

Západočeská univerzita v Plzni
Fakulta aplikovaných věd
Katedra informatiky a výpočetní techniky

Diplomová práce

Detekce vybraných aktivit diabetického pacienta 1. typu

ZÁPADOČESKÁ UNIVERZITA V PLZNI

Fakulta aplikovaných věd

Akademický rok: 2020/2021

ZADÁNÍ DIPLOMOVÉ PRÁCE

(projektu, uměleckého díla, uměleckého výkonu)

Jméno a příjmení:	Bc. David PIVOVAR
Osobní číslo:	A19N0073P
Studijní program:	N3902 Inženýrská informatika
Studijní obor:	Medicínská informatika
Téma práce:	Detekce vybraných aktivit diabetického pacienta 1. typu
Zadávací katedra:	Katedra informatiky a výpočetní techniky

Zásady pro vypracování

1. Seznamte se s problematikou diabetu 1. typu a kontinuální monitorací glukózy, konkrétně u scénářů příjmu karbohydrátů a zvýšené fyzické aktivity.
2. Prostudujte existující způsoby detekce příjmu karbohydrátů a zvýšené fyzické aktivity.
3. Seznamte se se systémem SmartCGMS primárně vyvíjeným na Katedře informatiky a výpočetní techniky.
4. Ve formě plug-inu pro systém SmartCGMS implementujte detekci příjmu karbohydrátů.
5. Ověřte detekci příjmu karbohydrátů vůči referenčním hodnotám.
6. Implementujte detekci zvýšené fyzické aktivity jako další plug-in.
7. Zhodnoťte dosažené výsledky.

Rozsah diplomové práce: **doporuč. 50 s. původního textu**
Rozsah grafických prací: **dle potřeby**
Forma zpracování diplomové práce: **tištěná**

Seznam doporučené literatury:

dodá vedoucí diplomové práce

Vedoucí diplomové práce: **Doc. Ing. Tomáš Koutný, Ph.D.**
Katedra informatiky a výpočetní techniky

Datum zadání diplomové práce: **11. září 2020**
Termín odevzdání diplomové práce: **20. května 2021**

L.S.

Doc. Dr. Ing. Vlasta Radová
děkanka

Doc. Ing. Přemysl Brada, MSc., Ph.D.
vedoucí katedry

Prohlášení

Prohlašuji, že jsem diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů.

V Plzni dne 26. srpna 2021

Bc. David Pivovar

Abstract

This thesis deals with carbohydrate and physical activity detection of a type 1 diabetic patient. The aim of the thesis was to evaluate existing detection methods and implement detection as filters in the SmartCGMS application. Two methods were proposed and implemented for carbohydrate detection. The first method uses recurrent neural networks, while the second detects the edges of the interstitial glucose waveform measured by a continuous glucose monitoring sensor. Physical activity detection is based on heart rate and motion data.

Abstrakt

Tato diplomová práce se zabývá detekcí karbohydrátů a fyzické aktivity diabetického pacienta 1. typu. Cílem práce bylo zhodnotit existující metody detekce a implementovat vlastní řešení jako filtry do aplikace SmartCGMS. Pro detekci karbohydrátů byly navrženy a implementovány dvě metody. První využívá rekurentní neuronové sítě, druhá detekuje hrany průběhu intersticiální glukózy měřené senzorem kontinuální monitorace glukózy. Detekce fyzické aktivity je na základě hodnot srdečního tepu a pohybových dat.

Obsah

1	Úvod	7
2	Diabetes mellitus	8
2.1	Dělení diabetu	9
2.2	Léčba	10
2.2.1	Inzulinové režimy	12
2.2.2	Inzulinová pumpa	13
2.2.3	Příjem karbohydrátů	13
2.2.4	Fyzická aktivita	14
2.3	Monitorace glukózy	14
2.3.1	CGMS	15
3	Analýza metod detekce příjmu karbohydrátů	17
3.1	Bergmanův minimální model	17
3.2	Model-based metody	18
3.3	Data-driven metody	24
3.4	Metody využívající neuronové sítě	30
3.5	Porovnání jednotlivých metod	33
4	Analýza metod detekce fyzické aktivity	34
4.1	Metody	34
4.2	Porovnání metod	41
5	SmartCGMS	43
6	Návrh metody detekce karbohydrátů	46
6.1	Rekurentní neuronové sítě	46
6.1.1	Long short-term memory	47
6.1.2	Gated recurrent unit	48
6.1.3	Model	48
6.2	Lineární a kvadratická diskriminační analýza	49
6.2.1	Množina prvků	50
6.2.2	Časové okno	51
6.3	Detekce hran průběhu intersticiální glukózy	51
7	Návrh metody detekce fyzické aktivity	54

8 Implementace	55
8.1 Použité nástroje	55
8.2 Dataset BGLP	55
8.2.1 Analýza dat	57
8.3 Filtr pro SmartCGMS	58
8.3.1 Detekce příjmu karbohydrátů	59
8.3.2 Detekce fyzické aktivity	60
8.3.3 Možnosti rozšíření	60
9 Výsledky	61
9.1 Detekce karbohydrátů	61
9.2 Detekce fyzické aktivity	63
10 Závěr	67
Přehled zkratk	68
Literatura	69
Příloha A: Tabulky	74
Příloha B	80
Adresářová struktura	80
Knihovny, kompilace, spuštění	80
Nastavení filtrů	81
Python skripty	83

1 Úvod

Diabetes mellitus je rozšířené chronické metabolické onemocnění. Vyznačuje se zvýšenou koncentrací cukru v krvi (glykémie), která vniká z nedostatku inzulínu nebo rezistencí vůči jeho působení. Hlavní součástí léčby je monitorace koncentrace glukózy a podávání inzulínu. V dnešní době je možné sledovat vývoj glykémie díky systémům kontinuální monitorace glukózy (CGMS) a kontinuálnímu dávkování inzulínu díky inzulínovým pumpám. Integrace těchto dvou systémů umožňuje autonomní řízení dávkování inzulínu v závislosti na aktuální koncentraci glukózy.

