 kvalitních PPG signálů ukázala, že v datasetu stále zůstávají některé signály, které jsou sice dostatečně kvalitní pro stanovení TF, ale nekvalitní pro odhad glykémie. Z toho důvodu bylo 193 algoritmem označených kvalitních signálů dodatečně oanotováno ručně.



Obrázek 7.3 : Ukázka GUI pro anotaci PPG signálů (invertované pro lepší vizualizaci)

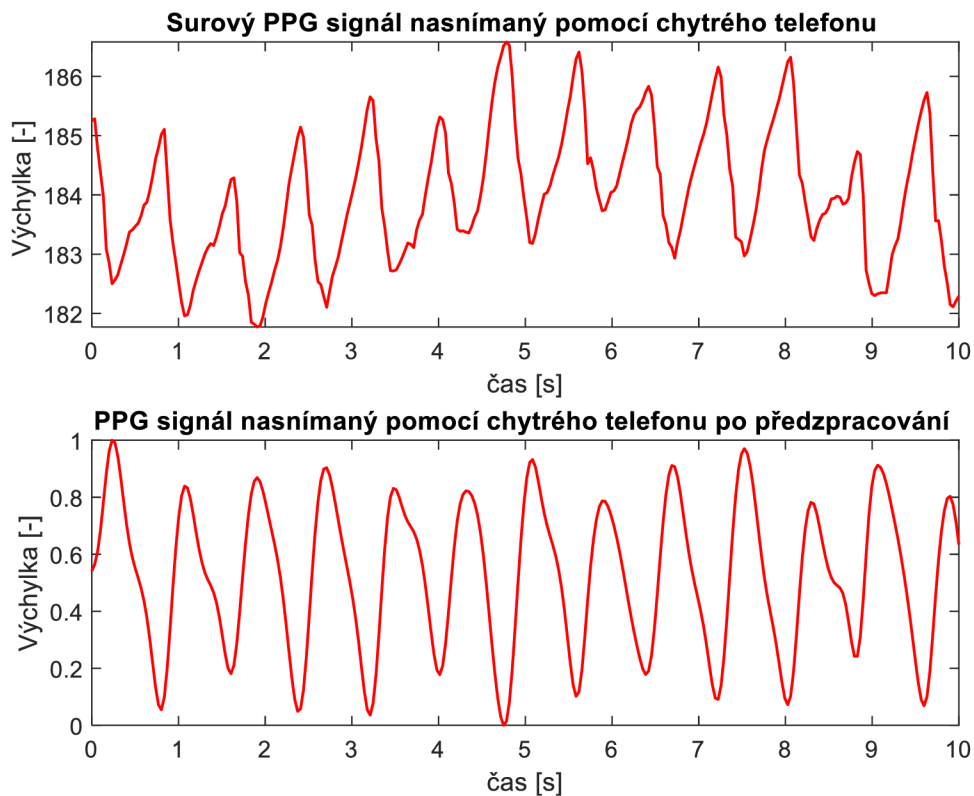
Ruční anotací bylo ze 193 PPG signálů označeno dalších 64 signálů jako nevhodných pro klasifikaci glykémie. Celkem tedy dataset nasnímaný chytrým telefonem obsahuje 129 záznamů. Zastoupení dat do jednotlivých tříd shrnuje Tabulka 7.3.

Tabulka 7.3: Zastoupení počtu 10 s záznamů do skupin nízká/vysoká hodnota glykémie (chytrý telefon)

Počet 10 s záznamů	Nízká glykémie ($\leq 7,2$ mmol/l)	Vysoká glykémie ($>7,2$ mmol/l)
129	51	78

7.3 Předzpracování PPG signálů

Veškeré PPG signály byly načtené do programového prostředí Matlab. PPG signály nasnímané pomocí chytrého telefonu je třeba nejdříve invertovat, jelikož se během snímání pracuje se světlem odraženým. Signály nasnímané pomocí chytrého náramku Empatica jsou již invertované. Následně byly všechny PPG signály filtrovány a normalizovány. Šum byl zredukován použitím Butterworthovy pásmové propusti 2. řádu v rozsahu 36 tepů/minutu až 216 tepů/minutu. Rozsah hodnot po max-min normalizaci je od 0 do 1 (Obrázek 7.4).



Obrázek 7.4: Ukázka PPG signálu č.8 před a po předzpracování

8. PŘÍZNAKY

V této kapitole se budu zaměřovat na příznaky, které mohou být vhodné pro stanovení glykémie z PPG signálů. Autoři článku [6] zabývající se predikcí glykémie využívají příznaky vycházející přímo z morfologie PPG vln. Naopak autoři [59] zabývající se klasifikací glykémie do dvou skupin využívají příznaků vycházejících z HRV analýzy. Pro účely této práce jsem se rozhodla využít příznaky vycházející z tvaru PPG křivky i z HRV analýzy. Celkem bylo vybráno 27 různých příznaků. Jejich souhrn lze vidět v Tabulka 8.1. Příznaky vycházející z HRV analýzy jsou podbarvené šedě a jejich souvislost s diabetem popisuje článek [46], kterému se věnuje kapitola 3.6.

Dalším krokem po výpočtu příznaků je odstranění odlehlých hodnot (kapitola 8.1). Následně byla provedena statistická analýza parametrů (kapitola 8.2), jejímž úkolem bylo odhalit příznaky, které nejsou schopné dobře odlišit data do dvou klasifikačních tříd. Poté byla dodatečně provedena selekce vhodných příznaků pomocí mRMR algoritmu (Maximum Redundancy Minimum Relevance) popsána v kapitole 8.3. Příznaky vybrané selekcí jsou v Tabulka 8.1 zvýrazněny tučně a blíže popsány v kapitole 8.4. Pro efektivní učení je také velice důležité příznaky standardizovat. Pro tuto práci byla zvolena standardizace Z-score. Z důvodu použití křížové validace je proces standardizace příznaků popsán až v kapitole 9.1.

8.1 Odlehlé hodnoty

Po výpočtu příznaků byly nalezeny a nahrazeny odlehlé hodnoty, které mohly vzniknout při měření signálů. Odlehlé hodnoty mohou způsobovat komplikace během učení modelu. Za odlehlou hodnotu je považována hodnota, která je vzdálená od horního/dolního kvartilu o více než 1,5násobek mezikvartilového rozpětí. V případě, že je daná hodnota nižší než dolní kvartil, je nahrazena nejnižší hodnotou povoleného rozmezí. V případě, že je hodnota vyšší než horní kvartil, je nahrazena nejvyšší hodnotou povoleného rozmezí.

Tabulka 8.1: Popis příznaků

	Příznaky	Popis příznaků
1	Špičatost	udává do jaké míry jsou hodnoty koncentrovány kolem střední hodnoty
2	Entropie	udává míru s jakou se rozdělení pravděpodobnosti odchyluje od rovnoměrného rozdělení
3	Šikmost	popisuje rozdělení souboru dat kolem střední hodnoty
4	P2rms	poměr největší absolutní hodnoty signálu k RMS
5	RMSSD	odmocnina z průměrného rozdílu každých dvou po sobě následujících RR (PP) umocněných na druhou
6	Směrodatná odchylka	směrodatná odchylka celého PPG signálu
7	Rozptyl	rozptyl celého PPG signálu
8	IQR	mezikvartilové rozpětí ukazuje rozdíl mezi 75. kvantilem a 25. kvantilem
9	Poměr TKEO	poměr minimální a maximální hodnoty obálky získané pomocí Teagerova-Kaiserova energetického operátoru (TKEO)
10	Počet peaků	počet peaků detekovaných v daném signálu
11	Průměrná výška PPG peaků	odpovídá průměrné výšce všech detekovaných peaků v PPG signálu
12	Medián výšky PPG peaků	odpovídá mediánu výšky všech detekovaných peaků v PPG signálu
13	Std výšky PPG peaků	odpovídá směrodatné odchylce výšky PPG peaků
14	Rozdíl maximální a minimální výšky PPG	jedná se o rozdíl maximální a minimální výšky peaků v PPG signálu
15	Nejkratší PP interval	odpovídá nejkratšímu PP intervalu ze všech PP intervalů v PPG signálu
16	Nejdelší PP interval	odpovídá nejdelšímu PP intervalu ze všech PP intervalů v PPG signálu
17	Std 1. derivace PPG	jde o směrodatnou odchylku 1. derivace PPG signálu
18	Průměrná vzdálenost PP intervalu	průměrná vzdálenost PP intervalů z PPG signálu
19	CV	také variační koeficient, což je směrodatná odchylka souboru RR (PP) intervalů podělená průměrem tohoto souboru intervalů
20	SDNN	směrodatná odchylka délky NN intervalů (PP intervalů)
21	NN50	počet NN intervalů, které se oproti následujícím liší o více než 50 ms
22	pNN50	počet NN50 vydělený celkovým počtem NN intervalů v záznamu
23	Poměr SD1/SD2	poměr krátkodobé variability PP intervalů SD1 a dlouhodobé variability PP intervalů SD2
24	Poměr minima a maxima 1. derivace PPG	jedná se o poměr minima a maxima 1. derivace PPG
25	Průměr 1. derivace PPG signálu	jde o průměr 1. derivace PPG signálu
26	Poměr minima a maxima 2. derivace PPG	odpovídá poměru minima a maxima 2. derivace PPG
27	Průměr 2. derivace PPG signálu	odpovídá průměru 2. derivace PPG signálu

8.2 Statistická analýza příznaků

Statistická analýza příznaků má odhalit příznaky, které nejsou vhodné pro klasifikaci záznamů do tříd nízká/vysoká glykémie. Nejdříve byl aplikován Kolmogorův-Smirnovův test na hladině významnosti $\alpha = 0,05$. Nulová hypotéza říká, že data mají normální rozložení. Naopak alternativní hypotéza říká, že data nepochází z normálního rozložení. V případě, že je nulová hypotéza zamítnuta (p -hodnota $< 0,05$), data nepochází z normálního rozložení. Test ukázal, že žádný příznak nemá normální rozložení. [65]

Následně byl na data aplikován Mann-Whitneyův neparametrický test, který se používá v případě, že data nemají normální rozložení. Pokud by některý z příznaků měl normální rozložení byl by použit nepárový (dvouvýběrový) t -test. Test by měly ukázat, jak moc se liší hodnoty každého příznaku v obou třídách. Hladina významnosti byla opět nastavena na $\alpha = 0,05$. Nulová hypotéza říká, že se data pro daný příznak v obou třídách neliší. V případě, že je nulová hypotéza zamítnuta, znamená to, že se hodnoty pro daný příznak v obou skupinách liší. V takovém případě je daný příznak vybrán jako vhodný pro klasifikaci. Tímto způsobem byly vyřazeny 4 příznaky - příznak číslo 16 (nejdelší PP interval), 25 (Průměr 1. derivace PPG signálu), 26 (Poměr minima a maxima 2. derivace PPG), 27 (Průměr 2. derivace PPG). [66]

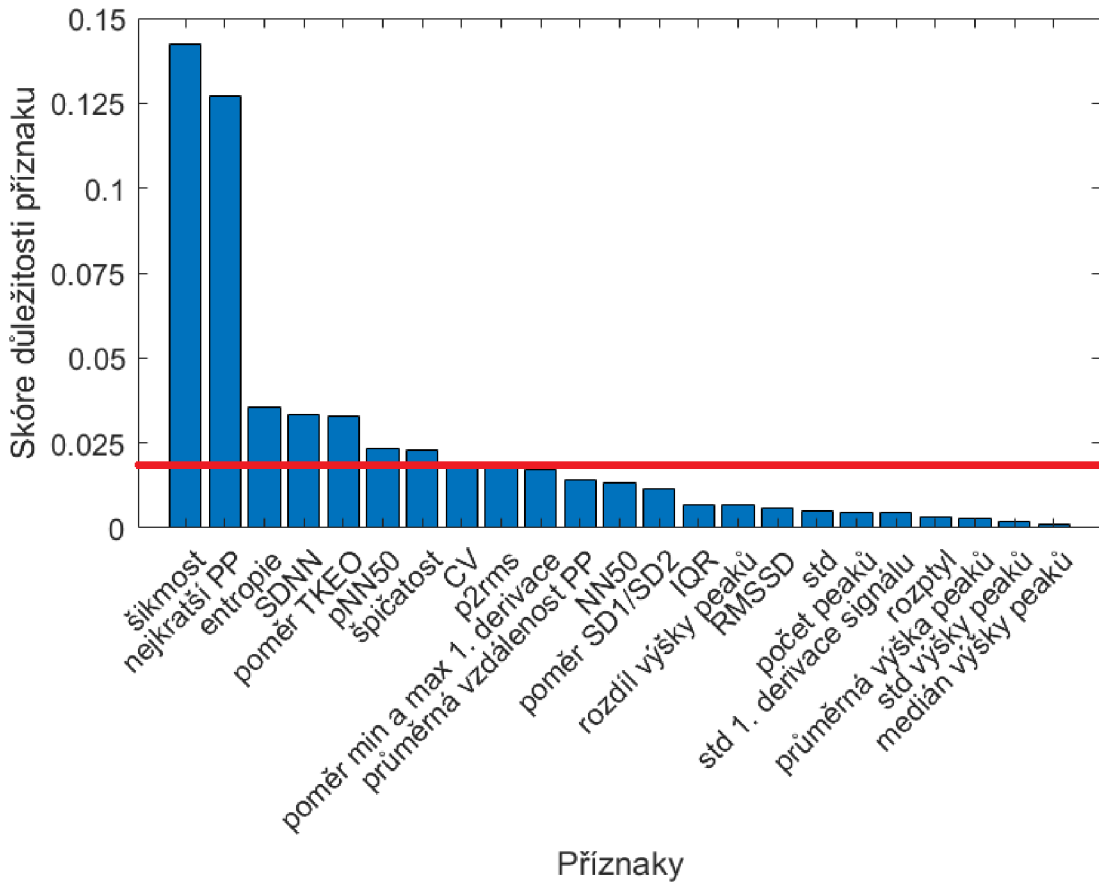
Zde je dobré zmínit, že analýza a následný výběr příznaků (kapitola 8.3) je prováděna pouze na datech získaných z chytrého náramku. Stejný proces a stejné příznaky pak budou také využity pro ověření algoritmu na datech z chytrého telefonu.

8.3 Výběr příznaků

Se zvyšující se dimenzionalitou dat roste exponenciálně množství dat potřebných ke správné analýze. Navíc se zvyšující se dimenzionalitou dat roste také výpočetní náročnost algoritmu. Z výše zmíněných důvodů je vhodné provést před učením redukcí dat. Existují dva základní způsoby: selekce příznaků a extrakce příznaků. Pod pojmem selekce příznaků rozumíme výběr příznaků z dané množiny příznaků. Pojem extrakce příznaků znamená snížení počtu příznaků kombinací všech příznaků. V případě extrakce příznaků nelze následně jednotlivé příznaky interpretovat. Právě z tohoto důvodu je pro účel DP zvolena jedna z metod selekce příznaků – minimální redundance maximální relevance (mRMR). [67]

Cílem metody mRMR je najít parametry, které charakterizují nejlépe jednotlivé třídy. Metoda nejdříve nahlíží na parametry ve smyslu maximální relevance, tzn. nejprve jsou vybrány parametry, které jsou nejvýznamnější pro danou úlohu. Následně se sleduje redundance vůči již vybraným parametrům, která by měla být minimální. Relevanci i redundanci lze popsat pomocí tzv. vzájemné informace. Vzájemná informace je množství informace, které má jedna náhodná veličina o jiné náhodné veličině. Celý algoritmus tedy funguje iterativně. Nejdříve je vybrán příznak

s nejvyšší vzájemnou informací ke třídám. Dále jsou příznaky vybírány postupně na základě minimální redundance s již vybraným příznakem/příznaky. [68]



Obrázek 8.1: Skóre důležitosti příznaků

Obrázek 8.1 ukazuje grafické znázornění skóre důležitosti příznaku pro jednotlivé příznaky. Na základě tohoto grafu byly vybrány příznaky, jejichž skóre důležitosti je vyšší, než heuristicky zvolený práh 0,02 (červená linie na Obrázek 8.1). Tímto způsobem bylo vybráno 10 příznaků, které jsou následně použity pro učení. Popis vybraných příznaků je v následující kapitole 8.4.

8.4 Popis vybraných příznaků

V této kapitole jsou popsány příznaky, které byly vybrány jako vhodné pro klasifikaci hodnot glykémie do dvou tříd a následnou predikci konkrétní hodnoty glykémie. Každý příznak je popsán a následně je pro každý příznak na ukázkou zobrazen krabicový graf, který slouží k vizualizaci rozložení hodnot daného příznaku ještě před standardizací. Krabicový graf ukazuje nejen minimální a maximální hodnotu v souboru daného příznaku, ale také medián, horní a dolní kvartil. Vodorovná červená čára v krabicovém grafu odpovídá mediánu, horní a dolní strana obdélníku poté hornímu a dolnímu

kvartilu. Svislé úsečky, též dolní a horní fous, odpovídají hodnotám, které jsou od daných kvartilů vzdálené méně než o 1,5násobek výšky krabice. Krabicové grafy demonstrují, jak moc se daný příznak liší u záznamů s nízkou a vysokou hodnotou glykémie. [69]

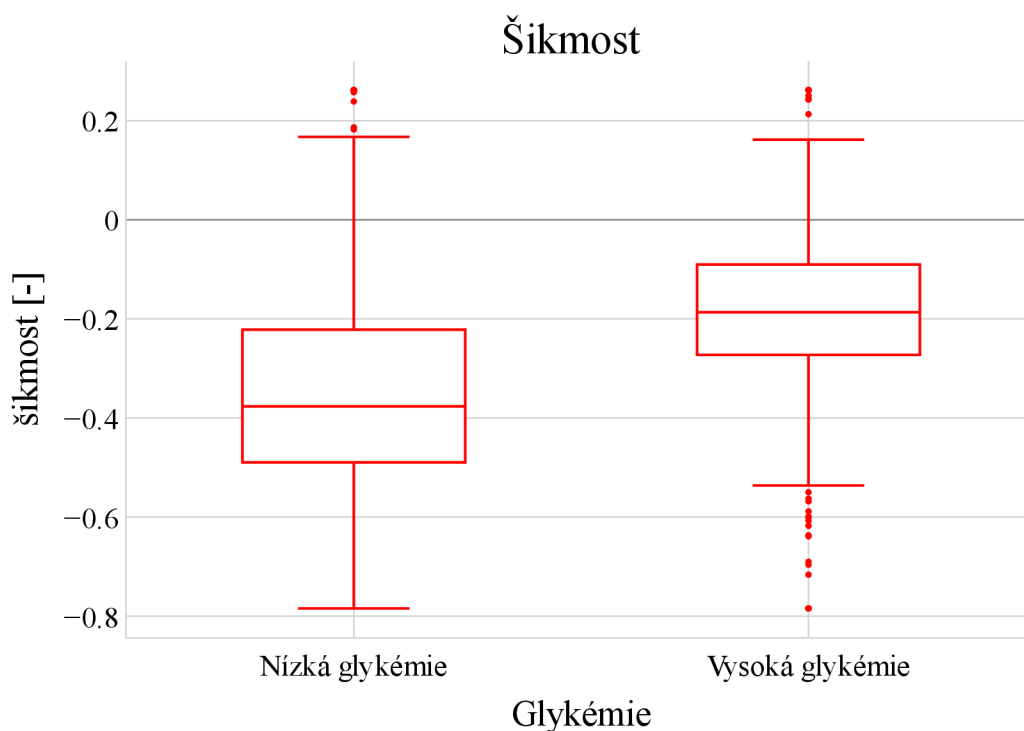
Šikmost

Šikmost popisuje rozdělení souboru dat kolem střední hodnoty a říká nám, zda jsou data rozdělena symetricky či nikoliv. Kladná šikmost značí odlehlejší hodnoty směrem vpravo od střední hodnoty. Šikmost může být potenciálně použita ke zlepšení diagnostiky a monitorování různých abnormalit. Šikmost lze vypočítat použitím vzorce:

$$\text{Šikmost} = 1/N \sum_{i=1}^N [x_i - \hat{\mu}_x / \sigma]^3, \quad (8.1)$$

kde $\hat{\mu}_x$ a σ jsou empirické odhady průměru a směrodatné odchylky proměnné x a N je počet vzorků PPG signálu. [70]

Krabicový graf pro šikmost je na Obrázek 8.2.



Obrázek 8.2: Krabicový graf šikmost

Špičatost

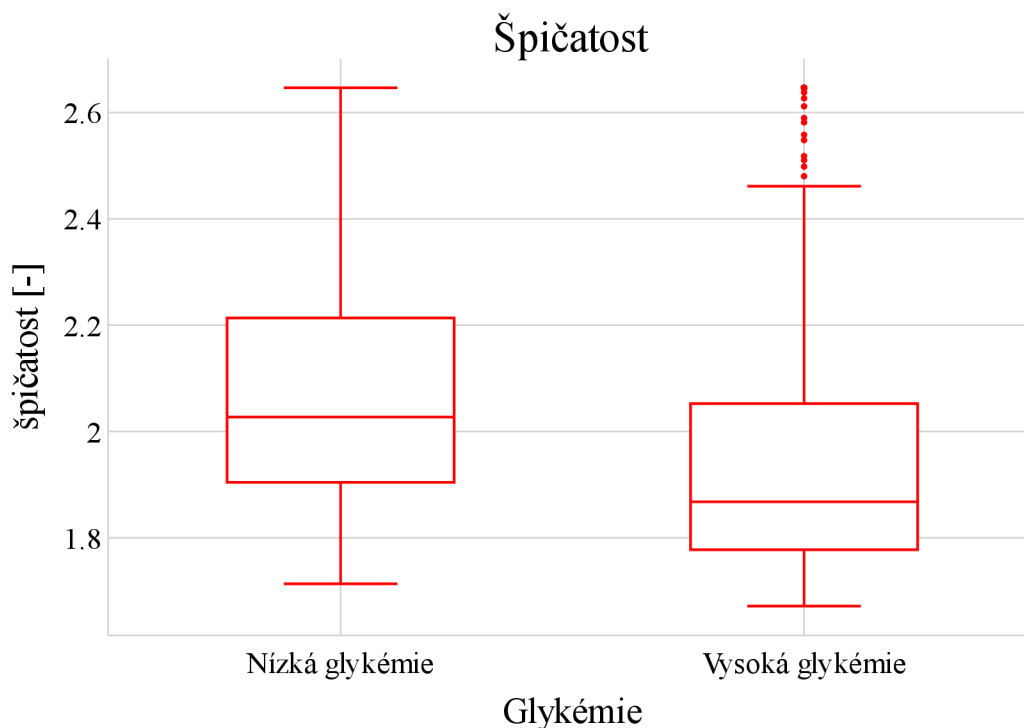
Špičatost říká, do jaké míry jsou hodnoty koncentrovány kolem střední hodnoty. Kladná špičatost odpovídá situaci, kde je většina hodnot koncentrována kolem střední hodnoty. Koeficient špičatosti je indikátorem odlehlých hodnot v PPG signálu, které mohou odpovídat abnormalitám souvisejícím například s vysokou hodnotou glykémie.

Špičatost lze získat dosazením do rovnice:

$$\text{Špičatost} = 1/N \sum_{i=1}^N [x_i - \hat{\mu}_x / \sigma]^4, \quad (8.2)$$

kde $\hat{\mu}_x$ a σ jsou empirické odhady průměru a směrodatné odchylky proměnné x a N je počet vzorků PPG signálu. [70][71]

Krabicový graf pro špičatost je na Obrázek 8.3.



Obrázek 8.3: Krabicový graf špičatost

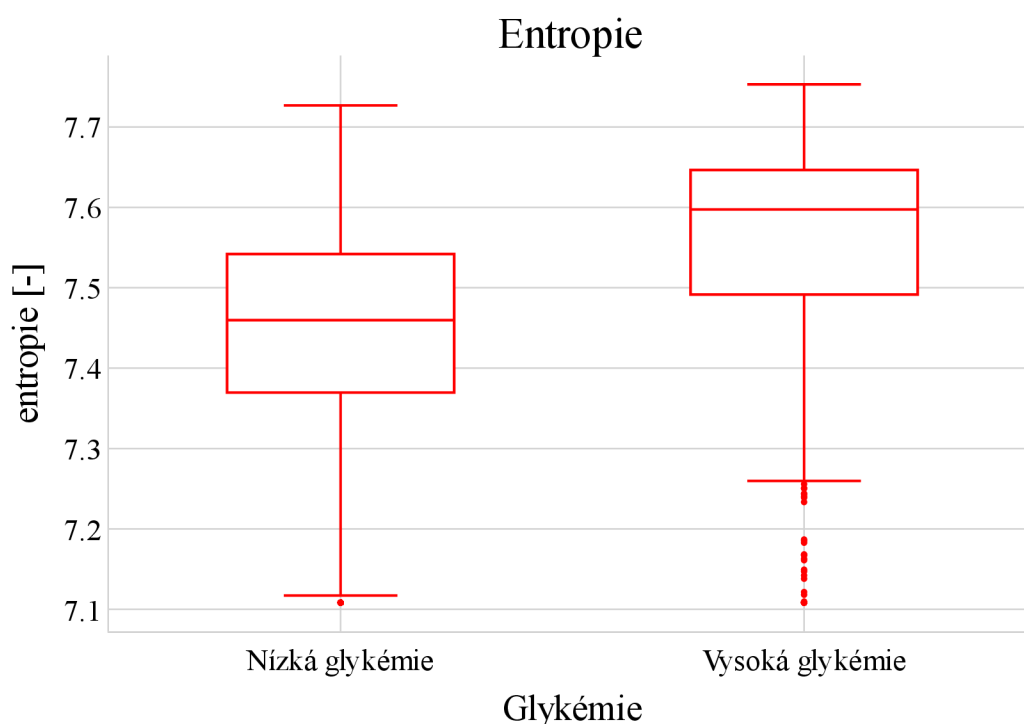
Shannonova entropie

Entropie je průměrné množství informace na jednu hodnotu, množství informace je nezáporné a roste s klesající pravděpodobností výskytu. Pokud nabývá vysokých hodnot, lze říci, že soubor dat je víceméně náhodný a neobsahuje žádné pravidelné vzory. Dosazením do vzorce (8.3) získáme entropii. [70]

$$SE = - \sum_{i=1}^N x[n]^2 \log_e(x[n]^2), \quad (8.3)$$

kde x je PPG signál a N odpovídá jeho délce.

Krabicový graf pro Shannonovu entropii je na Obrázek 8.4.



Obrázek 8.4: Krabicový graf entropie

SDNN

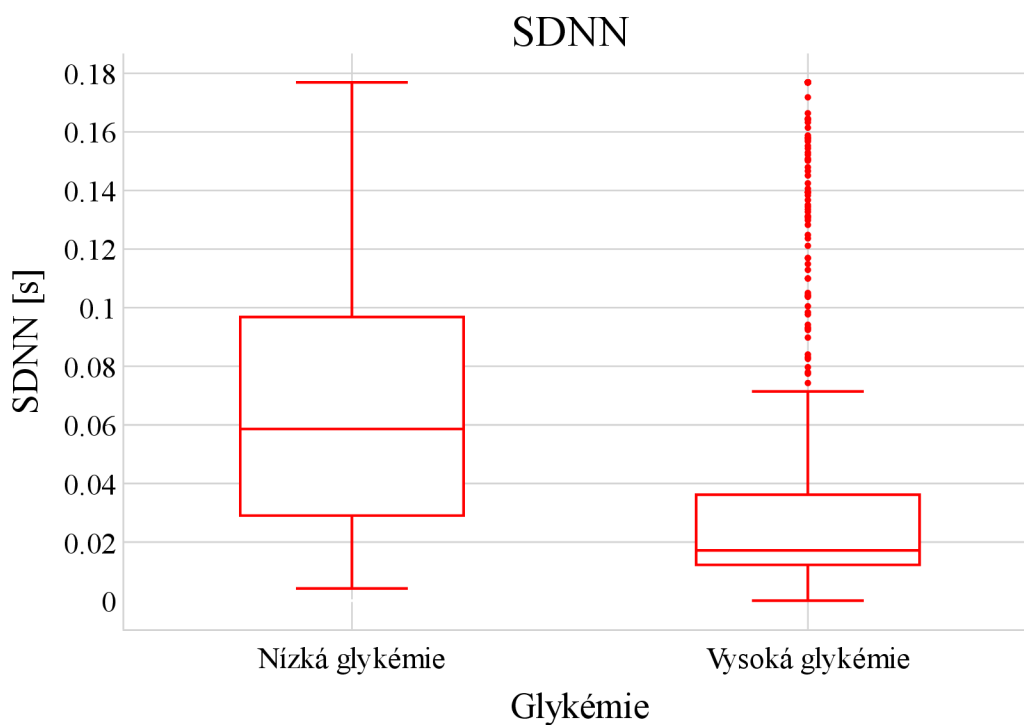
Jedná se o směrodatnou odchylku NN (PP) intervalů. Jelikož je tato veličina závislá na délce PP intervalů, její jednotkou jsou ms. Dle různých článků zabývajících se HRV analýzou by měli mít diabetici nižší hodnotu SDNN, právě v důsledku snížení srdeční variability způsobené sníženou parasympatickou aktivitou. [46][72][73][74].

Výpočet SDNN:

$$SDNN = \sqrt{\frac{1}{M-1} \sum_{i=1}^M (NN_i - \mu_{NN})^2}, \quad (8.4)$$

kde NN_i u PPG signálů odpovídá i -tému PP intervalu, μ_{NN} je průměrná hodnota NN (PP) intervalů a M je celkový počet intervalů.

Krabicový graf pro SDNN je na Obrázek 8.5.



Obrázek 8.5: Krabicový graf SDNN

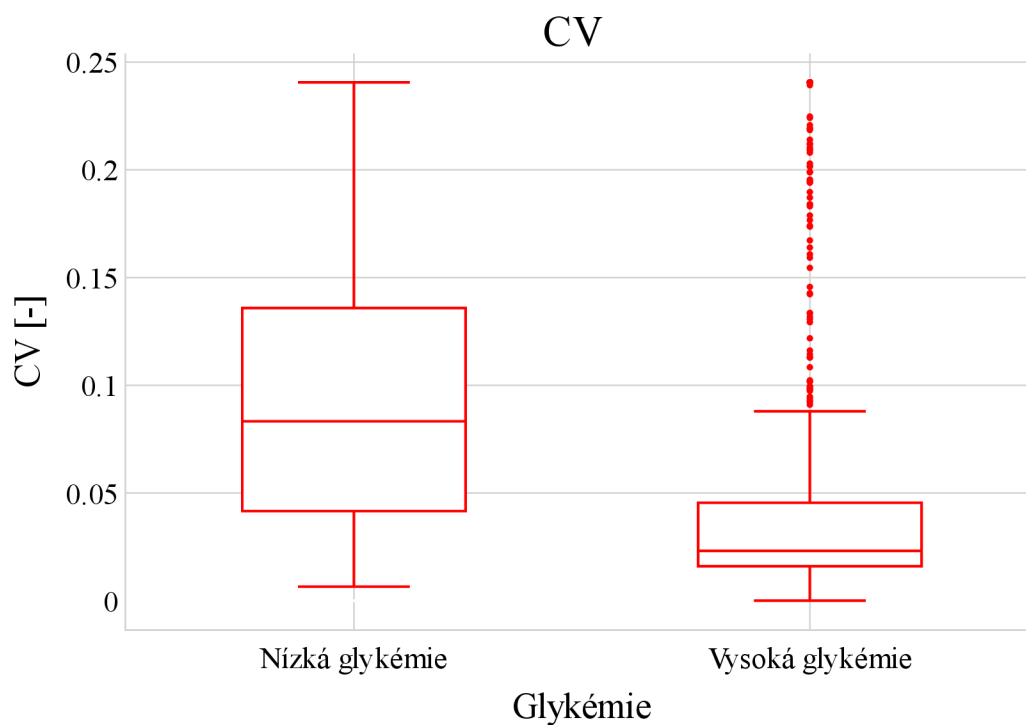
Variační koeficient (CV)

Variační koeficient lze spočítat dle rovnice (8.5).

$$CV = \frac{\sigma_{RR}}{\mu_{RR}}, \quad (8.5)$$

kde σ_{RR} je směrodatná odchylka RR (PP) intervalů a μ_{RR} je průměr tohoto souboru. Opět na základě článků platí, že parametr CV dosahuje nižších hodnot u diabetiků a osob s vyšší hodnotou glykémie.[46][72][73][74]

Krabicový graf pro CV je na Obrázek 8.6.

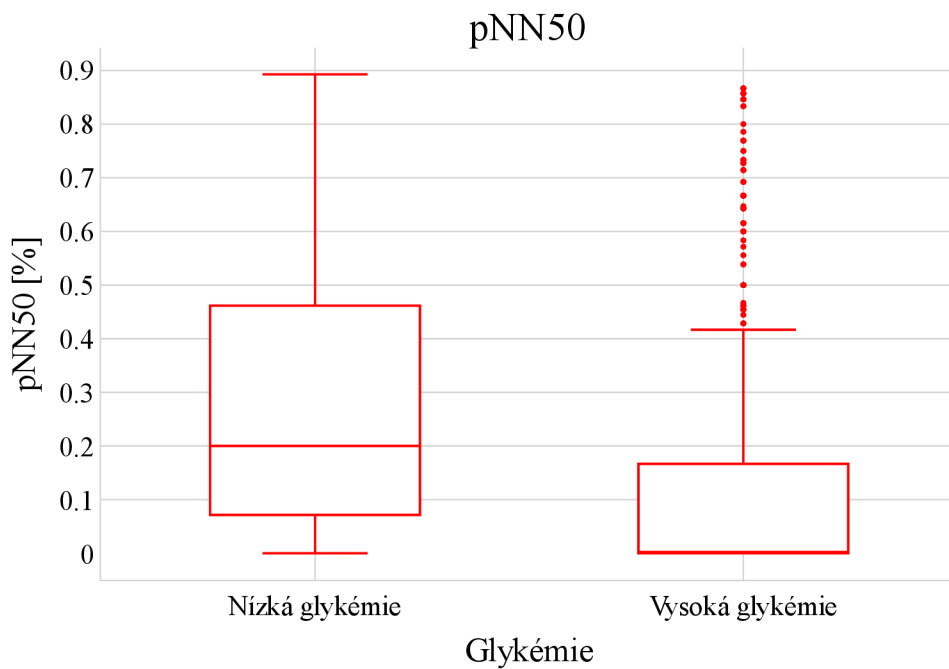


Obrázek 8.6: Krabicový graf CV

pNN50

Tento příznak vychází z parametru NN50, který odpovídá počtu párů NN (PP) intervalů, které se oproti následujícím liší o více než 50 ms. Parametr pNN50 je potom počet NN50 vydělený celkovým počtem NN (PP) intervalů v záznamu. Opět by tato hodnota měla být nižší u lidí s diabetem a vyšší hodnotou glykémie. [46][72][73][74]

Krabicový graf pro pNN50 je na Obrázek 8.7.

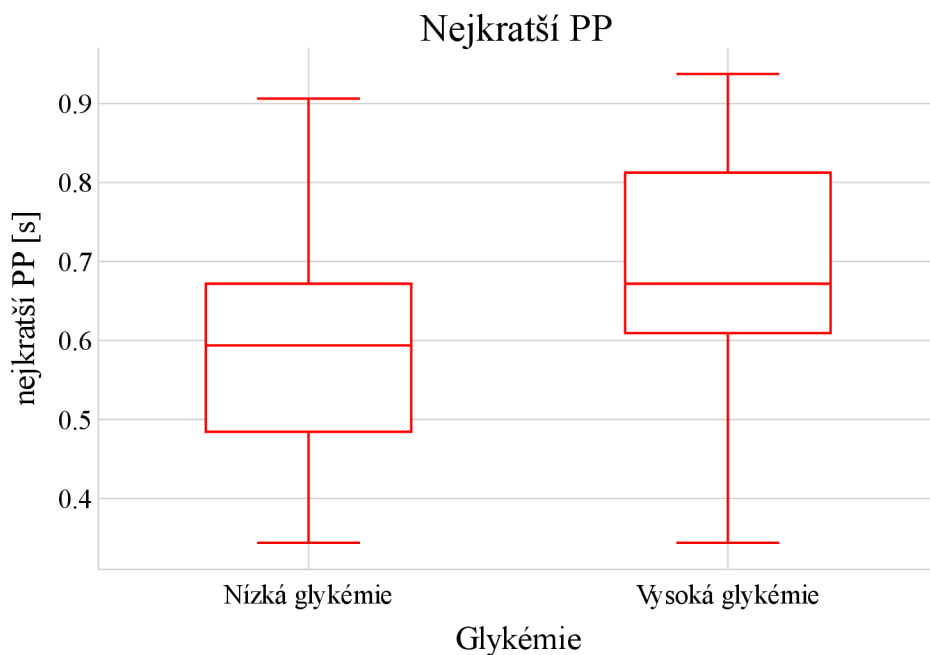


Obrázek 8.7: Krabicový graf pNN50

Nejkratší PP interval

Tento příznak odpovídá nejkratšímu PP intervalu ze všech PP intervalů v jednom PPG signálu. Tento příznak na základě krabicového grafu nabývá vyšších hodnot u diabetiků a osob se zvýšenou glykemií. Některé studie dokonce ukazují, že RR (PP) interval může být prodloužený v přítomnosti diabetu. [75]

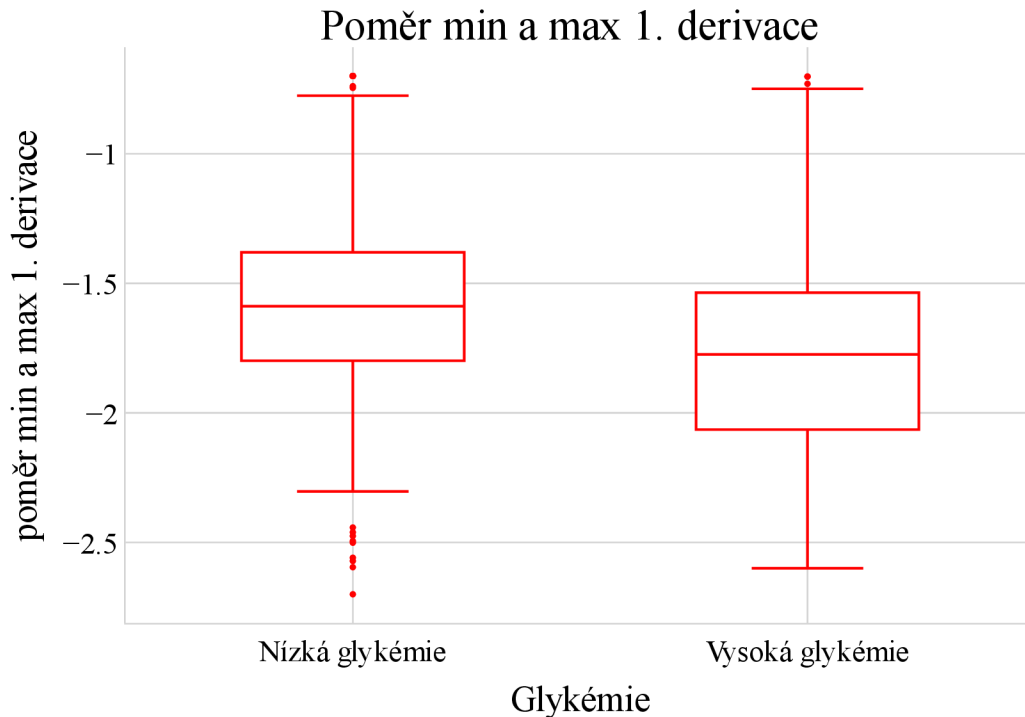
Krabicový graf nejkratšího PP intervalu je na Obrázek 8.8.



Obrázek 8.8: Krabicový graf minimální vzdálenosti PP intervalů

Poměr minima a maxima 1. derivace signálu

Příznak vyjadřuje poměr minimální a maximální hodnoty 1. derivace daného PPG signálu (Obrázek 8.9.).



Obrázek 8.9: Krabicový graf poměru minima a maxima 1. derivace PPG signálu

p2rms

Poměr největší absolutní hodnoty signálu k odmocnině průměru čtverců ze signálu. Krabicový graf znázorňující p2rms je na Obrázek 8.10.

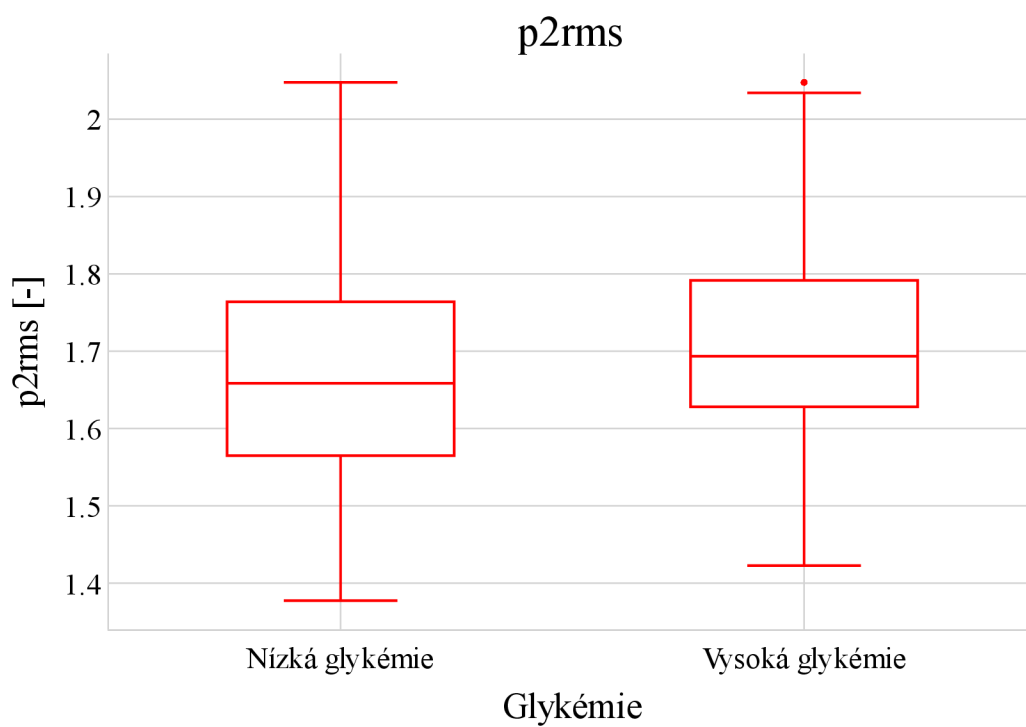
Poměr TKEO

TKEO odpovídá odhadu okamžité energie signálu. Pro účely DP se počítá poměr minimální a maximální hodnoty obálky získané pomocí TKEO.