Koncentraci glukózy v krvi ovlivňuje mnoho faktorů. Dva nejvýznamnější faktory jsou příjem karbohydrátů a fyzická aktivita. Tyto aktivity je třeba kompenzovat zvýšením či snížením množství podávaného inzulínu. V současné době aktivity do CGMS zadává pacient, který si zároveň následně upravuje dávku inzulínu. Monitorace těchto aktivit slouží pro správné nastavení léčby diabetologem. Jejich včasná detekce by umožnila přesnější monitoraci vývoje glykémie a lepší řízení autonomních systémů s minimálním zásahem pacienta.

Tato práce se zabývá možnostmi detekce příjmu karbohydrátů a fyzické aktivity z dat senzoru CGMS. Následně jsou metody implementovány do aplikace SmartCGMS vyvíjené na katedře Informatiky a výpočetní techniky na ZČU.

2 Diabetes mellitus

Diabetes mellitus je chronické metabolické onemocnění. Vyznačuje se zvýšenou hladinou cukru v krvi, která se nazývá hyperglykémie. V normálu je hladina glykémie (koncentrace glukózy v krvi) mezi 3,8 a 5,6 mmol/l. Po jídle může glykémie stoupnout, neměla by však přesáhnout 7,8 mmol/l. O diabetes se jedná pokud jsou hodnoty vyšší než 11,1 mmol/l.

Hyperglykémie vzniká z nedostatku inzulínu nebo rezistencí buněk vůči jeho působení. Inzulin je hormon, který se tvoří v beta buňkách Langerhansových ostrůvků ve tkáni slinivky břišní. Tento hormon slouží k transportu glukózy z krve do tkáně, kde je využita k tvorbě energie nebo se zde ukládá ve formě glykogenu (v játrech) a tuků. Snižuje tak hladinu cukru v krevním řečišti. Pokud je hladina cukru nízká, produkce inzulínu se okamžitě zastaví a hormony glukagon a adrenalin uvolní glukózu ze zásobních zdrojů pro zajištění fungování důležitých orgánů, především mozku, nervů a svalů.

Nedostatek inzulínu vede k nedostatečné utilizaci glukózy, která se hromadí v krvi. Dochází k poruše tvorby bílkovin, zvýšené tvorbě ketolátů, které se hromadí a způsobují překyselení organismu, a celkovému metabolickému rozvratu. Při hladině glykémie vyšší než 10-12 mmol/l začnou ledviny uvolňovat glukózu do moči spolu s dalšími látkami a vodou. To vše vyvolává dehydrataci, nevolnost, pokles krevního tlaku, poruchy vědomí.

Dlouhodobě přetrvávající hyperglykémie vede k poškození orgánů, především cév, nervového systému, ledvin a očí. Poškozením velkých cév může dojít k srdečnímu infarktu, mozkové příhodě nebo uzavření tepen v oblasti dolních končetin (ischemická choroba dolních končetin). U diabetických pacientů mají tato onemocnění komplikovanější průběh a vyšší úmrtnost. Nervové poškození je označováno jako diabetická neuropatie. Nejčastěji se objevuje v oblasti dolních končetin. končetina ztrácí citlivost, kůže vysychá a tvoří se defekt do kterého se snadno dostane infekce. Pokud není včas léčen, rozvíjí se syndrom diabetické nohy, který může vést k amputaci končetiny. poškození ledvin může vést k dialýze až k transplantaci ledvin. Poškození očí je příčinou vzniku šedého zákalu.

Bez léčby inzulínem může toto onemocnění vést k smrti. [23, 32]

2.1 Dělení diabetu

Diabetes se dělí do čtyř základních skupin: diabetes 1. typu, diabetes 2. typu, gestační (těhotenský) diabetes, jiné specifické typy diabetu.

Diabetes mellitus 1. typu je autoimunitní onemocnění, kdy vlastní imunitní systém zničí beta buňky Langerhansových ostrůvků produkující inzulín. Langerhansových ostrůvků je ve slinivce kolem jednoho milionu a každý obsahuje zhruba 2000 beta buněk. Přesný počet se u každého liší. Při zničení 75-85 % beta buněk nastává absolutní nedostatek inzulínu a objevují se zvýšené hodnoty glykémie.

Diabetes 1. typu se nejčastěji projeví již v dětském věku, ale může dlouho probíhat v latentní formě a projevit se až v pozdějším věku (LADA - latentní autoimunitní diabetes u dospělých). O diabetes se jedná pokud hodnota glykémie přesahuje 7 mmol/l nalačno. Po podání 75 g cukru pak glykémie přesahuje 11,1 mmol/l. Zda se jedná o diabetes 1. typu a ne diabetes 2. typu či jiné onemocnění způsobující vyšší hodnoty glykémie se určí stanovením hodnot glykovaného hemoglobinu a detekcí autoprotilátek, které vznikají při poškození beta buněk. Příčina vzniku diabetu 1. typu není známá a proto není známa ani účinná prevence. Vlohy jsou často dědičné.

Nemoc se projevuje častým močením, které je v důsledku neschopnosti ledvin absorbovat přebytečnou glukózu. Z toho je pacient dehydratován, trpí žízní, nechutenstvím a celkovou slabostí. V důsledku hromadění ketolátek dochází k překyselení organismu, což může vést k poruchám vědomí.

Diabetes mellitus 2. typu je metabolickou poruchou, kdy jsou buňky rezistentní vůči vlastnímu inzulínu, tj. buňky nejsou schopny vychytat inzulín v krevním oběhu, který by použili ke zpracování glukózy a úpravě hladiny glykémie. V počátcích beta buňky slinivky břišní reagují zvýšenou produkcí inzulínu (bezpříznaková inzulínová rezistence). Postupem času ale nejsou schopny dodávat potřebné množství inzulínu, nastává relativní nedostatek inzulínu a vzniká prediabetes a diabetes 2. typu. Příznaky jsou únava, žízeň, časté močení, pocit hladovění, hubnutí, infekce, špatné hojení a zhoršení zraku. Diabetes mellitus 2. typu vzniká v dospělosti na základě genetických předpokladů (inzulinová rezistence, nízká produkce inzulínu), rizikových faktorů a špatného životního stylu. Mezi ty patří nedostatek pohybu, nezdravá strava, nadváha a obezita (90 % diabetiků 2. typu trpí nadváhou nebo je obézních), kouření, vysoká hladina cholesterolu a vysoký krevní tlak (hypertenze).