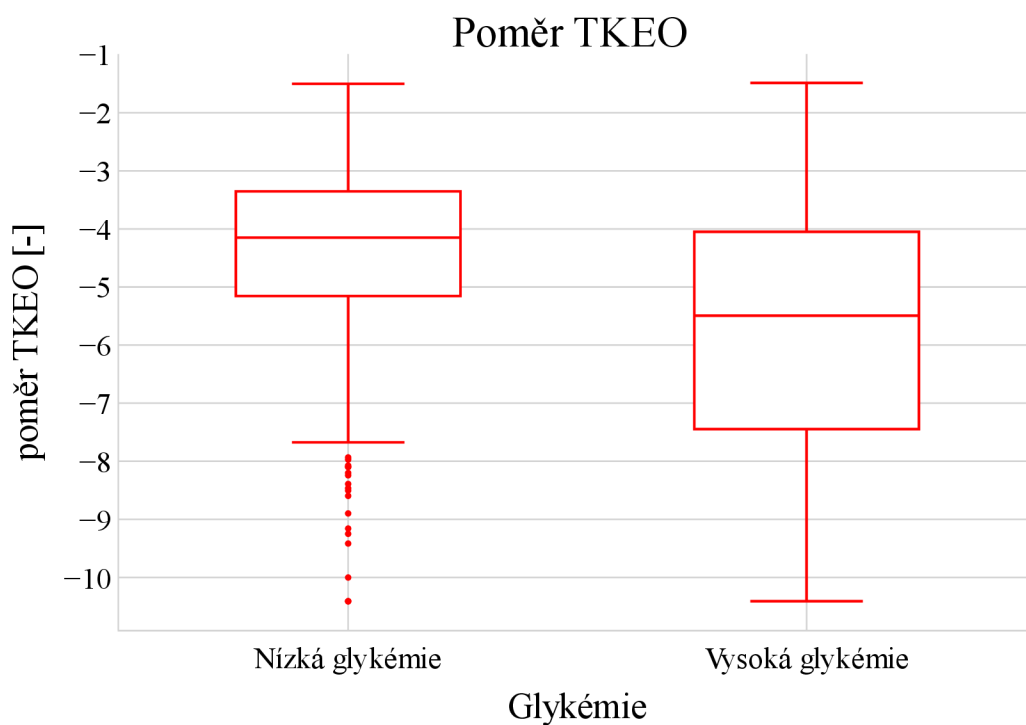
$$TKEO[x(n)] = x^2(n) - x(n+1)x(n-1), \quad (8.6)$$

kde n značí počet vzorků a x je PPG signál.

Krabicový graf poměru maxima a minima obálky TKEO lze vidět na Obrázek 8.11.

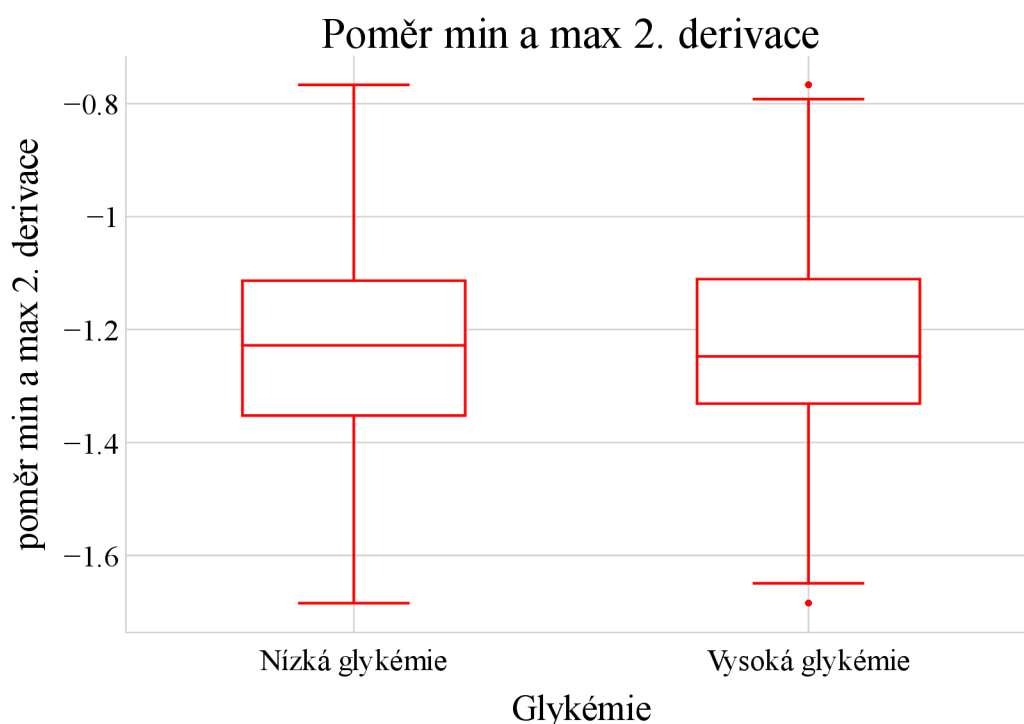


Obrázek 8.10: Krabicový graf p2rms



Obrázek 8.11: Krabicový graf poměru minima a maxima TKEO obálky

Na základě boxplotů se může zdát, že některé příznaky nabývají podobných hodnot v obou skupinách (nízká / vysoká glykémie). U všech výše zmíněných příznaků však byla na základě statistických testů zamítnuta nulová hypotéza na hladině významnosti 0,05 a dané příznaky se tak v obou třídách liší. Na následujícím krabicovém grafu (Obrázek 8.12) lze vidět příklad příznaku, u nějž nulová hypotéza zamítnuta nebyla a nebyl tak tento příznak zařazen do užšího výběru příznaků, jak je zmíněno i v kapitole 8.2.



Obrázek 8.12: Krabicový graf poměru minima a maxima 2. derivace PPG signálu

V dosavadních kapitolách byla popsána analýza a selekce příznaků prováděna na datech z chytrého náramku. Vybraných 10 příznaků bylo následně spočítáno také pro data z chytrého telefonu. Nahrazení odlehlých hodnot a následná standardizace (popsána v kapitole 9.1) byla provedena stejně. Další kapitola se již věnuje rozdělení dat na trénovací a testovací sadu a tvorbě modelu.

9. MODELÝ PRO KLASIFIKACI A PREDIKCI

Tato kapitola se nejdříve věnuje rozdělení dat z chytrého náramku na trénovací a testovací sadu a křížové validaci. Následující podkapitoly 9.2 a 9.3 se věnují tvorbě modelů pro klasifikační i predikční úlohu v programovém prostředí Python. Funkčnost modelů byla následně ověřena na sadě dat nasnímaných chytrým telefonem. Dosažené výsledky jsou vyhodnoceny a diskutovány v kapitole 10.

9.1 Rozdělení a standardizace dat

Pro tvorbu modelů se tedy v tuto chvíli pracuje pouze s daty z chytrého náramku. Data byla rozdělena na trénovací a testovací sadu. Trénovací sada dat je tvořena cca 75 % veškerých dat nasnímaných chytrým náramkem. Zbýlých 25 % dat je využito pro testování. Trénovací sada obsahuje data od 12 osob. Celkově tvoří trénovací sadu 15 měření, z toho 7 záznamů je od osob s diabetem / prediabetem (Tabulka 9.1).

Tabulka 9.1: Trénovací sada (chytrý náramek)

Trénovací sada (chytrý náramek)		
Celkový počet 10 s záznamů	864	
Celkový počet měření	15	
	Zdraví	Diabetes / prediabetes
Počet měření	8	7
Počet osob	8	4

Testovací sada obsahuje celkem 5 měření od 3 zdravých osob a 1 osoby s diabetem, přičemž diabetici byli měřeni 2krát (Tabulka 9.2).

Tabulka 9.2: Testovací sada (chytrý náramek)

Testovací sada (chytrý náramek)		
Celkový počet 10 s záznamů	279	
Celkový počet měření	5	
	Zdraví	Diabetes / prediabetes
Počet měření	3	2
Počet osob	3	1

Aby byla zaručena generalizační schopnost jednotlivých modelů, byla využita křížová validace. Proces rozdělení dat na trénovací a testovací sadu, tak jak je popsán výše, byl proveden tedy 4krát. Tímto způsobem byly vytvořeny 4 naprosto odlišné testovací sady, obsahující měření od různých osob. Pro všechna rozdělení dat

je zajištěno, aby se data od stejné osoby neobjevila v trénovací a testovací sadě zároveň. Výsledky z jednotlivých iterací křížové validace se nakonec průměrují. Proces křížové validace je použit pro nalezení optimálních parametrů modelu a také pro věrohodnější vyhodnocení výkonnosti modelu.

Jak bylo popsáno v kapitole 8.1, odlehlé hodnoty, které by mohly negativně ovlivnit proces učení, byly již nahrazeny. Pro efektivnější učení je vhodné převést příznaky s odlišnou škálou hodnot na příznaky standardizované. V DP byla vybrána standardizace příznaků pomocí z-skóre. Tato transformace je lineární, škálu posunuje, ale mění pouze měřítko příznaků, nikoliv vzdálenosti mezi jednotlivými hodnotami. Z - skóre lze spočítat dle rovnice:

$$Z = \frac{X - \mu}{\sigma}, \quad (9.1)$$

kde μ odpovídá aritmetickému průměru a σ vyjadřuje směrodatnou odchylku. [76]

Standardizace byla provedena v rámci křížové validace. Pro každou trénovací sadu dat byl tedy spočítán průměr a směrodatná odchylka. Následně byly tyto hodnoty použity pro standardizaci příslušné testovací sady dat.

9.2 Klasifikační modely

V následující kapitole jsou stručně popsány všechny klasifikační algoritmy, které byly pro DP použity. Úkolem těchto modelů je rozřadit vstupní data na základě příznaků do dvou skupin – nízká / vysoká hodnota glykémie. Byly využity dvě sady příznaků, jedna sada obsahuje pouze 10 příznaků vybraných metodou mRMR. Jelikož byl práh pro tuto metodu stanoven heuristicky byla také vyzkoušena druhá sada příznaků, která obsahuje všech 23 příznaků vybraných pomocí statistických testů. Pro některé modely je při procesu učení zásadní nastavit některé parametry. Popis jednotlivých hyperparametrů modelů a nalezení jejich optimální hodnoty je v jednotlivých kapitolách diskutován.

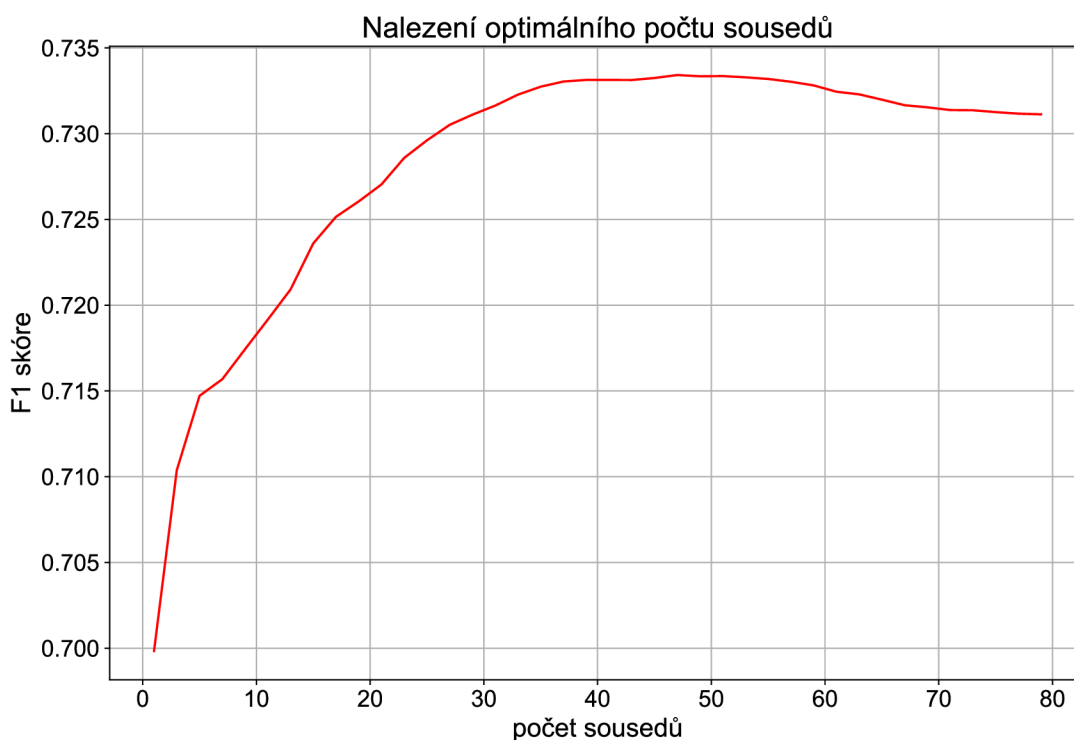
9.2.1 K-nejbližších sousedů

Metoda k-nejbližších sousedů (k-nearest neighbor, dále jen KNN) je založena na podobnosti sledovaného objektu s jeho nejbližšími sousedy. Podobnost, tedy vzdálenost k jednotlivým objektům bývá nejčastěji vyjádřena pomocí Euklidovské vzdálenosti.

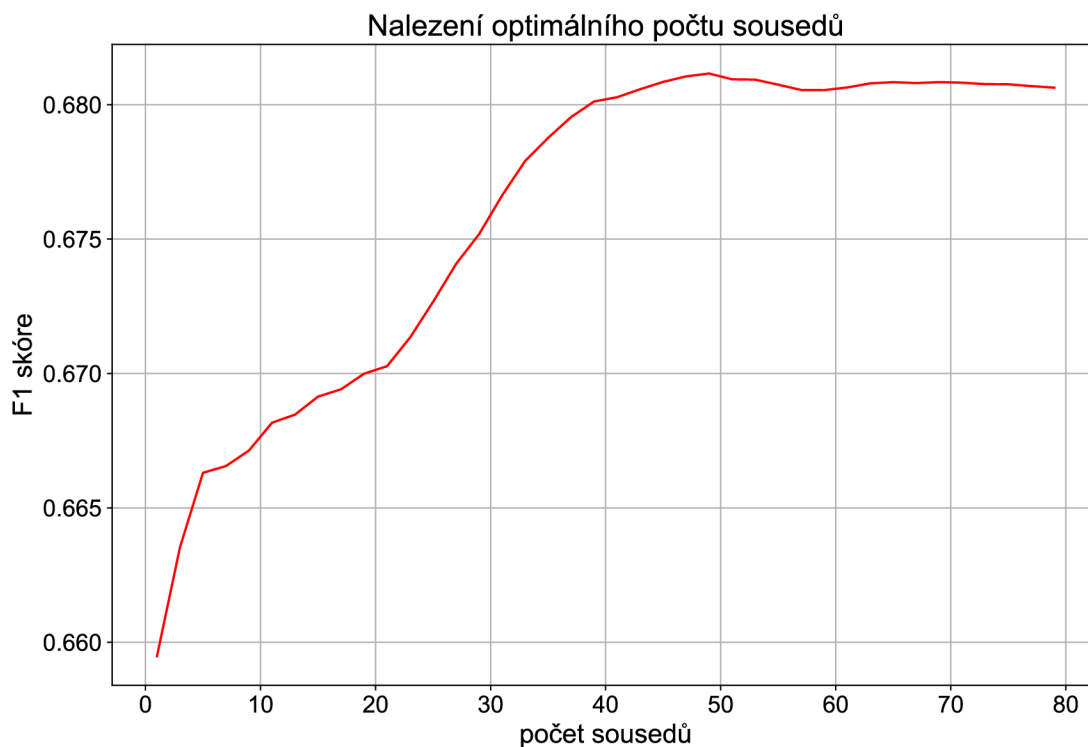
KNN algoritmus lze použít pro klasifikační i regresní úlohu. V případě, že je počet sousedů roven 1 ($k = 1$), bude testovací objekt zařazen do stejné třídy jako nejbližší objekt z trénovací množiny dat. V tomto případě je však metoda náchylná k šumu. Proto se obvykle volí $k > 2$, kde algoritmus zjistí příslušné třídy všech nejbližších sousedů a následně zařadí sledovaný objekt do třídy, která je nejvíce zastoupena.

Je vhodné volit lichý počet sousedů pro vyřešení situací s nerozhodným hlasováním. Algoritmus lze aplikovat také na regresní úlohu, kde výstupní hodnota testovaného objektu bude spočítána jako průměr z nejbližších sousedů. [77]

Metoda KNN je jedna z metod, která byla použita pro klasifikační úlohu. Jelikož je rozložení dat v klasifikačních skupinách poměrně rovnoměrné, nemělo by docházet k tomu, že prvek z jedné množiny bude obklopen prvky převážně z jiné klasifikační skupiny, což by mohlo vést ke špatnému přiřazení prvku do klasifikační třídy. Výsledky klasifikace může značně ovlivnit počet sousedů k . Pro nalezení nejlepšího počtu byla použita křížová validace. Někdy se počet sousedů volí také jako druhá odmocnina z celkového počtu bodů v tréninkové sadě [78]. Počet k se volil v rozsahu 1-80, přičemž se braly v úvahu pouze liché hodnoty. Optimální počet sousedů pro 23 příznaků vybraných statistickými testy ($k = 49$) a pro příznaky vybrané metodou mRMR ($k = 39$) odpovídá takovému počtu sousedů, kde průměrné F1 skóre (blíže popsáno v kapitole 10.1) z křížové validace na testovací sadě (chytrý náramek) dosahuje nejlepších výsledků (Obrázek 9.1 a Obrázek 9.2). Jak lze vidět z grafu hodnota F1 skóre by se ještě do jisté míry nepatrně zvyšovala i s vyšším počtem k . Je třeba však brát na vědomí, že se zvyšujícím se počtem sousedů musí algoritmus pro každé nové pozorování vypočítat vzdálenosti ke všem svým sousedům v trénovací sadě. Tudiž se zvyšuje výpočetní náročnost. Při volbě vyššího rozsahu již nedochází k zlepšení F1 skóre, což naznačuje, že další zvyšování hodnoty k již není pro tuto klasifikační úlohu přínosné a nevede k žádnému zásadnímu zlepšení výkonnosti algoritmu. Tabulka 10.1 shrnuje výsledky klasifikace i s využitím KNN.