Gestační diabetes se objevuje v druhé polovině těhotenství a končí po

porodu. Postihuje 17 % těhotných žen, které k ní mají vrozenou dispozici. Neléčený diabetes má negativní dopad na vývoj plodu, který se musí vyrovnat se zvýšeným přísunem glukózy. Plod, který má produkci inzulínu v pořádku, pak výrazně rychleji roste. To může vést k předčasnému porodu, vrozeným vývojovým vadám, dechovým obtížím, poruchám srdečního rytmu, poruchám vývoje mozku, sklon k obezitě a rozvoje diabetu 2. typu.

Jiné specifické typy diabetu se označují jako sekundární diabetes, protože k rozvoji diabetu dochází v důsledku jiného onemocnění (např. zánět slinivky břišní) nebo v důsledku genetické poruchy (takzvané MODY).

V České republice je evidováno více než 900 000 pacientů s diabetem, kdy 92 % připadá na diabetes 2. typu, 7 % na diabetes 1. typu. Ročně přibude 60 tisíc nových diabetických pacientů a 22 tisíc umírá. Náklady na léčbu jednoho pacienta jsou okolo 26 000 Kč za rok. Ročně se tak vydá na léčbu diabetu 23 miliard Kč. [23, 32]

Celosvětově je Světovou zdravotnickou organizací evidováno více než 422 milionů lidí s diabetem [37].

V této práci se budu zabývat detekcí karbohydrátů a fyzické aktivity u pacientů s diabetem 1. typu.

2.2 Léčba

Diabetes mellitus 1. typu je onemocnění z nedostatečné produkce inzulínu slinivkou břišní a vždy se léčí podáváním inzulínu. Způsob léčby a množství podávaného inzulínu je rozebráno v kapitole 2.2.1, způsoby podání v kapitole 2.2.2. Další možností pro pacienty diabetu 1. typu je transplantace slinivky břišní nebo transplantace samotných Langerhansových ostrůvků.

Při léčbě inzulínem se musí dbát na to, aby nenastala hypoglykémie. Hypoglykémie je stav, kdy dochází k poklesu cukru v krvi pod 3,9 mmol/l. Příznaky jsou pocení, třes, bušení srdce, slabost. Při těžké hypoglykémii dochází k poruchám vědomí a je nutná pomoc jiné osoby. Může nastat i smrt. K hypoglykémii často dochází při podání příliš velké dávky inzulínu snahou regulovat hladinu glukózy co nejvíce k normálním hodnotám. Z toho důvodu by pacient měl mít u sebe menší množství rychlých cukrů pro doplnění při pozorování prvních příznaků.

Zvýšené riziko těžké hypoglykémie je při poruše tvorby adrenalinu v nadledvinách. Adrenalin se uvolňuje při nízké hladině cukru v krvi a je zod-

povědný za příznaky hypoglykémie. Pacient pak není schopen rozpoznat rozvíjející se hypoglykémii. Tato porucha se označuje jako syndrom porušeného vnímání hypoglykémie.

Diabetes mellitus 2. typu se nejprve léčí úpravou životosprávy a kombinací perorálně podávaných antidiabetik (léky zvyšující produkci inzulinu a/nebo citlivost buněk k inzulinu). Pokud se těmito prostředky nedaří regulovat hladina cukru v krvi, přidá se k podávaným antidiabetikám léčba inzulinem.

Společnými režimovými opatřeními pro oba typy diabetu je monitorace glukózy (viz kapitola 2.3), dieta a fyzická aktivita.

Glykemie je závislá na jídle, druhu a frekvenci stravy, proto patří dieta mezi základní opatření při léčbě cukrovky. Pacient by měl znát složení jídla, především množství zkonsumovaných sacharidů. Strava by měla být vyvážená a v pravidelných intervalech.

Pohyb je důležitý protože napomáhá účinnosti inzulinu. Při pohybu se také využije více glukózy na energii, je prevencí proti nadváze a obezitě a celkově prospívá organismu. Fyzická aktivita tak výrazně snižuje koncentraci glukózy v krvi. Ideální hodnota glykémie před fyzickou aktivitou je 6-7 mmol/l. Při nižších hodnotách hrozí riziko hypoglykémie. Naopak při hodnotách >15 mmol/l se postupuje jako při hyperglykémii.

Cílem léčby je udržení doporučených hodnot vybraných ukazatelů popsaných v tabulce 2.1. V závorkách jsou uvedeny přijatelné hodnoty rizikových pacientů. Pakliže se pacientovi dlouhodobě daří udržovat tyto doporučené hodnoty, sniží riziko komplikací spojených s diabetem. [23]

Tabulka 2.1: Doporučené hodnoty vybraných ukazatelů

Ukazatel	Doporučené hodnoty
glykovaný hemoglobin	pod 45 (60) mmol/l
glykémie naměřená před jídlem	4-6 (pod 8) mmol/l
glykémie naměřená po jídlem	5-7,5 (pod 9) mmol/l
krevní tlak	pod 130/80 (pod 140/90) mmHg
celkový cholesterol	pod 4,5 mmol/l
LDL-cholesterol	pod 2,5 mmol/l
HDL-cholesterol (muži/ženy)	nad 1/1,2 mmol/l
Celková denní dávka inzulinu	pod 0,6 j./kg

Zdroj: www.cukrovka.cz [22]

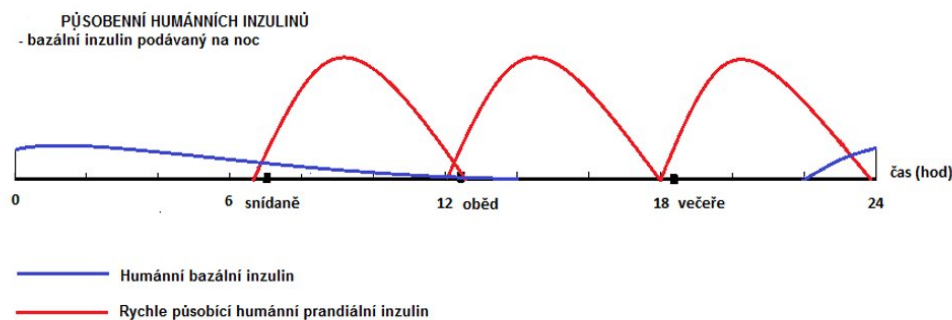
2.2.1 Inzulinové režimy

Schéma léčby inzulinem se volí individuálně. Cílem je korigovat deficit inzulinu a co nejpřesněji simulovat fyziologickou sekreci inzulinu. Pro každého diabetického pacienta je toto jiné z důvodu vlastního metabolismu, životního stylu a aktivit.

Zpravidla dělíme dávkování inzulinu na bazální hladinu inzulinu a bolusy. Bazální hladina inzulinu zajišťuje správnou hladinu glykémie mezi 5-9 mmol/l během dne a v noci. Dávka podaného inzulinu musí být nastavena tak, aby nedošlo k hypoglykémii. Inzulin se podává buď v jedné dávce na noc, nebo ve dvou menších dávkách ráno a večer. Účinnost podání bazálního inzulinu v jedné dávce večer je vidět na obrázku 2.1 modře. Používají se středně nebo dlouho působící inzuliny.

Bolus slouží k aktuálnímu snížení hladiny glykémie. Používá se rychle působící prandiální inzulin, který začíná působit v krátkém čase a vrchol působení je mezi 1-2 hodinami. Bolus inzulinu se aplikuje zpravidla před jídlem, nebo pokud se naměří zvýšená hladina glykémie. Podání bolusu je znázorněno na obrázku 2.1 červeně. Odhad poklesu glykémie po podání jedné jednotky inzulinu se vypočte 83 děleno celkovou denní dávkou inzulinu. [19, 26]

Obrázek 2.1: Působení inzulinu



Zdroj: www.cukrovka.cz [22]

Další možností je postupné dávkování inzulinu v malých dávkách během celého dne pomocí inzulinové pumpy (viz kapitola 2.2.2).

2.2.2 Inzulinová pumpa

Inzulin se podává injekčně do podkoží. Jednou z možností jsou inzulinové pera (aktuálně nejrozšířenější způsob podání inzulinu). Jedná se o lepší variantu klasické injekční stříkačky schopné dávkovat inzulin s přesností na 0,5 jednotky inzulinu.

Další možností aplikace inzulinu je inzulinová pumpa. Ta umožňuje kontinuální podání bazální dávky rychle působícího inzulinu. Simuluje tak reálnou funkci zdravé slinivky, která by za normálních okolností produkovala malé množství inzulinu pro udržení glykémie. Příjem jídla se vykryje nastavením požadovaného bolusu.

Na některých pumpách lze nastavit více bazálních dávek inzulinu přizpůsobených režimu daného dne (například víkendové aktivity nebo sport). Některé lze také propojit se systémem kontinuální monitorace glukózy, viz kapitola 2.3. [21, 26]

2.2.3 Příjem karbohydrátů

Karbohydráty (jiným názvem sacharidy či cukry) se po přijetí štěpí na glukózu, která je hlavním zdrojem energie pro lidské buňky. Při nedostatečném množství inzulinu nejsou buňky schopny glukózu zpracovat a ta se hromadí v krvi. Koncentrace glukózy v krvi se označuje jako glykémie.

Diabetický pacient musí hlídat množství přijatých karbohydrátů, které je nutné úměrně kompenzovat. Přijaté karbohydráty se projevují zvýšením koncentrace glukózy v krvi po dobu několika hodin (dle množství přijatých karbohydrátů). Pacient by měl rozdělit jídlo do šesti malých porcí v průběhu celého dne, z čehož 45 % by měly být sacharidy. To spolu se správným dávkováním inzulinu pomáhá regulaci glykémie. Různé potraviny mají rozdílný glykemický index (jak rychle zvyšují hladinu glukózy v krvi). Je doporučena konzumace potravin s nízkým glykemickým indexem. Rychlé cukry s vysokým glykemickým indexem by se měly konzumovat v případě poklesu koncentrace glukózy v krvi (například příliš velká dávka inzulinu nebo fyzická aktivita).

Před jídlem se podává rychle působící prandiální inzulin, bolus. Inzulin se podává 30 minut před jídlem a vrchol působení je mezi 1-2 hodinami. Jeho množství by mělo odpovídat velikosti jídla. Pro výpočet množství karbohydrátů pokrytých 1 jednotkou inzulinu aplikovaného před jídlem vydělíme číslo 500 celkovou denní dávkou inzulinu. Pro výpočet bolusu pak vydělíme

množství přijatých sacharidů vypočtenou hodnotou.

2.2.4 Fyzická aktivita

Během fyzické aktivity svalové buňky spotřebovávají glukózu a dochází tak k poklesu glykémie. Při fyzické zátěži je nutné častěji monitorovat hladinu glykémie, snížit množství dávkovaného inzulínu a případně podat sacharidy. Glykémie před cvičením by měla být mezi 6-7 mmol/l. Při nižších hodnotách je riziko hypoglykémie.

Hodnoty glykémie se měří před, během i po fyzické zátěži. Ideální je pro toto měření kontinuální monitorace glukózy (viz kapitola 2.3.1). Měření glykémie po cvičení je důležité kvůli riziku vzniku takzvané pozdní hypoglykémie.

Před fyzickou zátěží je potřeba snížit množství podávaného inzulínu. V případě jednorázové aktivity se snižuje bolus inzulínu podávaný před jídlem o 20-50 % dle náročnosti. Při dlouhodobější činnosti se snižuje množství bazálního inzulínu o 20-30 %. Před a během zátěže je nutné pravidelně doplňovat sacharidy pro redukci rizika hypoglykémie.

2.3 Monitorace glukózy

Monitoring je důležitý pro správné určení léčby diabetologem.