Obrázek 9.1: Optimální počet sousedů (10 příznaků)



Obrázek 9.2: Optimální počet sousedů (23 příznaků)

9.2.2 Logistická regrese

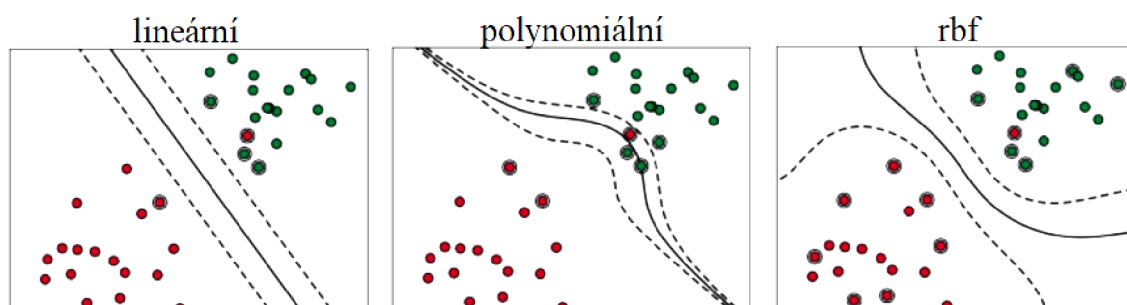
Logistická regrese (LR) je statistická metoda strojového učení používaná pro klasifikační úlohy. Jak plyne z názvu, využívá logistické funkce, též nazývané sigmoidální funkce. Hodnoty této funkce leží mezi jedničkou a nulou. [79]

Základní myšlenkou LR je předpovědět pravděpodobnost příslušnosti testovaného objektu k jedné ze dvou či více tříd. Během trénování se počítají koeficienty pro všechny příznaky tak, aby se minimalizovala chyba predikce k příslušné třídě. Tyto koeficienty jsou následně využity pro výpočet pravděpodobností příslušnosti k jednotlivým třídám pro nová vstupní data (testovací data). Testovaný objekt je následně přiřazen ke třídě, pro kterou je pravděpodobnost příslušnosti nejvyšší. Výhodou této metody je, že je velice jednoduchá na implementaci a je výpočetně nenáročná. Jedna z nevýhod LR spočívá v tom, že předpokládá lineární vztah mezi vstupními proměnnými a výstupem. Dále je tato metoda citlivá na odlehlé hodnoty, které však byly nahrazeny, jak je popsáno v kapitole 8.1. Jak dobře LR dokázala rozlišit mezi jednotlivými třídami popisuje Tabulka 10.1. [79]

9.2.3 Metoda podpůrných vektorů

Metoda podpůrných vektorů (anglicky support vectore machines, SVM) je metoda strojového učení, která je vhodná i pro lineárně neseparabilní data. V případě lineárně separabilních dat hledá přímku, kolem které se nachází tzv. toleranční pásmo. Ideálně metoda SVM nalezne přímku, která má co nejširší toleranční pásmo, ve kterém se zároveň nachází co nejmenší počet dat z trénovací množiny. Trénovací data náležící odlišným klasifikačním třídám pak leží na opačných stranách přímky. Pokud je toleranční pásmo příliš úzké, model je přeučeny a nová (testovací) data klasifikuje špatně. Body ležící na hranici tolerančního pásma se nazývají podpůrné vektory, z čehož plyne také název této metody. [80]

Jestliže jsou data lineárně neseparabilní, data jsou transformována do vyšších dimenzí, kde je možné je opět oddělit lineárně. V takovém případě lze místo původní množiny příznaků použít jejich kvadratické nebo kubické formy. Místo přímky pak hledáme optimální nadrovinu. Nelineární SVM transformuje prostor pomocí jader tzv. kernels. Použit lze například polynomiální jádro, radiální bázové jádro aj. (Obrázek 9.3). Obecně lze tvrdit, že N dat je možné rozdělit v prostoru o $N-1$ či více dimenzích. [80]



Obrázek 9.3: Data rozdělená pomocí různých jader SVM

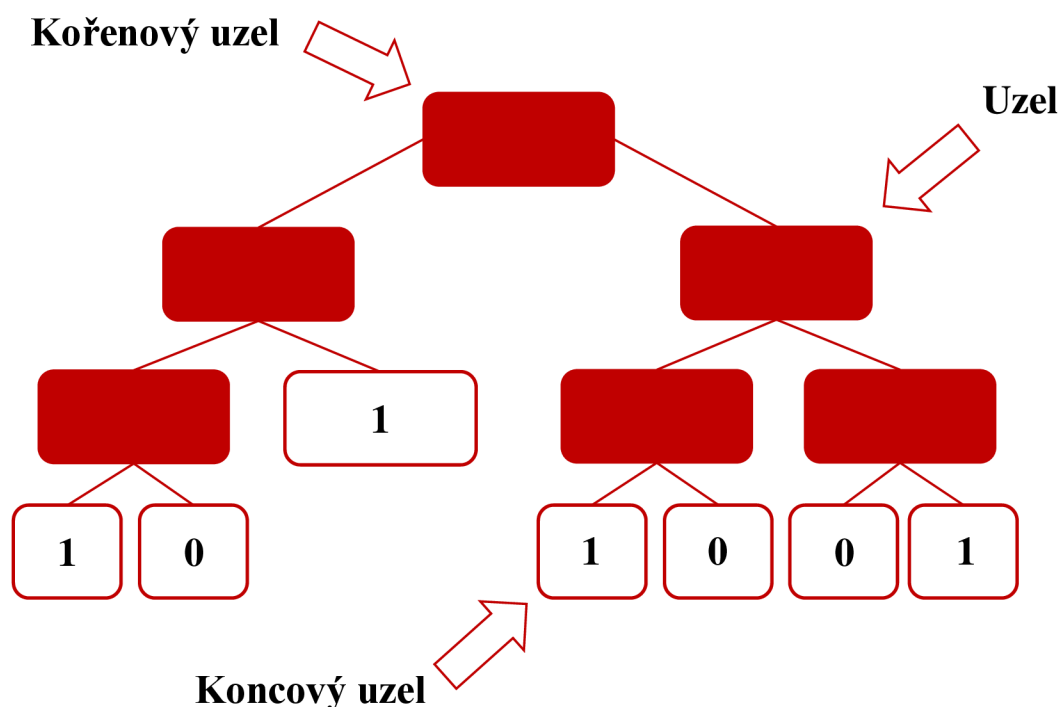
Zmiňovaná metoda SVM byla vyzkoušena i pro účely této DP, jelikož je poměrně jednoduchá a obecně dosahuje dobrých výsledků. Bylo vyzkoušeno SVM s jádrem lineárním, polynomiálním i rbf. SVM s rbf jádrem se ve všech případech osvědčilo jako nejvýkonnější.

9.2.4 Náhodný les

Náhodný les (Random Forest, RF) je algoritmus ze skupiny rozhodovacích stromů, který lze využít pro klasifikaci i regresi.

Rozhodovací stromy jsou hierarchické algoritmy vytvářející stromovou strukturu, která je tvořena uzly a větvemi. Každý jeden uzel odpovídá otázce (rozhodovacímu pravidlu dle vybraného příznaku). Na základě vylučujících se odpovědí na danou otázku jsou z tohoto uzlu vedeny větve do dceřiných uzlů, které opět představují nějakou odlišnou otázku. Objekt je postupně klasifikován od nejvrchnějšího uzlu (kořenu) až k poslednímu uzlu, kde dochází ke klasifikaci daného objektu (případně predikci). Schéma rozhodovacího stromu, který klasifikuje objekty do dvou skupin je na Obrázek 9.4.

RF je metoda strojového učení s učitelem. RF je založen na existenci několika různých rozhodovacích stromů. Trénovací sada pro každý strom je vytvořena náhodným výběrem dat z celé trénovací sady dat (tzv. bootstrapping). Dále je pro každý strom náhodně vybráno několik příznaků. Každý strom klasifikuje data z trénovací sady nezávisle na stromech ostatních. Výsledná klasifikace je poté určena ze všech “slabých” stromů na základě tzv. hlasování většiny. Právě kombinací “slabých” stromů vzniká robustnější a spolehlivější model. [81][82]



Obrázek 9.4: Schéma klasifikačního stromu

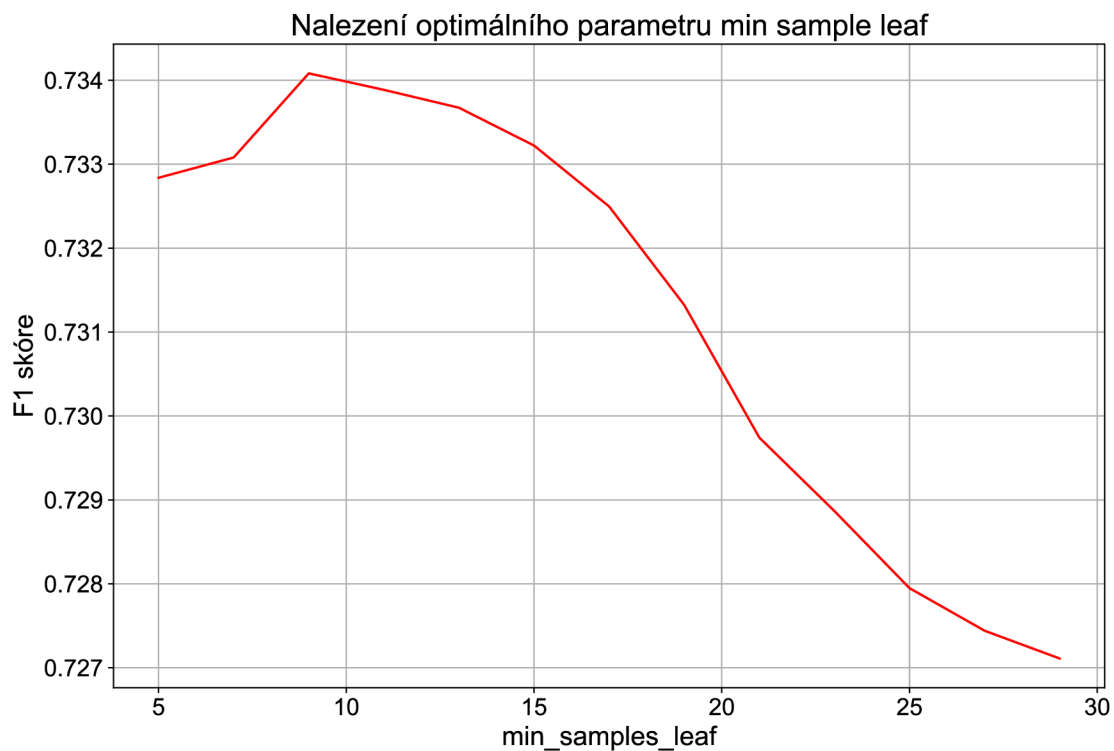
Důležité je zmínit, že trénování každého jednoho stromu probíhá na dvou třetinách dat – “In Bag” (IB). Zbývá jedna třetina dat odpovídá tzv. “Out Of Bag” (OOB) vzorkům, které lze využít k testování. Klasifikace každého vzorku je prováděna pouze

na těch stromech, u nichž se daný vzorek nachází právě v OOB. Poté je tedy výsledná klasifikace určena na základě hlasování většiny pouze z těchto stromů. Random Forest lze použít také pro predikci, celý proces je naprosto stejný. Liší se pouze v tom, že predikce vzorku je dána průměrnou predikcí z použitých stromů, nikoliv hlasováním většiny. [81][82]

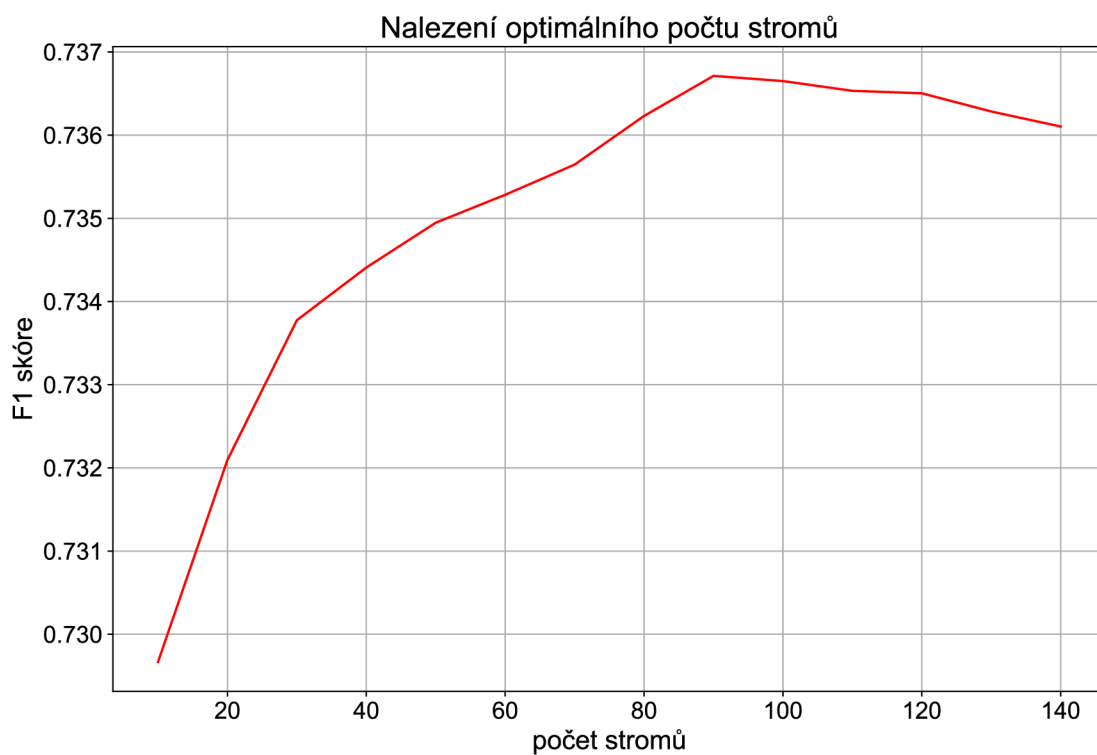
Náhodný les má díky kombinaci různých stromů dobrou generalizační schopnost. Neklade žádné požadavky na předzpracování dat a učení. [81]

Výsledky klasifikace můžeme opět ovlivnit různým nastavením parametrů jako je například počet stromů, nebo tzv. *min sample leaf*. Obecně platí, že čím více bude mít model stromů, tím bude stabilnější. Na druhou stranu se zvyšujícím se počtem stromů se bude zvyšovat také výpočetní náročnost algoritmu. Parametr *min sample leaf* určuje minimální počet případů, které musí být obsaženy v listu (koncovém uzlu), aby byl skutečně považován za koncový uzel. Pokud bude tento parametr příliš nízký, může dojít k přeučení. V opačném případě může být model náchylný k nedoučení a nebude schopen zachytit vztahy v datech.

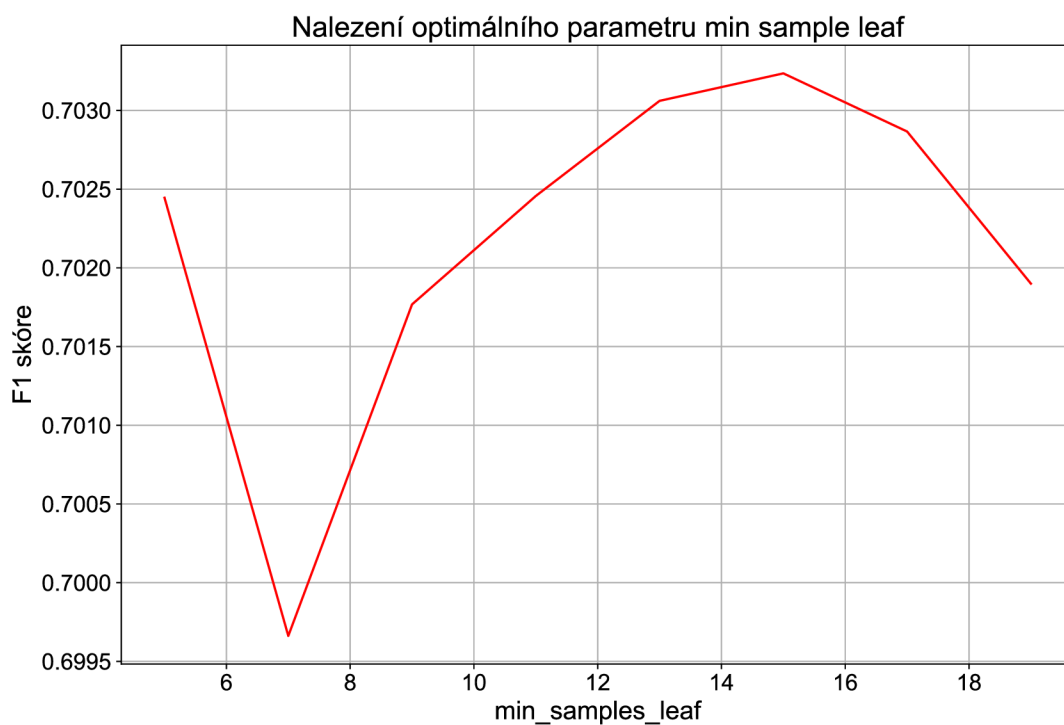
Pro nalezení optimálních parametrů byla opět využita křížová validace a průměrná hodnota F1 skóre. Nejvyšší hodnoty F1 skóre nabýval parametr *min sample leaf* = 9 a optimální *počet stromů* je pak 90 (Obrázek 9.5 a Obrázek 9.6) pro příznaky vybrané metodou mRMR. Optimální parametry pro všech 23 příznaků vybraných pomocí statistických testů je *min sample leaf* = 15 a *počet stromů* = 130 (Obrázek 9.7 a Obrázek 9.8). Výsledky klasifikace s využitím optimálních parametrů jsou v Tabulka 10.1.



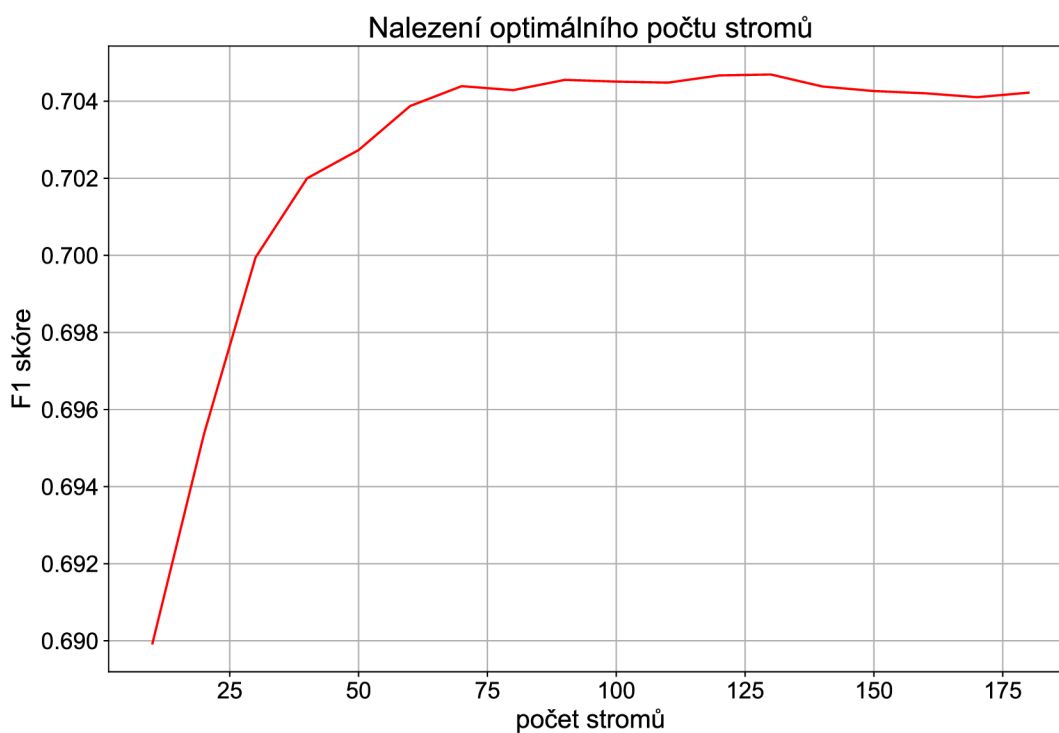
Obrázek 9.5: Nalezení optimálního parametru *min sample leaf* (10 příznaků)



Obrázek 9.6: Nalezení optimálního počtu stromů (10 příznaků)



Obrázek 9.7: Nalezení optimálního parametru *min sample leaf* (23 příznaků)



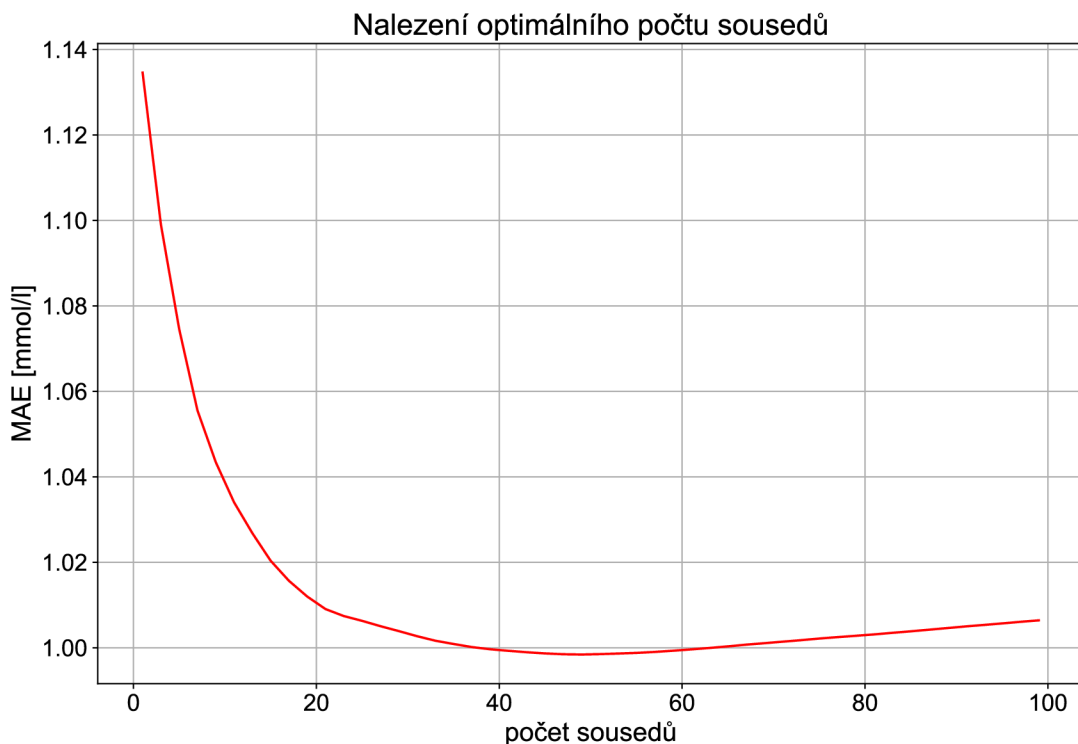
Obrázek 9.8: Nalezení optimálního počtu stromů (23 příznaků)