Součástí monitorace je měření koncentrace glukózy a ketolátek v krvi a moči. Dále se kontroluje krevní tlak, hmotnost (BMI). Zaznamenávat by se měly denní dávky inzulínu, příjem karbohydrátů, hypoglykémie a hyperglykémie a situace vyžadující úpravu dávkování inzulínu, jako je zvýšená fyzická aktivita nebo nemoc. Při pravidelné kontrole u lékaře se vyšetřuje množství glykovaného hemoglobinu, cholesterolu, činnost ledvin a neuropatie.

Frekvence monitoringu je individuální a závislá na mnoha faktorech a potřebách daného pacienta. Selfmonitoring by se měl provádět denně před aplikací inzulínu, zhruba 3-4x denně, v noci v případě rizika hypoglykémie a v době nástupu dawn fenoménu (ranní hyperglykémie). Výsledek běžného denního měření se nazývá malý glykemický profil. Častější monitoring je třeba v případě zvláštních situací vyžadujících úpravu dávkování inzulínu a v těhotenství. Takové měření se nazývá malý glykemický profil.

Nejčastěji se glykémie měří glukometrem z kapky krve nanesené na di-

agnostický proužek. Tato metoda, která zahrnuje vpich do konečku prstu může být pro pacienta bolestivá a odradit ho od častějších měření. [20]

2.3.1 CGMS

Alternativou k jednorázovým měřením je kontinuální monitorace glukózy (CGMS - Continuous Glucose Monitoring System). Pro měření se zavede elektrochemický glukózový senzor do podkoží, kde se měří koncentrace glukózy v plazmě v intersticiální (mezibuněčné) tekutině. Tento systém umožňuje měření hladiny glukózy po celý den v minutových nebo pětiminutových intervalech (na českém trhu jsou dostupné přístroje měřící v pětiminutových intervalech). To umožňuje sledovat vývoj glykémie během dne a stanovit tak glykemický profil pacienta.

Hodnoty ze senzoru a hodnoty naměřené z krve se mohou lišit. Protože měření probíhá v intersticiální tekutině, kam se glukóza dostává z krve s menším zpožděním, jsou naměřené hodnoty také se zpožděním vůči reálné hodnotě v krvi. V některých případech jsou koncentrace v intersticiální tekutině nižší než v krvi, zejména v noci. Senzor se musí pro správnou funkci denně kalibrovat (alespoň 2x) hodnotami naměřenými z krve při stabilní glykémii. [21]

Přístroje disponují alarmem signalizujícím vzestup či pokles glykémie. Do CGMS uživatel ručně zadává množství přijatých karbohydrátů z jídla a fyzická aktivita. Z dat CGMS pak lze lépe určit vývoj glykémie v závislosti na daných aktivitách.

Některé přístroje umožňují integraci s kompatibilní inzulinovou pumpou a zadání požadované bazální hladiny inzulinu a bolusů. Aktuální hybridní systémy jsou schopny upravovat hladinu bazálu podle naměřených hodnot. Inzulinová pumpa tak dává potřebné množství inzulinu bez nutnosti častého zásahu pacienta. Takové systémy se někdy označují jako umělá slinivka. Příkladem je systém OpenAPS, který ale není oficiálně schválený jako zdravotnický prostředek a tudíž není zaručena jeho správná funkcionality. Dalším příkladem aplikace pro řízení inzulinové pumpy na základě dat CGMS je SmartCGMS, který je vyvíjen na katedře Informatiky a výpočetní techniky na ZČU (viz kapitola 5). Plně autonomní systém zatím neexistuje z důvodu, že hladinu glukózy ovlivňuje mnoho faktorů, které přístroje na trhu nejsou schopny spolehlivě rozeznat. Jedná se především o rychlé zvýšení glykémie po jídle a podání odpovídajícího bolusu.

Rizikem autonomních systémů je podání příliš velkého množství inzu-

linu a stavu hypoglykémie, která je mnohonásobně horší než riziko hyperglykémie. Řešením by mohl být dávkovač glukagonu, který by při poklesu glykémie hladinu opět vyrovnal.

3 Analýza metod detekce příjmu karbohydrátů

Detekce příjmu karbohydrátů je důležitou součástí autonomních systémů CGMS a inzulinové pumpy. Při příjmu karbohydrátů se zvyšuje koncentrace glukózy v krvi, kterou je nutno korigovat bolusem inzulinu, aby se předešlo stavu hyperglykémie. Tato dávka inzulinu musí odpovídat množství přijatých karbohydrátů. Příliš velká dávka může vést k hypoglykémii.

Metody detekce karbohydrátů mohou být model-based nebo data-driven.

Většina studií je založená na implementaci fyzického modelu (většinou Bergmanův minimální model) a aplikaci predikčního algoritmu, jako je například Kalmanův filtr pro predikci jednotlivých stavů (glykémie a karbohydráty). Detekce karbohydrátů je pak porovnáním pozorovaných stavů a modelu vůči definovanému thresholdu, výpočtu cross-covariance nebo aplikováním rozhodovacích pravidel.

U data-driven metod se vychází z naměřených dat. Extrakce vlastností je u těchto metod kvantitativní nebo kvalitativní. Kvantitativní metoda je například analýza hlavních komponent. V kvalitativních modelech jsou časová data převedena na sekvenci kvalitativních proměnných k vytvoření kvalitativní reprezentace dat. Data-driven metody jsou méně závislé na přesnosti fyzického modelu, ale je potřeba pro jejich natrénování velkého množství vzorků. Mezi data-driven metody patří i neuronové sítě.

3.1 Bergmanův minimální model

Bergmanův minimální model [34] je nelineární dynamický model koncentrace glukózy v plazmě. Tento model popisuje vývoj hladiny glukózy v plazmě na základě koncentrace a účinnosti inzulinu a přijatých karbohydrátů. Model určuje sada diferenciálních rovnic prvního řádu:

$$\begin{aligned}\frac{dF(t)}{dt} &= -p_1 G(t) - p_4 I_{eff}(t) G(t) + p_1 G_b + R_a(t) \\ \frac{dI_{eff}(t)}{dt} &= -p_2 I_{eff}(t) + p_3 I_p(t)\end{aligned}$$

kde G_b je koncentrace glukózy v plazmě, I_p je koncentrace inzulinu v plazmě, I_{eff} je koncentrace efektivního inzulinu, p_1, p_2, p_3, p_4 jsou parametry a $R_a(t)$

je míra výskytu glukózy definovaná jako:

$$R_a(t) = \frac{C(t)}{V\tau^2}te^{-\frac{t}{\tau}}$$

kde $C(t)$ je množství přijatých karbohydrátů, V je distribuční objem a τ je vrchol absorpce karbohydrátů [35].

3.2 Model-based metody

3.2.1 Real-time insulin bolusing for unannounced meals with artificial pancreas

Pro detekční algoritmus této metody použili Turksoy et al. [35] upravený Bergmanův minimální model. Unscented Kalmanův filtr je použit pro odhad stavů a parametrů minimálního modelu.

Model definuje rychlost výskytu glukózy $\mathbf{Ra(t)}$ na základě množství přijatých karbohydrátů $C(t)$, distribučního objemu V a maximální doby absorpce jídla τ :

$$R_a(t) = \frac{C(t)}{V\tau^2}te^{-\frac{t}{\tau}}$$

Rychlost výskytu glukózy $\mathbf{Ra(t)}$ je použita pro výpočet příjmu karbohydrátů. Bolus karbohydrátů je detekován pokud $Ra(t)$ je větší než 2 mg/dl/min a naměřená hodnota z CGM je větší než 100mg/dl. Další příjem karbohydrátů může být detekován až když $Ra(t)$ klesne pod hranici 2 mg/dl/min a uplyne alespoň 30 minut od posledního bolusu.

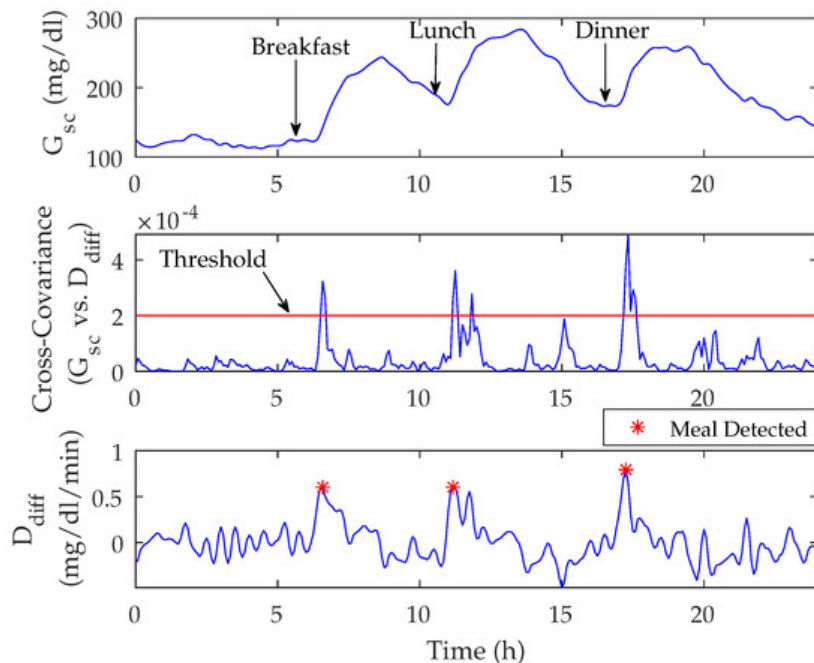
Testování bylo provedeno na sedmi reálných pacientech, kdy v první části experimentu si pacient aplikoval inzulin sám na základě doporučení diabetologa a ve druhé části bylo dávkování inzulinu řízeno algoritmem pro detekci karbohydrátů. Nutno podotknout, že experimentu se zúčastnili mladí lidé s optimálním glykemickým profilem. Experiment ukázal, že výsledky jsou statisticky ekvivaletní.

3.2.2 Unannounced Meals in the Artificial Pancreas: Detection Using Continuous Glucose Monitoring

Autoři Ramkisson et al. [25] počítají cross-covarianci mezi naměřenými hodnotami glukózy a jejich odhadem dopředného rozdílu chybového parametru D_{diff} (Unscented Kalman filter) přes tři posuvná časová okna. Pro

každé okno je aplikován jiný threshold pro detekci, přičemž se snižuje riziko falešně pozitivní detekce.

Obrázek 3.1: Naměřené hodnoty glukózy (graf 1), cross-covariance (graf 2), D_{diff} parameter (graf 3)



Zdroj: Unannounced Meals in the Artificial Pancreas: Detection Using Continuous Glucose Monitoring [25]

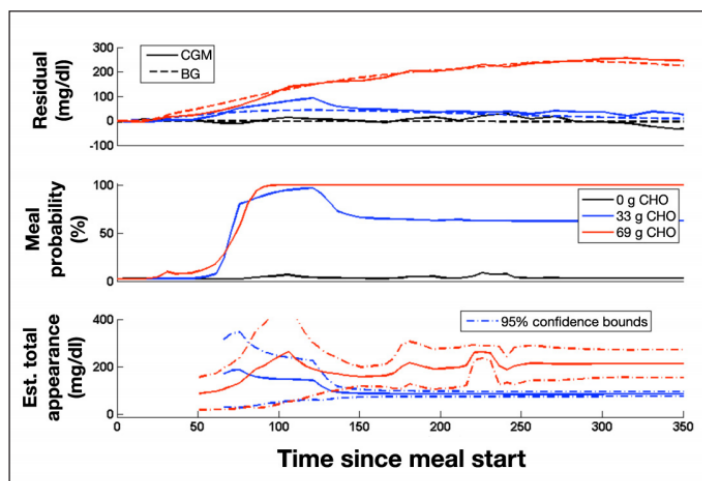
První okno mělo vysokou míru pravdivě pozitivních detekcí, ale také falešně pozitivních. Druhé a třetí okno razantně snižuje množství falešně pozitivních detekcí, ale také se sníží množství detekovaných jídel. Zpoždění detekce bylo v průměru 30 minut u prvního okna, s každým dalším oknem se zpoždění zvyšuje.