9.3 Regresní modely

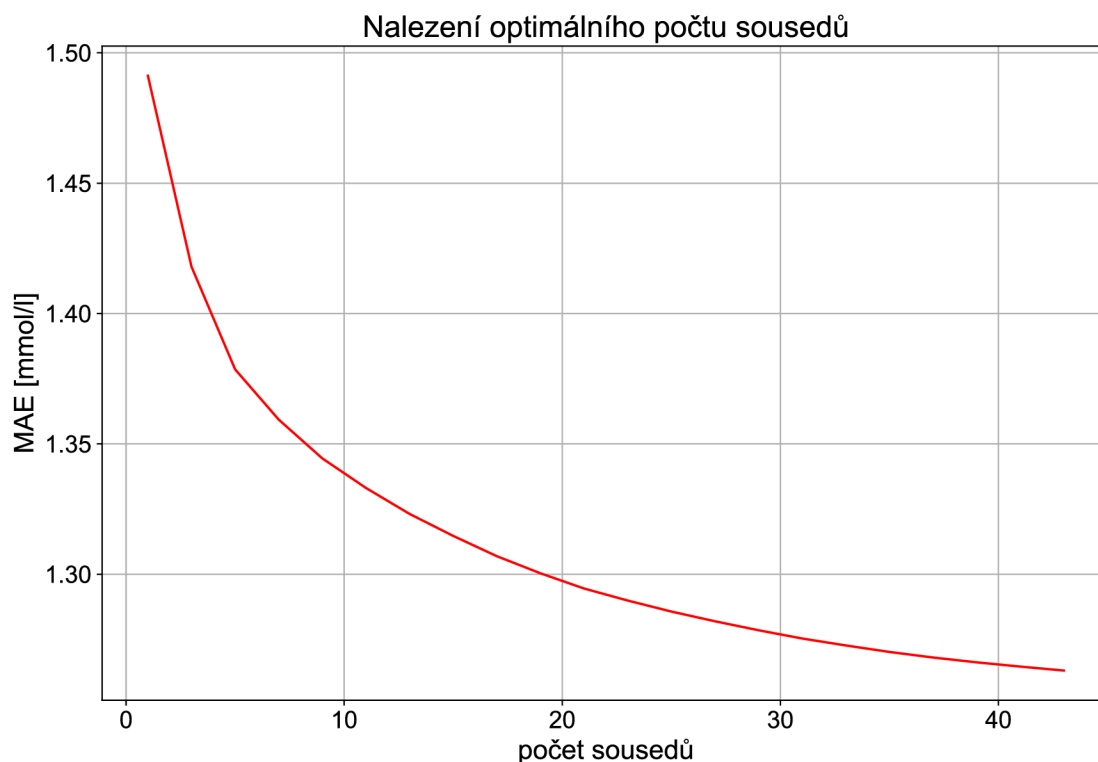
Tato kapitola se zabývá stručným popisem modelů a nastavením optimálních hyperparametrů pro regresní úlohu. Opět je využito obou sad dat (23 i 10 příznaků). Pro trénování je použita stejná sada dat i křížová validace jako v případě klasifikační úlohy. Výstupem však tentokrát nejsou třídy, ale spojité hodnoty. Cílem těchto modelů je tedy predikovat konkrétní hodnotu glykémie.

9.3.1 K-nejbližších sousedů

Algoritmus KNN byl již popsán v kapitole 9.2.1. Lze jej použít i pro regresní úlohy, kde výstupní hodnota testovaného objektu bude spočítána jako průměr z nejbližších sousedů. Opět byl s využitím křížové validace zjišťován optimální počet sousedů. Jak lze vidět z grafu, nejnižší MAE (blíže popsáno v kapitole 10.2), dosahuje počet sousedů $k=41$ v případě 23 příznaků (Obrázek 9.9). Pro příznaky vybrané metodou mRMR je optimální k zvoleno jako 29 (Obrázek 9.10). Se zvyšujícím se počtem sousedů by hodnota MAE nepatrně sice ještě klesala, ale na úkor výpočetní náročnosti, proto byl tedy počet sousedů nastaven na 29.



Obrázek 9.9: Nalezení optimálního počtu sousedů pro regresní úlohu (23 příznaků)



Obrázek 9.10: Nalezení optimálního počtu sousedů pro regresní úlohu (10 příznaků)

Tabulka 10.3 shrnuje výsledky modelu KNN pro regresní úlohu s optimálním počtem sousedů.

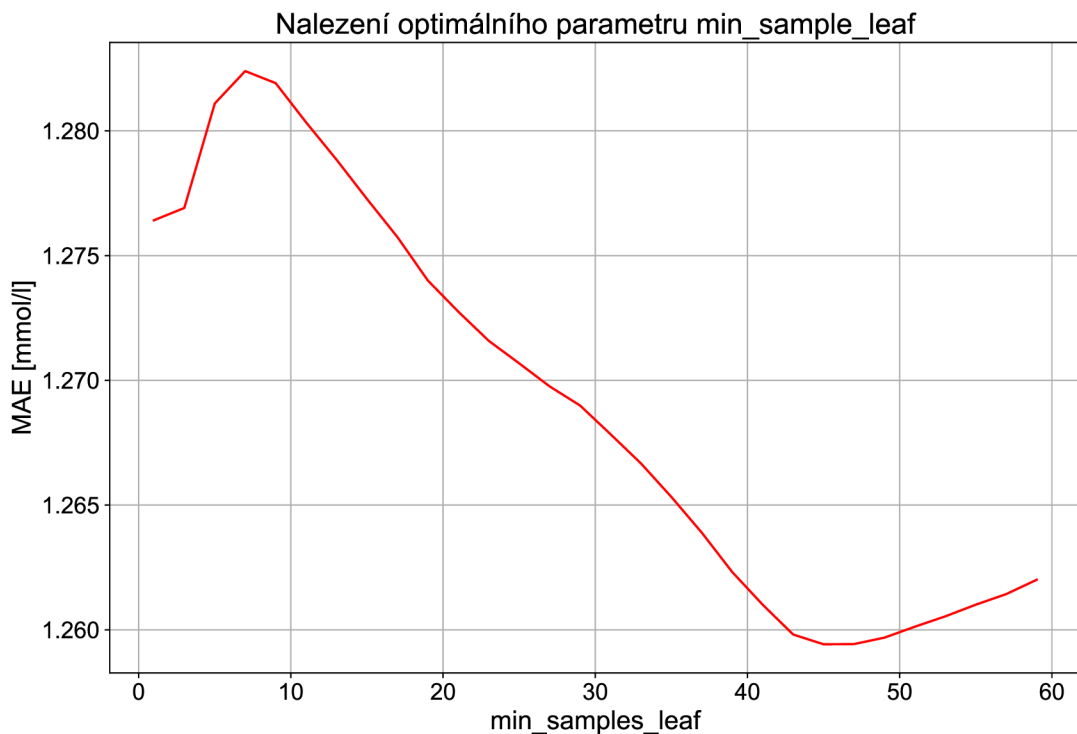
9.3.2 Metoda podpůrných vektorů

Metodu podpůrných vektorů lze využít také pro regresní úlohu (SVR). Princip metody SVR je ve své podstatě stejný jako princip SVM využívané pro klasifikaci. Cílem SVR je nalézt co nejlepší aproximaci mezi vstupními a výstupními daty, přičemž výstupem algoritmu je spojitá hodnota, nikoliv klasifikační třída. Cílem je nalézt nadrovinu, která minimalizuje rozdíl mezi skutečnou hodnotou výstupu a hodnotou predikovanou. Zároveň se SVR snaží maximalizovat vzdálenost mezi nadrovinou a nejbližšími vstupními daty, což zajišťuje lepší generalizační schopnost vůči datům novým. Pro regresní úlohu bylo využito opět rbf, polynomiální i lineární jádro, přičemž se rbf jádro jeví jako nejvhodnější.

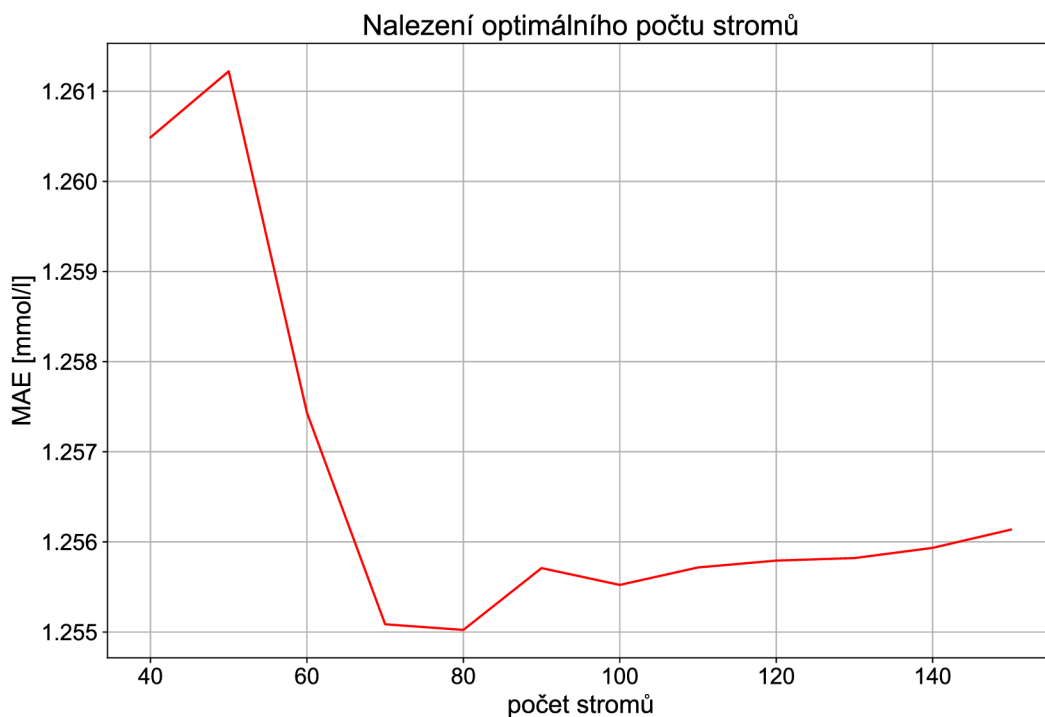
9.3.3 Náhodný les

Náhodný les (RF) lze použít také pro predikci, celý proces je naprosto stejný jako pro klasifikaci (kapitola 9.2.4). Liší se pouze v tom, že predikce vzorku je dána průměrnou predikcí z použitých stromů, nikoliv hlasováním většiny. [81][82]

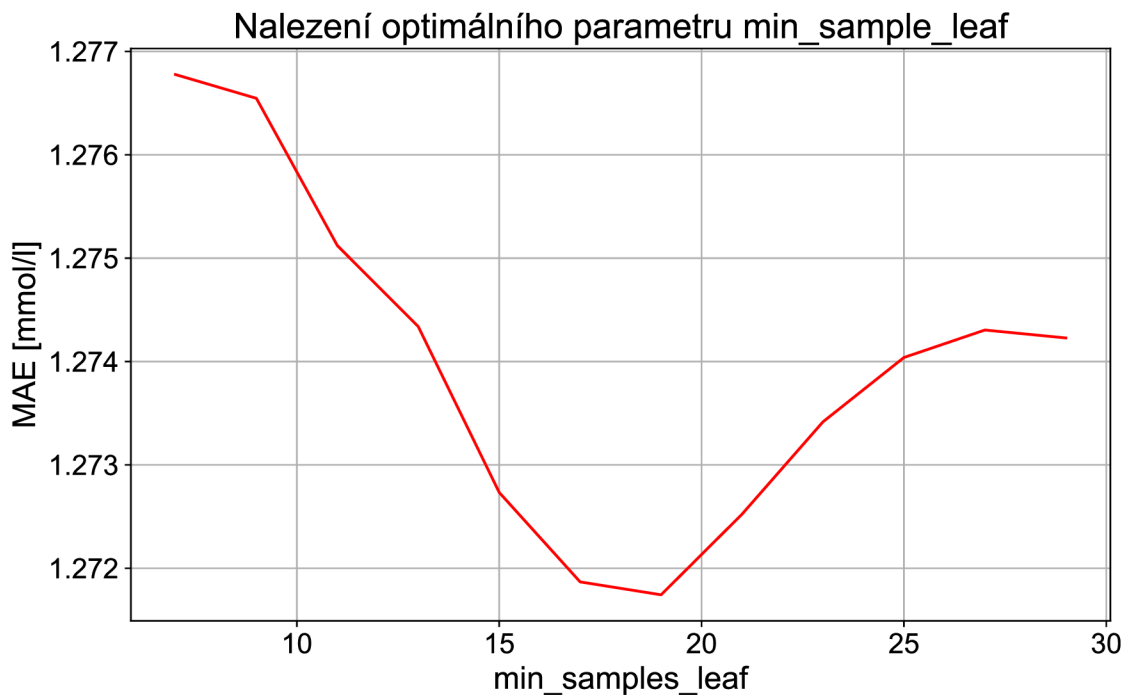
Pro predikci konkrétní hodnoty glykémie s využitím modelu RF byly opět nalezeny optimální parametry pro všech 23 příznaků (*min sample leaf* = 45). Tato hodnota byla stanovena dle nejnižší průměrné MAE pro testovací sadu dat z chytrého náramku s využitím křížové validace (Obrázek 9.11). Počet stromů byl nastaven na 80 (Obrázek 9.12). Pro sadu dat obsahujících pouze 10 příznaků byla parametr *min sample leaf* nastaven na 19 (Obrázek 9.13) a *počet stromů* na 100. Při volbě vyššího počtu stromů nedochází k výraznému zlepšení MAE, což naznačuje, že další zvyšování počtu stromů již není pro tuto predikci konkrétní hodnoty glykémie přínosné a nevede k žádnému zásadnímu zlepšení výkonnosti algoritmu (Obrázek 9.14). Zároveň nižší počet stromů odpovídá nižší výpočetní náročnosti.



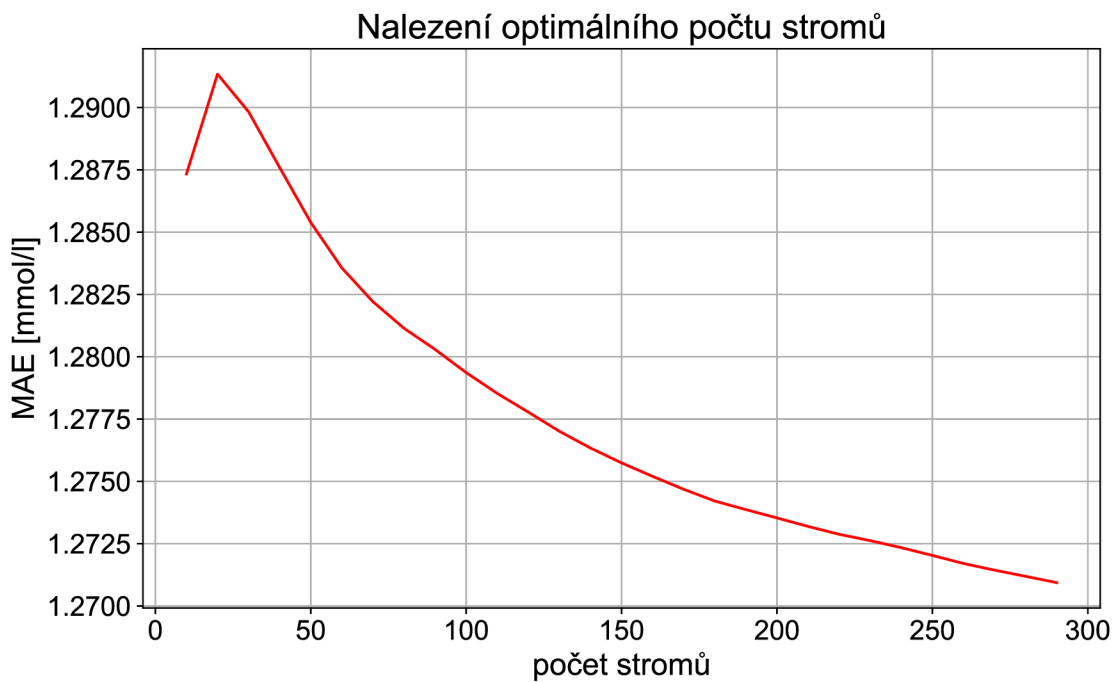
Obrázek 9.11: Nalezení optimálního počtu parametru *min samples leaf* pro regresi (23 příznaků)



Obrázek 9.12: Nalezení optimálního počtu stromů pro regresi (23 příznaků)



Obrázek 9.13: Nalezení optimálního počtu parametru min samples leaf pro regresi (10 příznaků)



Obrázek 9.14: Nalezení optimálního počtu stromů pro regresní úlohu (10 příznaků)

Úspěšnost RF pro predikci konkrétní hodnoty glykémie na testovací sadě dat z chytrého náramku shrnuje Tabulka 10.3.

10. ÚSPĚŠNOST JEDNOTLIVÝCH MODELŮ

Tato kapitola je věnována vyhodnocení úspěšnosti jednotlivých modelů pro klasifikaci i predikci. Nejdříve jsou uvedeny dosažené výsledky pro klasifikaci glykémie do dvou skupin – nízká / vysoká glykémie. Následující podkapitola se poté týká vyhodnocení úspěšnosti modelů predikujících konkrétní hodnotu glykémie. Na závěr kapitoly jsou dosažené výsledky diskutovány.

10.1 Výsledky klasifikace

Výsledky klasifikace pomocí všech výše zmíněných metod shrnuje Tabulka 10.1, kde je zahrnuta úspěšnost na trénovacích i testovacích datech. Testovací sada vždy obsahuje 279 PPG záznamů nasnímaných pomocí chytrého náramku. Tyto signály prošly stejným předzpracováním jako data z trénovací sady a byly pro ně vypočítané stejné příznaky. V rámci vyhodnocení úspěšnosti klasifikace je použito nejdříve všech 23 příznaků vybraných statistickou analýzou a poté 10 příznaků vybraných metodou mRMR, které se jeví jako vhodnější.

Pro posouzení úspěšnosti klasifikace záznamů byla použita matice záměn, která je tvořena čtyřmi základními hodnotami: TP, TN, FP a FN. True positive (TP) vyjadřuje počet záznamů, které byly správně klasifikovány do skupiny „vysoká glykémie“. True negative (TN) odpovídá počtu záznamů, které byly správně klasifikovány do skupiny „nízká glykémie“. False positive (FP) udává počet záznamů, které byly klasifikovány do skupiny „vysoká glykémie“, ve skutečnosti však nabývají hodnot nižších, než je daná prahová hodnota 7,2 mmol/l. False negative (FN) pak odpovídá počtu špatně klasifikovaných záznamů do skupiny „nízká glykémie“. [69]

Z výše zmíněných hodnot TP, TN, FP a FN se počítají ukazatelé správnosti – senzitivita, F1-skóre, přesnost a pozitivní prediktivní hodnota. Senzitivita (Se) udává pravděpodobnost, že bude záznam označen správně jako záznam patřící do třídy „vysoká glykémie“ (10.1). Přesnost (Acc) vyjadřuje kolik záznamů je správně klasifikováno (10.3). Specificita (Sp) je pravděpodobnost, že bude záznam správně označen jako záznam patřící do třídy „nízká glykémie“ (10.2). Pozitivní prediktivní hodnota (PPV) odpovídá procentu správně zařazených záznamů do třídy „vysoká glykémie“ ze všech záznamů klasifikovaných do této třídy (10.4). F1-skóre (F1) představuje celkovou přesnost a počítá se ze Se a PPV (10.5). Vzorce pro výpočet jednotlivých ukazatelů správnosti popisují následující rovnice [69]:

$$Se = \frac{TP}{TP+FN} \quad (10.1)$$

$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \quad (10.2)$$

$$Sp = \frac{TN}{TN+FP} \quad (10.3)$$

$$PPV = \frac{TP}{TP+FP} \quad (10.4)$$

$$F1 = 2 \frac{Se \times PPV}{Se + PPV} \quad (10.5)$$

Průměrné výsledky z křížové validace pro ukazatele spolehlivosti Acc a F1 skóre na trénovací i testovací sadě dat popisuje Tabulka 10.1. Je třeba zmínit, že se jedná o váhovanou formu těchto metrik, která bere v potaz počet vzorků v jednotlivých klasifikačních skupinách, jelikož i přes snahy vybalancovat datasety v průběhu měření, nejsou vzorky v jednotlivých skupinách úplně přesně rovnoměrně zastoupeny.

Tabulka 10.1: Výsledky klasifikace na trénovací a testovací sadě dat z Empaticy

Model			Trénovací sada (chytrý náramek)		Testovací sada (chytrý náramek)	
Klasifikátor	počet příznaků	optimální parametry	Acc [%]	F1 [%]	Acc [%]	F1 [%]
KNN	23	$k = 49$	81,2	81,9	69,4	68,4
KNN	10	$k = 39$	81,4	81,8	74,6	73,5
LR	23	-	78,8	79,6	71,9	71,3
LR	10	-	79,8	80,5	72,1	72,2
SVM (rbf)	23	-	88,2	88,6	74,7	74,9
SVM (rbf)	10	-	86,4	86,3	75,6	75,6
SVM (poly)	23	-	84,5	84,2	71,8	71,8
SVM (poly)	10	-	81,1	80,4	69,2	68,6
SVM (linear)	23	-	86,3	86,2	74,2	73,8
SVM (linear)	10	-	81,8	81,7	74	73,8
RF	23	$m_{sp} = 15$ $ps = 130$	88,9	89,3	69,8	69,4
RF	10	$m_{sp} = 9$ $ps = 90$	89,2	89,7	74,9	74,1

Následně byly vybrány nejlepší modely (zvýrazněné tučně) otestovány na sadě dat získané pomocí jiného zařízení – chytrého telefonu (Tabulka 10.2). Tato data prošla stejným předzpracováním jako data z chytrého náramku. Ověření funkčnosti na datech z chytrého telefonu proběhlo pro každý vybraný model (LR, KNN, SVM i RF)

následovně: během křížové validace ($k = 4$) byly tedy vytvořeny čtyři modely se stejným nastavením parametrů, které byly natrénované a poté otestované vždy na odlišných datech z chytrého náramku (Tabulka 10.1). Tyto 4 modely byly poté dodatečně otestovány na tentokrát již stejném datasetu z chytrého telefonu. Výsledná úspěšnost daného modelu na testovací sadě dat z chytrého telefonu je poté brána jako průměr z těchto 4 modelů vytvořených během křížové validace (Tabulka 10.2).

Tabulka 10.2: Výsledky klasifikace na testovací sadě dat z chytrého telefonu

Model			Testovací data (chytrý telefon)	
klasifikátor	počet příznaků	optimální parametry	Acc [%]	F1 [%]
KNN	10	$k = 39$	68	67
LR	10	-	67,3	68,2
SVM (rbf)	10	-	74	73,1
RF	10	$m_{sp} = 9$ $p_s = 90$	74,5	76,7

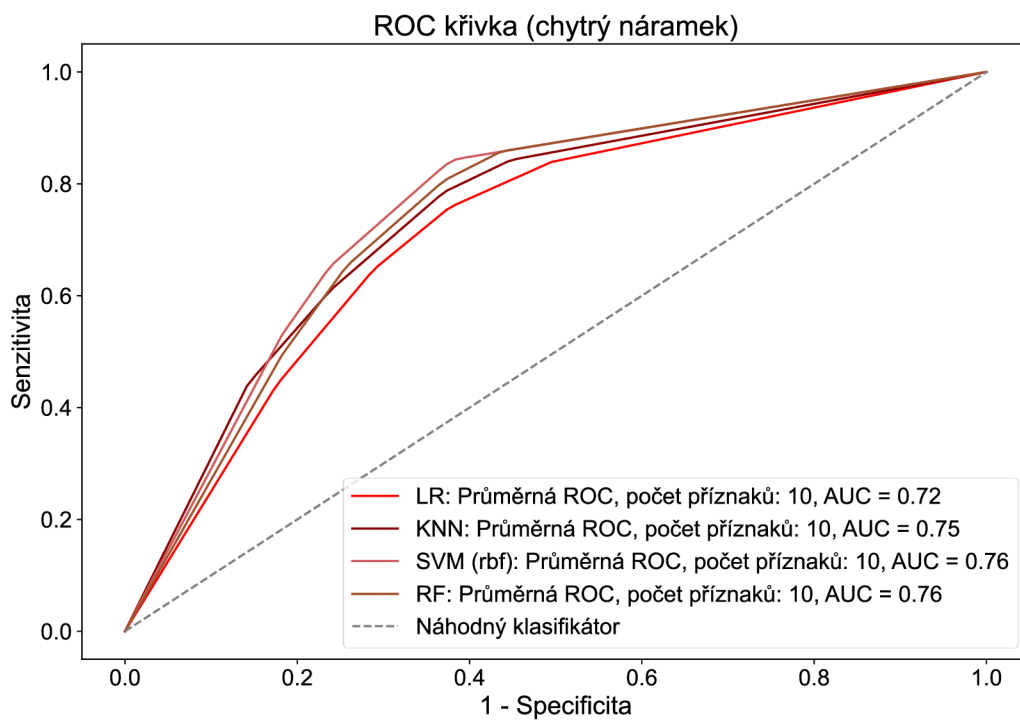
Pro posouzení kvality klasifikátoru se často také používá ROC křivka (receiver operating characteristic curve) a vyjádření plochy pod křivkou (AUC, area under the curve). Graf ROC křivky zobrazuje průběh závislosti negativní prediktivní hodnoty ($1 -$ specificita) na ose x a senzitivity na ose y pro všechny prahové hodnoty θ . Každému prahu θ přísluší jeden daný bod na ROC křivce. Křivka začíná v bodě (0,0) a končí v bodě (1,1), mezi těmito body je stoupající. Tvar ROC křivky závisí na zařazení objektů klasifikátorem do jednotlivých klasifikačních skupin. V ideálním případě, kdy klasifikátor zařadí všechny objekty správně, by ROC křivka měla procházet z bodu (0,0) do bodu (0,1) a následně do bodu (1,1). Čím je ROC křivka bližší tomuto průběhu, tím je klasifikátor kvalitnější. Diagonála v ROC prostoru poté odpovídá klasifikaci s pravděpodobností 0,5 a jde tedy o tzv. náhodný klasifikátor. [83]

Plocha pod křivkou je kvantitativní veličina, která popisuje ROC křivku. Lze ji spočítat dle rovnice (10.6). [83]

$$AUC = \int_0^1 ROC(p) dp \quad (10.6)$$

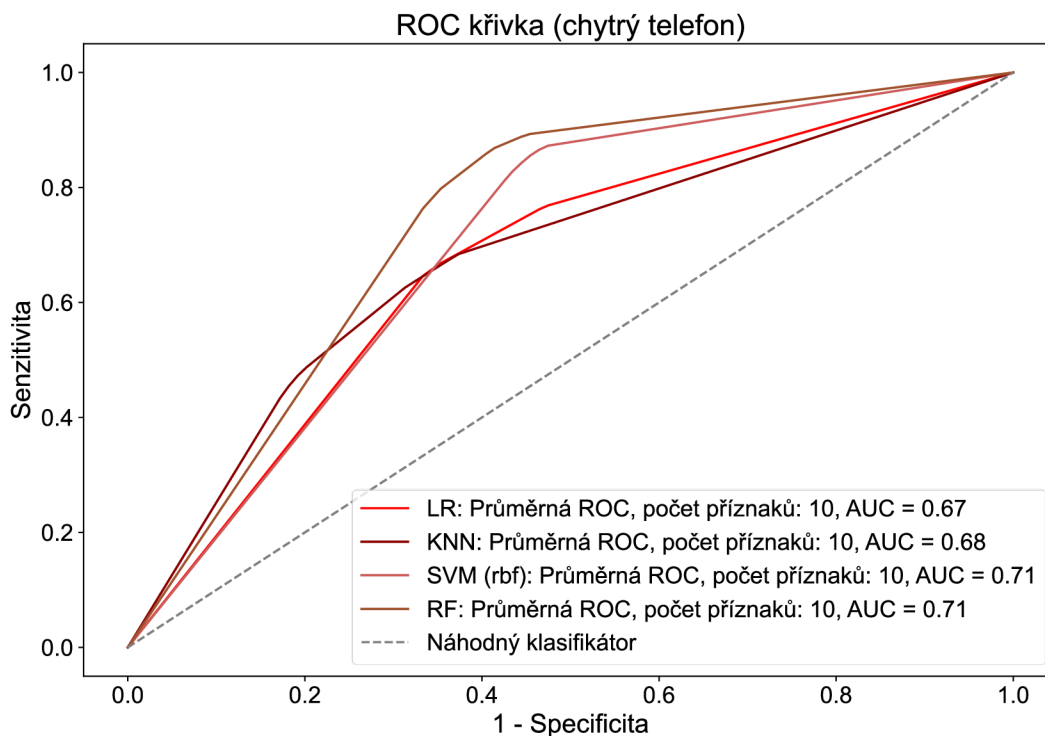
AUC nabývá hodnot $\langle 0, 1 \rangle$ a udává kvalitu daného modelu. Čím je hodnota blíže 1, tím je klasifikátor kvalitnější. [83]

Obrázek 10.1 zobrazuje průměrné ROC křivky z křížové validace pro jednotlivé vybrané modely na testovacích datech z chytrého náramku Empatica. Dále je v grafu znázorněna průměrná hodnota AUC pro jednotlivé klasifikátory. Je patrné, že dle AUC si nejlépe vede SVM s rbf jádrem a náhodný les (RF).



Obrázek 10.1: ROC křivky a hodnota AUC pro jednotlivé modely na testovací sadě dat z Empaticy

Obrázek 10.2 zobrazuje tedy průměrné ROC křivky pro jednotlivé vybrané modely na testovacích datech z chytrého telefonu. V grafu je také znázorněna průměrná hodnota AUC pro jednotlivé klasifikátory. Je patrné, že dle AUC si opět nejlépe vede SVM s rbf jádrem a náhodný les (RF). Avšak průměrná ROC křivka pro RF se více blíží ideálu.



Obrázek 10.2: ROC křivky a hodnota AUC pro jednotlivé modely na testovací sadě dat z chytrého telefonu

10.2 Výsledky predikce konkrétní hodnoty glykémie

Průměrné výsledky predikce konkrétní hodnoty glykémie pro trénovací i testovací sadu na datech získaných chytrým náramkem lze vidět v Tabulka 10.3.

Pro posouzení úspěšnosti predikce záznamů byla použita střední absolutní chyba, kterou lze spočítat dle vzorce:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (10.7)$$

kde y_i je predikovaná hodnota, \hat{y}_i je skutečná hodnota glykémie a n je počet vzorků.

Tabulka 10.3: Výsledky predikce glykémie na trénovací a testovací sadě dat z Empaticy

Model			Trénovací data (chytrý náramek)	Testovací data (chytrý náramek)
prediktor	počet příznaků	optimální parametry	MAE [mmol/l]	MAE [mmol/l]
KNN	23	$k = 41$	0,61	1,24
KNN	10	$k = 29$	0,89	1,23
SVR (rbf)	23	-	0,95	1,25
SVR (rbf)	10	-	0,78	1,26
SVR (poly)	23	-	0,95	1,75
SVR (poly)	10	-	1,03	1,7
SVR (linear)	23	-	0,98	1,28
SVR (linear)	10	-	1,1	1,4
RF	23	$m_{sp} = 45$ $p_s = 80$	0,86	1,25
RF	10	$m_{sp} = 19$ $p_s = 100$	0,73	1,27

Opět byly nejlepší vytvořené modely otestovány stejným způsobem na sadě dat získané pomocí jiného zařízení – chytrého telefonu (Tabulka 10.4).

Tabulka 10.4: Výsledky predikce glykémie nejvýkonnějších modelů na testovací sadě dat z chytrého telefonu

Model			Testovací data (chytrý telefon)
prediktor	počet příznaků	optimální parametry	MAE [mmol/l]
KNN	10	$k = 29$	1,58
SVR (rbf)	23	-	1,63
RF	23	$m_{sp} = 45$ $p_s = 80$	1,37

Přestože na testovací sadě dat z chytrého náramku model KNN dosahuje lepšího výsledku (MAE = 1,23 mmol/l) a model SVM s rbf jádrem stejného výsledku (MAE = 1,25 mmol/l), jako RF (MAE = 1,25 mmol/l), tak na testovací sadě dat z chytrého telefonu dosahuje podstatně nižší hodnoty MAE právě RF (MAE = 1,37 mmol/l). Z tohoto důvodu považuji za obecně lepší prediktor právě RF. V obou případech (RF i SVM) bylo trénování provedeno na 23 příznacích vybraných s využitím statistických testů, na rozdíl od klasifikačních modelů (pro nejlepší modely

použito 10 příznaků), což je dáno tím, že pro regresní úlohu může být vhodná jiná kombinace příznaků než pro úlohu klasifikační.

10.3 Diskuse

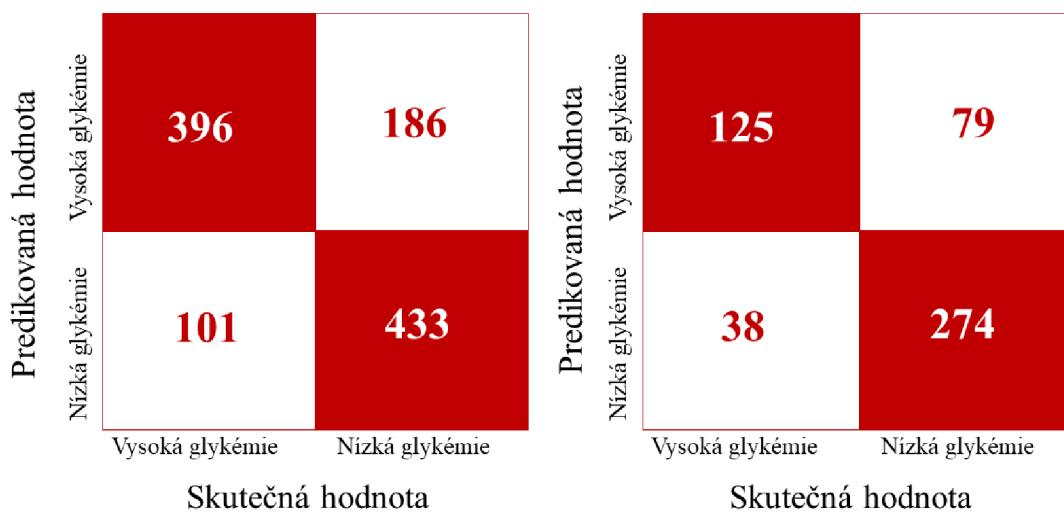
V této kapitole jsou diskutovány výsledky dosažené jednotlivými modely a vybrán nejlepší model pro klasifikaci i predikci. Dále je zde zmíněno porovnání modelů dosahujících nejlepších výsledků s jinými autory a limitace algoritmů.

Jedním z cílů práce je tedy neinvazivně klasifikovat hladinu glykémie z PPG signálů do dvou skupin – nízká a vysoká hladina glykémie na základě prahu 7,2 mmol/l, který byl určen dle manuálu glukometru (referenční hodnota glykémie). Tabulka 10.1 zobrazuje výsledky různých klasifikačních modelů. Nejlépe si vedl náhodný les (RF) a SVM s rbf jádrem. V obou případech byly nalezeny optimální parametry modelů. Pro účely klasifikace byly využity dvě sady příznaků (statistické metody – 23 příznaků a mRMR metoda – 10 příznaků). Pro oba klasifikační modely se jeví jako vhodná sada pouze 10ti příznaků. Modely byly otestovány na datech z chytrého náramku a dodatečně také na datech z jiného zařízení (chytrý telefon). V následující tabulce (Tabulka 10.5) jsou pro připomenutí opět uvedeny výsledky těchto modelů na testovací sadě dat z chytrého náramku i se směrodatnou odchylkou získanou z jednotlivých iterací křížové validace a výsledky na testovací sadě dat z chytrého telefonu.

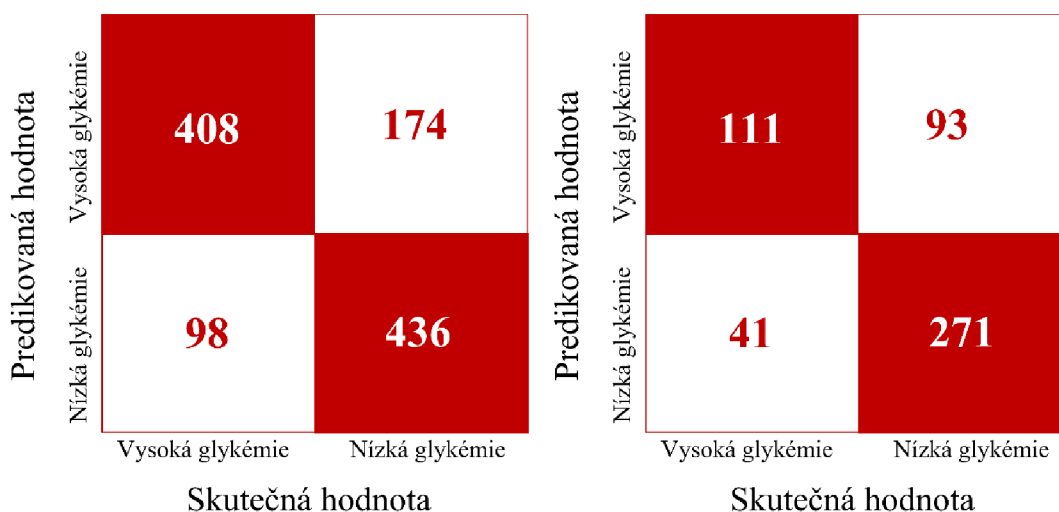
Tabulka 10.5: Výsledky klasifikace glykémie nejvýkonnějších modelů na obou testovacích sadách

		Testovací data (chytrý náramek)	Testovací data (chytrý telefon)
model	metrika		
SVM (rbf)	AUC	0,76	0,71
	Acc [%]	75,6	74
	std Acc [%]	3,8	-
	F1 [%]	75,6	73,1
	std F1 [%]	3,7	-
RF	AUC	0,76	0,71
	Acc [%]	74,9	74,5
	std Acc [%]	4,4	-
	F1 [%]	74,1	76,7
	std F1 [%]	4,8	-

Pro oba klasifikační modely jsou také zobrazeny matice záměn pro oba testovací datasety (Obrázek 10.3 a Obrázek 10.4).



Obrázek 10.3: Matice záměn RF klasifikátoru pro testovací sadu dat z chytrého náramku (vlevo) a testovací sadu dat z chytrého telefonu (vpravo)



Obrázek 10.4: Matice záměn SMV (rbf) klasifikátoru pro testovací sadu dat z chytrého náramku (vlevo) a testovací sadu dat z chytrého telefonu (vpravo)

Z matic záměn pro oba modely a obě testovací sady je patrné, že modely často klasifikují některé záznamy jako záznamy s vysokou glykemií, byť ve skutečnosti se jedná o záznamy s nízkou glykemií. To je pravděpodobně dáno tím, že některá data pochází ze stejného 15minutového měření, během kterého byly získány 3 hodnoty glykémie v rozsahu nad i pod prahovou hodnotou 7,2 mmol/l, nebo jsou velice blízko této prahové hodnotě. Přesnost obou algoritmů je na testovací sadě dat z chytrého náramku o něco vyšší než na testovací sadě dat z chytrého telefonu, což bylo předpokládáno, jelikož v druhém případě se jedná o data získaná z jiného zařízení (chytrého telefonu). Malé rozdíly v přesnosti však poukazují na schopnost algoritmu reagovat i na data z jiného zařízení,

než na kterém proběhlo trénování (chytrý náramek). Tabulka 10.6 zobrazuje porovnání výsledků klasifikace s jinými autory.