3.2.3 Probabilistic Evolving Meal Detection and Estimation of Meal Total Glucose Appearance

V této práci Fraser et al. [7] použili inzulin-glukózový model, který udává rychlost změny intersticiální glukózy závislé na působení inzulínu a endogenní produkci glukózy. Provede se analýza změn rozdílů mezi modelovanou rychlostí změny intersticiální glukózy a naměřenou CGM. Následně jsou navzorkovány experimentálně zjištěná data příjmu karbohydrátů od 0 g do 100 g (10 vzorků) a na základě pravděpodobnosti se určí nejlepší odhad.

Na obrázku 3.2 je znázorněn provedený experiment při příjmu 0,33 g a 69 g karbohydrátů. První graf znázorňuje rozdíl mezi předpokládanou modelovou hodnotou glukózy a naměřenou hodnotou. Prostřední graf ukazuje pravděpodobnost příjmu karbohydrátů. Spodní graf ukazuje odhadovaný celkový výskyt glukózy z přijatého jídla. Hranice pro detekci je pravděpodobnost 10 %.

Obrázek 3.2: Detekce karbohydrátů v čase



Zdroj: Probabilistic Evolving Meal Detection and Estimation of Meal Total Glucose Appearance [7]

Experiment byl proveden pro 99 scénářů. Pro 11 scénářů bez jídla nebyly žádné falešně pozitivní detekce. Pro scénáře, kdy byly přijaté karbohydráty, byl algoritmus úspěšnější při příjmu většího množství karbohydrátů.

3.2.4 An Unannounced Meal Detection Module for Artificial Pancreas Control Systems

Fathi et al. [5] navrhli metodu, která využívá lineárního Kalmanova filtru nad linearizovaným Bergmanovým modelem. Přijaté karbohydráty pak způsobují nekonzistentnost tohoto filtru. Jak je v práci podotknuto, lineární Kalmanův filtr není pro tuto úlohu ideální vzhledem k tomu, že glykoregulační systém diabetického pacienta je nelineární a hladina glukózy může během dne kolísat i bez zjevných příčin. Proto se v jiných pracích používá nelineární Unscented Kalmanův filtr. Úspěšnost a rychlost detekce pro různou velikost jídel je vidět v tabulce na obrázku 3.3.

Obrázek 3.3: Úspěšnost detekce vzhledem hladině glukózy a množství přijatých karbohydrátů

Sensitivity TP / (TP + FN)	93.23 %
Meal CHO = 40g	84.77 %
Meal CHO = 60g	96.29 %
Meal CHO = 80g	98.63 %
Number of false positives	64 (4.17 % of 1563)
Meal CHO = 40g	34 (6.64 % of 512)
Meal CHO = 60g	16 (3.13 % of 512)
Meal CHO = 80g	14 (2.73 % of 512)
Detection time	40 [30 - 50] min
Meal CHO = 40g	50 [40 - 60] min
Meal CHO = 60g	40 [30 - 50] min
Meal CHO = 80g	30 [30 - 40] min
Glucose increase at detection time	2.6 \pm 1.2 mmol/L
Meal CHO = 40g	2.4 \pm 1.5 mmol/L
Meal CHO = 60g	2.7 \pm 1.1 mmol/L
Meal CHO = 80g	2.8 \pm 1.0 mmol/L
Glucose increase 10 min before detection time	1.4 \pm 1.0 mmol/L
Meal CHO = 40g	1.5 \pm 1.3 mmol/L
Meal CHO = 60g	1.4 \pm 0.9 mmol/L
Meal CHO = 80g	1.2 \pm 0.7 mmol/L

Zdroj: An Unannounced Meal Detection Module for Artificial Pancreas Control Systems [5]

3.2.5 Committed Moving Horizon Estimation for Meal Detection and Estimation in Type 1 Diabetes

Committed Moving Horizon Estimation (CMHE) [1] umožňuje predikovat čas a velikost nehlášených jídel. Metoda je založena na Moving Horizon Estimation (MHE). MHE je metoda odhadů omezených stavů nelineárního diskrétního modelu, která predikuje sekvenci stavových proměnných a poruch δ (disturbance) tak, že minimalizuje chybu mezi predikovanými a měřenými hodnotami [1]. V případě detekce karbohydrátů je modelem glukózo regulační systém člověka (ve studii je použit Bergmanův minimální model) a poruchami se rozumí přijaté karbohydráty. MHE, na rozdíl od Kalmanova filtru, nepočítá s normálním rozdělením poruch.

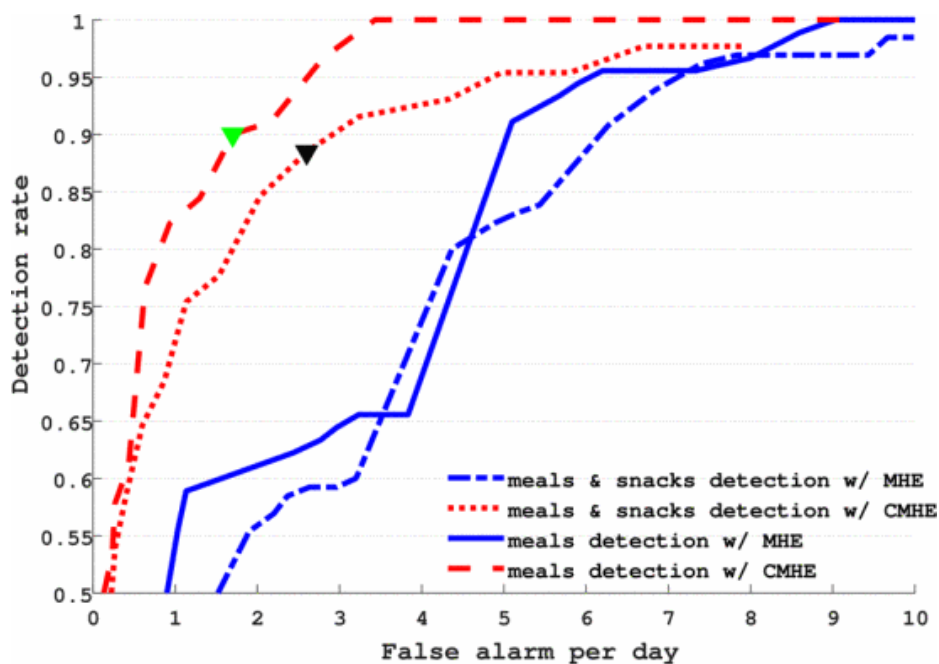
Pro každou instanci MHE (časové okno velikosti N) je N odhadů disturbance hodnoty (pro t až $t-N$). V každém časovém bodě t tak máme N odhadů v čase $t+N$. Problémem je, který z N odhadů vybrat. Prvotní odhad v čase t není přesný, protože tento odhad zohledňuje pouze naměřené hodnoty před časem t , zatímco zvýšení hladiny glukózy se projeví až s časovým odstupem od přijetí karbohydrátů v závislosti na metabolismu daného člověka. Od-