Tabulka 10.6: Porovnání výsledků s jinými autory (klasifikace)

Autor	PPG	Metoda	Acc [%]
DP	chytrý náramek	RF	74,9
DP	chytrý telefon	RF	74,5
DP	chytrý náramek	SVM (rbf)	75,6
DP	chytrý telefon	SVM (rbf)	74,0
G. Zhang et al. (2020) [84]	chytrý telefon	KNN	74,0
G. Zhang et al. (2020) [84]	chytrý telefon	Bagged Trees	71,0
Y. Zhang et al. [85]	chytrý telefon	Rozhodovací stromy	80,1

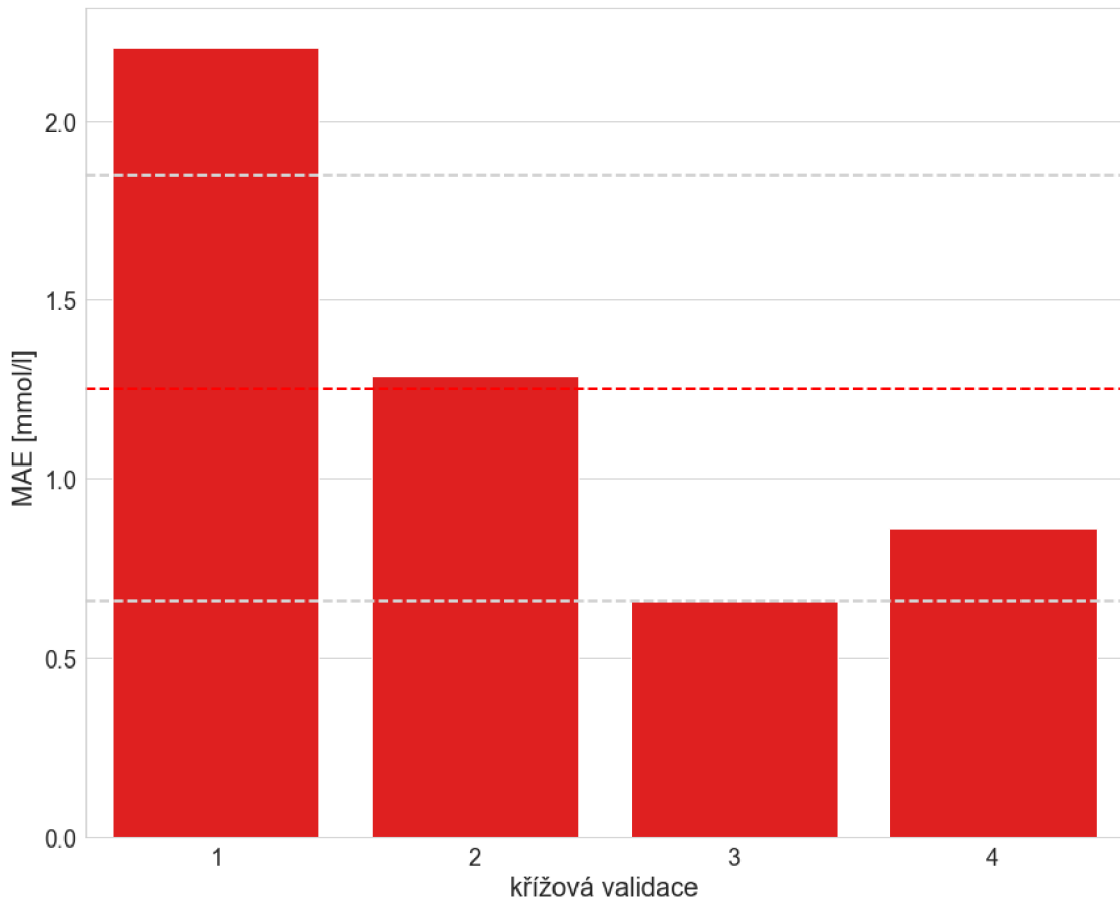
Dalším cílem této práce je pokusit se predikovat konkrétní hodnotu glykémie. Tabulka 10.3 shrnuje dosažené výsledky na testovací sadě dat z chytrého náramku. Funkčnost nejlepších modelů byla dodatečně ověřena na datech z chytrého telefonu (Tabulka 10.4). Tabulka 10.7 shrnuje výsledky nejlepšího predikčního modelu (RF). MAE na testovací sadě dat z chytrého telefonu dosahuje vyšších hodnot, pravděpodobně z toho důvodu, že model byl natrénován na datech z jiného zařízení (chytrý náramek).

Tabulka 10.7: Průměrná MAE nejlepšího modelu pro obě testovací sady dat

	Testovací data (chytrý náramek)	Testovací data (chytrý telefon)
model	MAE [mmol/l]	MAE [mmol/l]
RF	1,25	1,37

Obrázek 10.5 znázorňuje 4 hodnoty MAE z jednotlivých iterací křížové validace pro testovací sadu z chytrého náramku. Červená čára odpovídá průměrné hodnotě ze všech 4 iterací (1,25 mmol/l) a šedé čáry poté \pm směrodatné odchylce.

Srovnání MAE (křížová validace)



Obrázek 10.5: MAE pro jednotlivé iterace křížové validace na testovací sadě dat z chytrého náramku

Z grafu je patrné, že výsledná MAE na testovací sadě dat v první iteraci křížové validace je oproti ostatním poměrně vysoká a posouvá tak celkovou průměrnou hodnotu MAE na 1,25 mmol/l. Pokud by byly brány v potaz pouze následující tři iterace křížové validace, průměrná MAE by odpovídala 0,96 mmol/l. Právě tato skutečnost je zapříčiněna celkově nižším počtem subjektů (16) a také malým rozsahem naměřených hodnot glykémie (4,3–13,2 mmol/l). Například v první iteraci křížové validace dosahuje vysoké průměrné hodnoty MAE hned 1. subjekt v testovací sadě (3,3 mmol/l, což je nejvyšší průměrná MAE ze všech subjektů). Je to z toho důvodu, že se hladina glykémie této osoby pohybovala v rozsahu 4,3–4,5 mmol/l. Jedná se tedy o nejnižší hodnoty glykémie v celém rozsahu měřených hodnot glykémie, a tak nízká hodnota tedy v případě 1. iterace křížové validace nebyla přítomná v trénovací sadě dat. Výsledky z jednotlivých iterací jsou tedy ovlivněny subjekty v trénovací sadě a jejich celkovým počtem, což bylo očekáváno. Přesto je důležité zmínit, že práce je experimentální a dosažené výsledky ukazují, že tato metoda má potenciál.

Průměrná MAE na testovací sadě dat z chytrého telefonu je o něco vyšší a činí (1,37 mmol/l), což je dáno tím, že se jedná o data z jiného zařízení, než která byla použita pro trénování (chytrý náramek).

Tabulka 10.8: Porovnání výsledků s jinými autory (predikce konkrétní hodnoty glykémie)

Autor	Metoda	MAE [mmol/l]
DP	RF	1,25
DP	RF	1,37
L. Guzman et al. (2020) [86]	SVR	0,9
M. Shamim et al. (2022) [87]	CART	1,61
Z. Nie et al. (2023) [88]	RF	1,72
B. Manurung et al. (2019) [89]	DL	0,32

Tabulka 10.8 ukazuje porovnání výsledků predikce konkrétní hodnoty glykémie z PPG signálů s jinými autory. Z tabulky je patrné, že výsledky z algoritmů navržených v této práci jsou srovnatelné s jinými autory. Autoři [89] dosáhli velice dobrého výsledku (MAE = 0,32 mmol/l), avšak pro účely predikce použili metody hlubokého učení, která dává obecně lepší výsledky. Zde je důležité zmínit, že právě v důsledku nedostatku dat jsem se tímto směrem nemohla vydat.

I přes výše zmíněné limitace, práce dosahuje v klasifikační úloze i predikci konkrétní hodnoty glykémie dobrých výsledků, srovnatelných s jinými autory a poukazuje tak na potenciál využití chytrých zařízení a PPG signálů pro neinvazivní stanovení glykémie. Zároveň všechny modely vychází pouze z příznaků, které lze získat přímo z PPG signálu a nejsou tak závislé na žádných dalších informacích o měřených subjektech (věk, váha, výška aj.), což považuji za velkou výhodu.

ZÁVĚR

Tato práce se zabývá hodnocením zdraví pomocí chytrých zařízení, konkrétně neinvazivním stanovením glykémie z PPG signálů. První a druhá kapitola práce se věnuje literární rešerši v oblasti hodnocení zdraví pomocí chytrých zařízení. Třetí kapitola pojednává o glykémii a diabetu mellitu. Fotopletysmografií a možnostem stanovení glykémie z PPG signálů se věnuje kapitola čtvrtá a pátá.

Následující kapitoly se již věnují praktické části diplomové práce, jejímž cílem bylo pokusit se klasifikovat hodnoty glykémie z PPG signálů do skupin „nízká“ a „vysoká“ glykémie, a následně predikovat konkrétní hodnotu glykémie.

Pro účely této práce byly nasnímány dvě databáze PPG signálů. Jedna obsahuje data snímaná chytrým náramkem Empatica. Druhá databáze obsahuje PPG signály nasnímané chytrým telefonem a byla použita pouze k ověření funkčnosti navržených algoritmů na datech získaných z jiného zařízení. Součástí přípravy datasetu bylo také určení kvality PPG signálů snímaných chytrým telefonem, které bývají často nekvalitní.

Poté byly PPG signály z obou datasetů normalizovány a předzpracovány. Data nasnímaná z chytrého náramku sloužila k trénování i testování následně navržených algoritmů. Na těchto datech bylo vypočítáno 27 příznaků. Následně byl proveden výběr příznaků dle statistických testů (23 vybraných příznaků) a selekční metody mRMR (10 vybraných příznaků). Pro klasifikaci byly použity tyto modely: LR, KNN, SVM a RF. Pro predikci byly poté použity tyto modely: KNN, SVM pro regresi a RF. Zmíněné modely jsou popsány v deváté kapitole. Pro účely DP byla využita také křížová validace, která je použita jak pro hledání optimálních parametrů jednotlivých modelů, tak také pro věrohodnější vyhodnocení výkonnosti jednotlivých modelů.

Poslední kapitola se již věnuje dosaženým výsledkům pro klasifikační a regresní úlohu a ověření nejlepších modelů na datové sadě nasnímané pomocí chytrého telefonu. Nejlépe se s klasifikací PPG záznamů do dvou skupin na základě 10 příznaků vypořádaly modely RF a SVM s rbf jádrem. Tyto modely dosahují přesnosti Acc na testovací sadě dat z chytrého náramku 76 % (SVM) a 75 % (RF). Ověření funkčnosti navržených modelů je následně provedeno na testovací sadě dat z chytrého telefonu, kde oba modely dosahují obdobných výsledků: 74 % (SVM) a 75 % (RF).

Nejnižší MAE poté dosahoval RF trénovaný na 23 příznacích vybraných dle statistických testů. Tento model je schopen predikovat hladinu glykémie na testovací sadě dat z chytrého náramku s průměrnou absolutní chybou 1,25 mmol/l a na testovací sadě dat z chytrého telefonu poté s průměrnou absolutní chybou 1,37 mmol/l.

Dosažené výsledky pro klasifikační i regresní úlohu jsou srovnatelné s jinými autory, v některých případech i lepší. Jednou z limitací této práce je však počet měřených subjektů (16) a také rozsah měřené glykémie (4,3 – 13,2 mmol/l). V tuto chvíli je však důležité zmínit, že neexistuje žádná volně dostupná databáze PPG

signálů obsahující také referenční hodnoty glykémie, a tak musela být veškerá data naměřena. Na závěr bych ráda zmínila, že se jedná o experimentální práci, přičemž stanovení glykémie z PPG signálů zatím není tak prozkoumáno jako například stanovení jiného parametru zdraví (TF, KT, aj.). I přes výše zmíněné limitace, práce dosahuje dobrých výsledků a poukazuje na potenciál využití chytrých zařízení a PPG signálů pro neinvazivní stanovení glykémie.

Literatura

- [1] LU, Lin, Jiayao ZHANG, Yi XIE, Fei GAO, Song XU, Xinghuo WU a Zhewei YE. Wearable Health Devices in Health Care: Narrative Systematic Review. *JMIR mHealth and uHealth* [online]. 2020, 8(11). ISSN 2291-5222. Dostupné z: doi:10.2196/18907
- [2] *Wearable and Implantable Medical Devices* [online]. Elsevier, 2020. ISBN 9780128153697. Dostupné z: doi:10.1016/C2017-0-03249-4
- [3] DIAS, Duarte a João P. S. CUNHA. Wearable Health Devices—Vital Sign Monitoring, Systems and Technologies. *Sensors* [online]. 2018, **18**(8), 2414. ISSN 1424-8220. Dostupné z: doi:10.3390/s18082414
- [4] OMETOV, Aleksandr, Viktoriia SHUBINA, Lucie KLUS, Justyna SKIBIŃSKA, Salwa SAAFI, Pavel PASCACIO, Laura FLUERATORU, Darwin Quezada GAIBOR, Nadezhda CHUKHNO, Olga CHUKHNO, Asad ALI, Asma CHANNA, Ekaterina SVERTOKA, Waleed bin QAIM, Raúl CASANOVA-MARQUÉS, Sylvia HOLCER, Joaquín TORRES-SOSPEDRA, Sven CASTELEYN, Giuseppe RUGGERI, Giuseppe ARANITI, Radim BURGET, Jiri HOSEK a Elena Simona LOHAN. A Survey on Wearable Technology: History, State-of-the-Art and Current Challenges. *Computer Networks* [online]. 2021, **193**, 108074. ISSN 13891286. Dostupné z: doi:10.1016/j.comnet.2021.108074
- [5] SOUČEK, Miroslav a Petr SVAČINA. *Vnitřní lékařství v kostce*. 1. vydání. Praha: Grada Publishing, 2019. ISBN 9788027122899.
- [6] ISLAM, Tanvir Tazul, Md Sajid AHMED, Md HASSANUZZAMAN, Syed Athar BIN AMIR a Tanzilur RAHMAN. Blood Glucose Level Regression for Smartphone PPG Signals Using Machine Learning. *Applied Sciences* [online]. 2021, **11**(2), 618. ISSN 2076-3417. Dostupné z: doi:10.3390/app11020618
- [7] *International Diabetes Federation. IDF Diabetes Atlas*. 10th ed. Brussels, Belgium: <https://www.diabetesatlas.org>, 2021.
- [8] BANDAY, Mujeeb Z, Aga S SAMEER a Saniya NISSAR. Pathophysiology of diabetes: An overview. *Avicenna Journal of Medicine* [online]. 2020, **10**(04), 174–188. ISSN 2231-0770. Dostupné z: doi:10.4103/ajm.ajm_53_20
- [9] CHOI, Yosoon a Yeanjae KIM. Applications of Smart Helmet in Applied Sciences: A Systematic Review. *Applied Sciences* [online]. 2021, **11**(11), 5039. ISSN 2076-3417. Dostupné z: doi:10.3390/app11115039
- [10] VON ROSENBERG, Wilhelm, Theerasak CHANWIMALUEANG, Valentin GOVERDOVSKY, David LOONEY, David SHARP a Danilo P. MANDIC. Smart Helmet: Wearable Multichannel ECG and EEG. *IEEE Journal of Translational Engineering in Health and Medicine* [online]. 2016, **4**, 1–11. ISSN 2168-2372. Dostupné z: doi:10.1109/JTEHM.2016.2609927

- [11] MITRASINOVIC, Stefan, Elvis CAMACHO, Nirali TRIVEDI, Julia LOGAN, Colson CAMPBELL, Robert ZILINYI, Bryan LIEBER, Eliza BRUCE, Blake TAYLOR, David MARTINEAU, Emmanuel L.P. DUMONT, Geoff APPELBOOM a E. Sander CONNOLLY JR. Clinical and surgical applications of smart glasses. *Technology and Health Care* [online]. 2015, **23**(4), 381–401. ISSN 09287329. Dostupné z: doi:10.3233/THC-150910
- [12] REEDER, Blaine a Alexandria DAVID. Health at hand: A systematic review of smart watch uses for health and wellness. *Journal of Biomedical Informatics* [online]. 2016, **63**, 269–276. ISSN 15320464. Dostupné z: doi:10.1016/j.jbi.2016.09.001
- [13] BEANS, Carolyn. Wearable tech meets tattoo art in a bid to revolutionize both. *Proceedings of the National Academy of Sciences* [online]. 2018, **115**(14), 3504–3506. ISSN 0027-8424. Dostupné z: doi:10.1073/pnas.1803214115
- [14] ANGELUCCI, Alessandra, Matteo CAVICCHIOLI, Ilaria CINTORRINO, Giuseppe LAURICELLA, Chiara ROSSI, Sara STRATI a Andrea ALIVERTI. Smart Textiles and Sensorized Garments for Physiological Monitoring: A Review of Available Solutions and Techniques. *Sensors* [online]. 2021, **21**(3), 814. ISSN 1424-8220. Dostupné z: doi:10.3390/s21030814
- [15] BANAEI, Hadi, Mobyen AHMED a Amy LOUTFI. Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges. *Sensors* [online]. 2013, **13**(12), 17472–17500. ISSN 1424-8220. Dostupné z: doi:10.3390/s131217472
- [16] ALEXANDER, Sarah, Guy PERYER, Emma GRAY, Frederik BARKHOF a Jeremy CHATAWAY. Wearable technologies to measure clinical outcomes in multiple sclerosis: A scoping review. *Multiple Sclerosis Journal* [online]. 2021, **27**(11), 1643–1656. ISSN 1352-4585. Dostupné z: doi:10.1177/1352458520946005
- [17] KOLÁŘ, Radim. *Lékařská diagnostická technika, elektronická skripta*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství, 2007.
- [18] LANDREANI, Federica, Damien GOLIER, Amin HOSSEIN, Jérémy RABINEAU, Philippe VAN DE BORNE, Enrico CAIANI a Pierre-François MIGEOTTE. Heartbeat Detection Using Three-Axial Seismocardiogram Acquired by Mobile Phone. 2018. Dostupné z: doi:10.22489/CinC.2018.215
- [19] GAVRIEL, Constantinos, Kim H. PARKER a A. Aldo FAISAL. Smartphone as an ultra-low cost medical tricorder for real-time cardiological measurements via ballistocardiography. In: *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)* [online]. IEEE, 2015, s. 1–6. ISBN 978-1-4673-7201-5. Dostupné z: doi:10.1109/BSN.2015.7299425

- [20] MISRA, Aditi, Rohan BANERJEE, Anirban Dutta CHOUDHURY, Aniruddha SINHA a Arpan PAL. Novel peak detection to estimate HRV using smartphone audio. In: *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)* [online]. IEEE, 2015, s. 1–6. ISBN 978-1-4673-7201-5. Dostupné z: doi:10.1109/BSN.2015.7299378
- [21] MORAES, Jermana, Matheus ROCHA, Glauber VASCONCELOS, José VASCONCELOS FILHO, Victor DE ALBUQUERQUE a Auzuir ALEXANDRIA. Advances in Photoplethysmography Signal Analysis for Biomedical Applications. *Sensors* [online]. 2018, **18**(6), 1894. ISSN 1424-8220. Dostupné z: doi:10.3390/s18061894
- [22] CHOWDHURY, Moajjem Hossain, Md Nazmul Islam SHUZAN, Muhammad E.H. CHOWDHURY, Zaid B. MAHBUB, M. Monir UDDIN, Amith KHANDAKAR a Mamun Bin Ibne REAZ. Estimating Blood Pressure from the Photoplethysmogram Signal and Demographic Features Using Machine Learning Techniques. *Sensors* [online]. 2020, **20**(11), 3127. ISSN 1424-8220. Dostupné z: doi:10.3390/s20113127
- [23] NĚMCOVÁ, Andrea. Application of SWT for Heart Rate Monitoring Using Smartphone Camera. 2017.
- [24] SIDDIQUI, Sarah Ali, Yuan ZHANG, Zhiquan FENG a Anton KOS. A Pulse Rate Estimation Algorithm Using PPG and Smartphone Camera. *Journal of Medical Systems* [online]. 2016, **40**(5), 126. ISSN 0148-5598. Dostupné z: doi:10.1007/s10916-016-0485-6
- [25] CHARLTON, Peter H., Drew A. BIRRENKOTT, Timothy BONNICI, Marco A. F. PIMENTEL, Alistair E. W. JOHNSON, Jordi ALASTRUEY, Lionel TARASSENKO, Peter J. WATKINSON, Richard BEALE a David A. CLIFTON. Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Reviews in Biomedical Engineering* [online]. 2018, **11**, 2–20. ISSN 1937-3333. Dostupné z: doi:10.1109/RBME.2017.2763681
- [26] Vital Signs. In: *Cleveland Clinic* [online]. Dostupné z: <https://my.clevelandclinic.org/health/articles/10881-vital-signs>.
- [27] KOZUMPLÍK Jiří. *Analýza biologických signálů, elektronická skripta*. Brno: Ústav biomedicínského inženýrství FEKT VUT v Brně, 2012.
- [28] KOZUMPLIK, Jiri, Lukas SMITAL, Andrea NEMCOVA, Marina RONZHINA, Radovan SMISEK, Lucie MARSANOVA, Martin KRALIK a Martin VITEK. Respiratory Rate Estimation Using the Photoplethysmogram: Towards the Implementation in Wearables. In: *2021 Computing in Cardiology (CinC)* [online]. IEEE, 2021, s. 1–4. ISBN 978-1-6654-7916-5. Dostupné z: doi:10.23919/CinC53138.2021.9662674

- [29] JAAFAR, Rosmina a Mohd Aliff Azroy ROZALI. Estimation of breathing rate and heart rate from photoplethysmogram. In: *2017 6th International Conference on Electrical Engineering and Informatics (ICEEI)* [online]. IEEE, 2017, s. 1–4. ISBN 978-1-5386-0475-5. Dostupné z: doi:10.1109/ICEEI.2017.8312414
- [30] LABATE, Domenico, Fabio La FORESTA, Gianluigi OCCHIUTO, Francesco Carlo MORABITO, Aime LAY-EKUAKILLE a Patrizia VERGALLO. Empirical Mode Decomposition vs. Wavelet Decomposition for the Extraction of Respiratory Signal From Single-Channel ECG: A Comparison. *IEEE Sensors Journal* [online]. 2013, **13**(7), 2666–2674. ISSN 1530-437X. Dostupné z: doi:10.1109/JSEN.2013.2257742
- [31] SHARMA, Hemant, K.K. SHARMA a Om Lata BHAGAT. Respiratory rate extraction from single-lead ECG using homomorphic filtering. *Computers in Biology and Medicine* [online]. 2015, **59**, 80–86. ISSN 00104825. Dostupné z: doi:10.1016/j.combiomed.2015.01.024
- [32] CHARLTON, Peter H., Drew A. BIRRENKOTT, Timothy BONNICI, Marco A. F. PIMENTEL, Alistair E. W. JOHNSON, Jordi ALASTRUEY, Lionel TARASSENKO, Peter J. WATKINSON, Richard BEALE a David A. CLIFTON. Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review. *IEEE Reviews in Biomedical Engineering* [online]. 2018, **11**, 2–20. ISSN 1937-3333. Dostupné z: doi:10.1109/RBME.2017.2763681
- [33] ROKYTA, Richard. *Fyziologie*. Třetí vydání. Praha: Galén, 2016.
- [34] CHANDRASEKARAN, Vikram, Ram DANTU, Srikanth JONNADA, Shanti THIYAGARAJA a Kalyan Pathapati SUBBU. Cuffless Differential Blood Pressure Estimation Using Smart Phones. *IEEE Transactions on Biomedical Engineering* [online]. 2013, **60**(4), 1080–1089. ISSN 0018-9294. Dostupné z: doi:10.1109/TBME.2012.2211078
- [35] Blood Pressure Monitor on Smartwatch: How it Works and Is it Reliable? In: *Wearables Drop* [online]. Dostupné z: <https://wearablesdrop.com/how-blood-pressure-monitoring-works-in-smartwatch>
- [36] STEINMAN, Joe, Andrew BARSZCZYK, Hong-Shuo SUN, Kang LEE a Zhong-Ping FENG. Smartphones and Video Cameras: Future Methods for Blood Pressure Measurement. *Frontiers in Digital Health* [online]. 2021, **3**. ISSN 2673-253X. Dostupné z: doi:10.3389/fgdth.2021.770096
- [37] LAMONACA, Francesco, Domenico Luca CARNI, Domenico GRIMALDI, Alfonso NASTRO, Maria RICCIO a Vitaliano SPAGNOLO. Blood oxygen saturation measurement by smartphone camera. In: *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings* [online]. IEEE, 2015, s. 359–364. ISBN 978-1-4799-6477-2. Dostupné z: doi:10.1109/MeMeA.2015.7145228

- [38] O2Ring. In: *Wellue* [online]. Dostupné z: <https://getwellue.com/pages/o2ring-oxygen-monitor>
- [39] KREHEL, Marek, Martin WOLF, Luciano F. BOESEL, René M. ROSSI, Gian-Luca BONA a Lukas J. SCHERER. Development of a luminous textile for reflective pulse oximetry measurements. *Biomedical Optics Express* [online]. 2014, **5**(8), 2537. ISSN 2156-7085. Dostupné z: doi:10.1364/BOE.5.002537
- [40] HAMPTON John R. *EKG stručně, jasně, přehledně*. Vyd. 2., rozš. Praha: Grada, 2005.
- [41] GANONG William F. *Přehled lékařské fyziologie*. Dvacáté vydání. Praha: Galén, 2005.
- [42] SINHA, Raju. *An Approach for Classifying ECG Arrhythmia Based on Features Extracted from EMD and Wavelet Packet Domains*. 2012. Dostupné z: https://www.researchgate.net/publication/287200852_An_Approach_for_Classifying_ECG_Arrhythmia_Based_on_Features_Extracted_from_EMD_and_Wavelet_Packet_Domains
- [43] Pořízení EKG v aplikaci EKG na Apple Watch. In: *Apple* [online]. Dostupné z: <https://support.apple.com/cs-cz/HT208955>
- [44] KSIAŻCZYK, Marcin, Agnieszka DEBSKA-KOZŁOWSKA, Izabela WARCHOŁ a Andrzej LUBIŃSKI. Enhancing Healthcare Access—Smartphone Apps in Arrhythmia Screening: Viewpoint. *JMIR mHealth and uHealth* [online]. 2021, **9**(8), e23425. ISSN 2291-5222. Dostupné z: doi:10.2196/23425
- [45] DOSTÁL Jiří. *Biochemie: pro posluchače bakalářských oborů, skripta*. Brno: Masarykova univerzita, 2009.
- [46] SWAPNA, G., K. P. SOMAN a R. VINAYAKUMAR. Diabetes Detection Using ECG Signals: An Overview. In: DASH, Sujata, Biswa Ranjan ACHARYA, Mamta MITTAL, Ajith ABRAHAM a Arpad KELEMEN, ed. *Deep Learning Techniques for Biomedical and Health Informatics* [online]. Cham: Springer International Publishing, 2020, 2020-11-15, s. 299-327 [cit. 2023-05-19]. Studies in Big Data. ISBN 978-3-030-33965-4. Dostupné z: doi:10.1007/978-3-030-33966-1_14
- [47] OLOKOBA, Abdulfatai B., Olusegun A. OBATERU a Lateefat B. OLOKOBA. Type 2 Diabetes Mellitus: A Review of Current Trends. *Oman Medical Journal* [online]. 2012, **27**(4), 269–273. ISSN 1999768X. Dostupné z: doi:10.5001/omj.2012.68
- [48] MOTTL, Roman. Diabetes mellitus s vaskulárními komplikacemi. In: *Interní medicína pro praxi 2001 / 12* [online]. Dostupné z: <https://www.internimedicina.cz/pdfs/int/2001/12/05.pdf>
- [49] KUDLOVÁ, Pavla a Rudolf CHLUP. Selfmonitoring u osob s diabetem. In: *Interní medicína pro praxi 2006 / 12* [online]. Dostupné z: <https://www.internimedicina.cz/pdfs/int/2006/12/06.pdf>

- [50] EYTH, Emily a Roopa NAIK. *Hemoglobin A1C*. 2022. In: StatPearls [online]. Dostupné z: <https://www.ncbi.nlm.nih.gov/books/NBK549816/>
- [51] FUNTANILLA, Vienica D, Tina CALIENDO a Olga HILAS. Continuous Glucose Monitoring: A Review of Available Systems. *P & T: a peer-reviewed journal for formulary management*. 2019, **44**(9), 550–553. ISSN 1052-1372.
- [52] FreeStyle Libre. Dostupné z: <https://www.freestylelibre.cz/produkty>.
- [53] MOHANRAM, S. a A. SHIRLY EDWARD. Measurement of Blood Glucose using Non-Invasive Methods – A Review. *Journal of Physics: Conference Series* [online]. 2021, **1964**(6), 062022. ISSN 1742-6588. Dostupné z: doi:10.1088/1742-6596/1964/6/062022
- [54] SEYD, Patrick, V. I. Thajudin AHAMED, Jeevamma JACOB a K. PAULJOSEPH. Time and Frequency Domain Analysis of Heart Rate Variability and their Correlations in Diabetes Mellitus. *International Journal of Medical and Health Sciences*. 2008, **2**, 85–88.
- [55] SCHROEDER, Emily B., Lloyd E. CHAMBLESS, Duanping LIAO, Ronald J. PRINEAS, Gregory W. EVANS, Wayne D. ROSAMOND a Gerardo HEISS. Diabetes, Glucose, Insulin, and Heart Rate Variability. *Diabetes Care* [online]. 2005, **28**(3), 668–674. ISSN 0149-5992. Dostupné z: doi:10.2337/diacare.28.3.668
- [56] JIAN, Lee Wei a Teik-Cheng LIM. Automated Detection of Diabetes by Means of Higher Order Spectral Features Obtained from Heart Rate Signals. *Journal of Medical Imaging and Health Informatics* [online]. 2013, **3**(3), 440–447. ISSN 21567018. Dostupné z: doi:10.1166/jmih.2013.1178
- [57] ALLEN, John. Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement* [online]. 2007, **28**(3), R1–R39. ISSN 0967-3334. Dostupné z: doi:10.1088/0967-3334/28/3/R01
- [58] PARK, Junyung, Hyeon Seok SEOK, Sang-Su KIM a Hangsik SHIN. Photoplethysmogram Analysis and Applications: An Integrative Review. *Frontiers in Physiology* [online]. 2022, **12**. ISSN 1664-042X. Dostupné z: doi:10.3389/fphys.2021.808451
- [59] SUSANA, Ernia, Kalamullah RAMLI, Hendri MURFI a Nursama Heru APRIANTORO. Non-Invasive Classification of Blood Glucose Level for Early Detection Diabetes Based on Photoplethysmography Signal. *Information* [online]. 2022, **13**(2), 59. ISSN 2078-2489. Dostupné z: doi:10.3390/info13020059
- [60] SRINIVASAN, Venkatesh Bharadwaj a Foroohar FOROOZAN. Deep Learning based non-invasive diabetes predictor using Photoplethysmography signals. In: *2021 29th European Signal Processing Conference (EUSIPCO)* [online]. IEEE, 2021, s. 1256–1260. ISBN 978-9-0827-9706-0. Dostupné z: doi:10.23919/EUSIPCO54536.2021.9616351

- [61] SEN GUPTA, Shantanu, Tae-Ho KWON, Shifat HOSSAIN a Ki-Doo KIM. Towards non-invasive blood glucose measurement using machine learning: An all-purpose PPG system design. *Biomedical Signal Processing and Control* [online]. 2021, **68**, 102706. ISSN 17468094. Dostupné z: doi:10.1016/j.bspc.2021.102706
- [62] E4 wristband technical specifications. In: *Empatica Support* [online]. Dostupné z: <https://support.empatica.com/hc/en-us/articles/202581999-E4-wristband-technical-specifications>
- [63] *E4 data - BVP expected signal* [online]. Dostupné z: <https://support.empatica.com/hc/en-us/articles/360029719792-E4-data-BVP-expected-signal>
- [64] VARGOVÁ, Enikö. *Stanovení kvality a odhad tepové frekvence ze signálu PPG* [online]. Brno, 2021. Bakalářská práce. Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, Ústav biomedicínského inženýrství. Vedoucí práce Andrea Němcová.
- [65] MISHRA, Prabhaker, ChandraM PANDEY, Uttam SINGH, Anshul GUPTA, Chinmoy SAHU a Amit KESHRI. Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia* [online]. 2019, **22**(1), 67. ISSN 0971-9784. Dostupné z: doi:10.4103/aca.ACA_157_18
- [66] SUNDJAJA, Joshua Henrina, Rijen SHRESTHA a Kewal KRISHAN. *McNemar And Mann-Whitney U Tests*. 2022. In: StatPearls [online]. Dostupné z: <https://pubmed.ncbi.nlm.nih.gov/32809534/>
- [67] KALINA, Jan a Anna SCHLENKER. A Robust Supervised Variable Selection for Noisy High-Dimensional Data. *BioMed Research International* [online]. 2015, **2015**, 1–10. ISSN 2314-6133. Dostupné z: doi:10.1155/2015/320385
- [68] HANCHUAN PENG, FUHUI LONG a C. DING. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online]. 2005, **27**(8), 1226–1238. ISSN 0162-8828. Dostupné z: doi:10.1109/TPAMI.2005.159
- [69] PAVLÍK, Jiří. *Aplikovaná statistika*. Praha: Vysoká škola chemicko-technologická, 2005. ISBN 80-7080-569-2.
- [70] ELGENDI, Mohamed. Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering* [online]. 2016, **3**(4), 21. ISSN 2306-5354. Dostupné z: doi:10.3390/bioengineering3040021
- [71] ORPHANIDOU, Christina. *Signal Quality Assessment in Physiological Monitoring* [online]. Cham: Springer International Publishing, 2018. ISBN 978-3-319-68414-7. Dostupné z: doi:10.1007/978-3-319-68415-4

- [72] BENICHO, Thomas, Bruno PEREIRA, Martial MERMILLOD, Igor TAUVERON, Daniela PFABIGAN, Salwan MAQDASY a Frédéric DUTHEIL. Heart rate variability in type 2 diabetes mellitus: A systematic review and meta-analysis. *PLOS ONE* [online]. 2018, **13**(4), e0195166. ISSN 1932-6203. Dostupné z: doi:10.1371/journal.pone.0195166
- [73] RIBEIRO, Ícaro J.S., Rafael PEREIRA, Paulo F. VALENÇA NETO, Ivna v. FREIRE, Cezar A. CASOTTI a Mitermayer Galvão dos REIS. Relationship between diabetes mellitus and heart rate variability in community-dwelling elders. *Medicina* [online]. 2017, **53**(6), 375–379. ISSN 1010660X. Dostupné z: doi:10.1016/j.medici.2017.12.001
- [74] SIRISHA, Allampalli, Yerrabelli DHANALAKSHMI, Gopal KRUSHNA PAL, Mukta WYAWAHARE a Shashi AHUJA. Delayed Visual Evoked Response is Linked to Vascular Endothelial Growth Factor and Autonomic Dysfunction in Type 2 Diabetes Mellitus. *Journal of Diabetes & Metabolism* [online]. 2019, **10**(9). Dostupné z: doi:10.35248/2155-6156.19.10.833
- [75] ZAYAS-ARRABAL, Julian, Amaia ALQUIZA, Erkan TUNCAY, Belma TURAN, Monica GALLEGO a Oscar CASIS. Molecular and Electrophysiological Role of Diabetes-Associated Circulating Inflammatory Factors in Cardiac Arrhythmia Remodeling in a Metabolic-Induced Model of Type 2 Diabetic Rat. *International Journal of Molecular Sciences* [online]. 2021, **22**(13), 6827. ISSN 1422-0067. Dostupné z: doi:10.3390/ijms22136827
- [76] CURTIS, Alexander, Tanya SMITH, Bulat ZIGANSHIN a John ELEFTERIADES. The Mystery of the Z-Score. *AORTA* [online]. 2016, **04**(04), 124–130. ISSN 2325-4637. Dostupné z: doi:10.12945/j.aorta.2016.16.014
- [77] ZHANG, Zhongheng. Introduction to machine learning: k-nearest neighbors. *Annals of Translational Medicine* [online]. 2016, **4**(11), 218–218. ISSN 23055839. Dostupné z: doi:10.21037/atm.2016.03.37
- [78] NADKARNI, Prakash. Core Technologies: Data Mining and “Big Data”. In: *Clinical Research Computing* [online]. Elsevier, 2016, s. 187–204. Dostupné z: doi:10.1016/B978-0-12-803130-8.00010-5
- [79] DREISEITL, Stephan a Lucila OHNO-MACHADO. Logistic regression and artificial neural network classification models: a methodology review. *Journal of Biomedical Informatics* [online]. 2002, **35**(5–6), 352–359. ISSN 15320464. Dostupné z: doi:10.1016/S1532-0464(03)00034-0
- [80] MAMMONE, Alessia, Marco TURCHI a Nello CRISTIANINI. Support vector machines. *WIREs Computational Statistics* [online]. 2009, **1**(3), 283–289. ISSN 1939-5108. Dostupné z: doi:10.1002/wics.49
- [81] RODRIGUEZ-GALIANO, V.F., B. GHIMIRE, J. ROGAN, M. CHICA-OLMO a J.P. RIGOL-SANCHEZ. An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and*

- Remote Sensing* [online]. 2012, **67**, 93–104. ISSN 09242716. Dostupné z: doi:10.1016/j.isprsjprs.2011.11.002
- [82] WRIGHT, Marvin N. a Inke R. KÖNIG. Splitting on categorical predictors in random forests. *PeerJ* [online]. 2019, **7**, e6339. ISSN 2167-8359. Dostupné z: doi:10.7717/peerj.6339
- [83] HAJIAN-TILAKI, Karimollah. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian journal of internal medicine*. 2013, **4**(2), 627–35. ISSN 2008-6164.
- [84] ZHANG, Gaobo, Zhen MEI, Yuan ZHANG, Xuesheng MA, Benny LO, Dongyi CHEN a Yuanting ZHANG. A Noninvasive Blood Glucose Monitoring System Based on Smartphone PPG Signal Processing and Machine Learning. *IEEE Transactions on Industrial Informatics* [online]. 2020, **16**(11), 7209–7218. ISSN 1551-3203. Dostupné z: doi:10.1109/TII.2020.2975222
- [85] ZHANG, Yuwei, Yuan ZHANG, Sarah A. SIDDIQUI a Anton KOS. Non-invasive Blood-Glucose Estimation Using Smartphone PPG Signals and Subspace KNN Classifier. *Elektrotehniski Vestnik/Electrotechnical Review*. 2019, **86**, 68–74.
- [86] GUZMAN, Luis, Andres Marcelo GarzaFox CAZARES a Antonio MARTINEZ-TORTEYA. Model for Glycemic Level Detection using Heart Rate Variability in a Mexican Sample. In: *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)* [online]. B.m.: IEEE, 2021, s. 505–510. ISBN 978-1-7281-4245-6. Dostupné z: doi:10.1109/IECBES48179.2021.9398841
- [87] SHAMIM, Mohammed Zubair M., Sattam ALOTAIBI, Hany S. HUSSEIN, Mohammed FARRAG a Mohammad SHIBLEE. Diagnostic Accuracy of Smartphone-Connected Electrophysiological Biosensors for Prediction of Blood Glucose Level in a Type-2 Diabetic Patient Using Machine Learning: A Pilot Study. *IEEE Embedded Systems Letters* [online]. 2022, **14**(1), 27–30. ISSN 1943-0663. Dostupné z: doi:10.1109/LES.2021.3096717
- [88] NIE, Zihan, Meng RONG a Kaiyang LI. Blood glucose prediction based on imaging photoplethysmography in combination with Machine learning. *Biomedical Signal Processing and Control* [online]. 2023, **79**, 104179. ISSN 17468094. Dostupné z: doi:10.1016/j.bspc.2022.104179
- [89] MANURUNG, Betty Elisabeth, Hugi Reyhandani MUNGGARAN, Galih Fajar RAMADHAN a Allya Paramita KOESOEMA. Non-Invasive Blood Glucose Monitoring using Near-Infrared Spectroscopy based on Internet of Things using Machine Learning. In: *2019 IEEE R10 Humanitarian Technology Conference (R10-HTC)(47129)* [online]. IEEE, 2019, s. 5–11. ISBN 978-1-7281-0834-6. Dostupné z: doi:10.1109/R10-HTC47129.2019.9042479

SEZNAM SYMBOLŮ A ZKRATEK

AC	pulzující složka
Acc	přesnost
ANS	autonomní nervová soustava
AUC	plocha pod křivkou (area under the curve)
BSA	plocha povrchu těla
BVP	puls objemu krve (blood volume pulse)
CAN	kardiovaskulární autonomní neuropatie
CNS	centrální nervová soustava
DC	nepulzující složka
DF	dechová frekvence
DFA	metodu analýzy fluktuací
DM	diabetes mellitus
DP	diplomová práce
DTK	diastolický krevní tlak
EBTA	ensemble bagged trees algoritmus
EDA	elektrodermální aktivita
EKG	elektrokardiografie
EMG	elektromyografie
ET	ejekční čas
FEKT	Fakulta elektrotechniky a komunikačních technologií
FKG	fonokardiografie
FN	falešně negativní (false negative)
FP	falešně pozitivní (false positive)
HbAc1	glykovaný hemoglobin
IB	In Bag
IDF	International Diabetes Federation
I_{max}	maximální intenzita
I_{min}	minimální intenzita
KNN	k-nejbližších sousedů
KT	krevní tlak
LED	elektroluminiscenční dioda (light emitting diode)
LR	logistická regrese
MAE	střední absolutní chyba (mean absolute error)
mRMR	maximální redundance minimální relevance
NO	oxid dusnatý
oGTT	orální glukózový toleranční test
OOB	Out Of Bag
PCR	regrese hlavních komponent (principal component regression)

PLS	částečná regrese nejmenších čtverců
PPG	fotopletysmografie
PPV	pozitivní prediktivní hodnota
PTK	pulzní krevní tlak
PTT	Pulse Transit Time
rbf	radiální bázové jádro
RF	náhodný les (random forest)
RFR	regrese s využitím náhodného lesu (random forest regression)
RGB	červená-zelená-modrá (red-green-blue)
ROC	ROC křivka (receiver operating characteristic curve)
s	sekunda
Se	senzitivita
SE	Shannonova entropie
Sp	specificita
SpO ₂	saturace krve kyslíkem
std	směrodatná odchylka
STK	systolický krevní tlak
SV	tepový objem
SVM	metoda podpůrných vektorů
SVR	metoda podpůrných vektorů pro regresi
SWT	stacionární vlnková transformace
TF	tepová frekvence
TKEO	Teager-Kaiserův operátor (TKEO)
TN	skutečně negativní (true negative)
TP	skutečně pozitivní (true positive)
VTT	Vascular Transit Time
VUT	Vysoké učení technické v Brně
WHD	nositelná elektronika (Wearable Health Devices)