$$\Delta_{t-V} = \frac{(\sum_{i=t-V+1}^t W(i)^b \cdot \delta_{t-V}^i)}{\sum_{i=t-V+1}^t W(i)^b}$$

CMHE instance (min Ct) se optimalizují pomocí Mixed-integer Quadratic Programming (MIQP), který disturbance parametr reprezentuje jako čtvercovou vlnu (za předpokladu, že příjem karbohydrátů po dobu jídla je konstantní). Jídlo je detekováno pokud Δt je nad daným *thresholdem* alespoň 80 % času w.

22

Obrázek 3.5: Detekce MHE (modrá) a CMHE (červená)



Zdroj: *Committed Moving Horizon Estimation for Meal Detection and Estimation in Type 1 Diabetes [1]*

Tabulka 3.1: Výsledky

	breakfast	snack1	lunch	snack2	dinner	snack3
Probability of Occurrence	100%	50%	100%	50%	100%	50%
CHO amount (g)	40-60	5-25	70-110	5-25	55-75	5-15
Time of day (h)	1:00-5:00	5:00-8:00	8:00-12:00	12:00-15:00	15:00-19:00	19:00-21:00
MDE performance						
Onset Deviation (min)	22.43	30.67	17.59	25.00	20.68	12.04
CHO Deviation (g)	50.53	20.20	25.54	18.33	22.85	26.93
Detection rate	100%	92.31%	100%	91.67%	100%	27.78%

Zdroj: *Committed Moving Horizon Estimation for Meal Detection and Estimation in Type 1 Diabetes [1]*

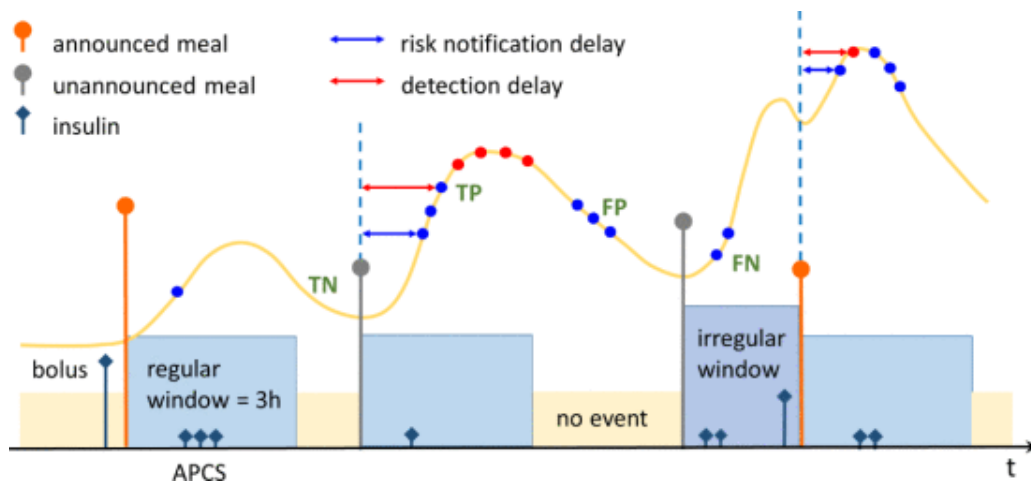
3.3 Data-driven metody

3.3.1 Unannounced Meal Detection for Artificial Pancreas Systems Using Extended Isolation Forest

Isolation forest je metoda, která detekuje a izoluje anomálie v datech. Data jsou rekurzivně rozdělena do stromové struktury (iTree - isolation tree) náhodným výběrem atributů z daných vlastností (feature) dokud nejsou jednotlivé instance izolovány. Toto náhodné rozdělení má pro anomálie signifikantně kratší cestu ve stromové struktuře [17].

Vlastnosti, které slouží jako vstup pro Extended isolation forest [39] jsou derivace naměřené glukózy v plazmě dGp , zadané karbohydráty (pokud byly zadány) a podaný inzulin. Pro výpočet derivace glukózy v čase t $dGp(t)$ je potřeba znát naměřenou glukózu v čase $t+1$ $Gp(t+1)$. Z toho důvodu nastává zpoždění detekce o 5 minut. Detekce anomálií je pak pomocí dvou thresholdů pro menší a větší míru rizika nezadaných karbohydrátů. Příklad detekce anomálií je na obrázku 3.6. Notifikace je vznešena pokud se objeví jedna anomálie s větší mírou rizika (červené), nebo 3 anomálie s menší mírou rizika (modré).

Obrázek 3.6: Úspěšnost detekce vzhledem hladině glukózy a množství přijatých karbohydrátů



Zdroj: *Unannounced Meal Detection for Artificial Pancreas Systems Using Extended Isolation Forest* [39]

Metoda byla testována na čtyřech virtuálních pacientech (pro jejich simulaci byl využit Hovorkovo model) V simulovaných pětidenních denních

scénářích přijal pacient průměrně 3,4 jídla za den. Pro scénář, kdy bylo jídlo v 50 % zadáno a v 50 % nezadáno, dosahovala metoda úspěšnosti 90,8 % a v 6,2 % případů byla falešně pozitivní. V případě, že jídlo nebylo zadáno vůbec, byla úspěšnost 90,0 % a v 11,47 % falešně pozitivní. Zpoždění detekce je 26-39 minut [39].

3.3.2 A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator

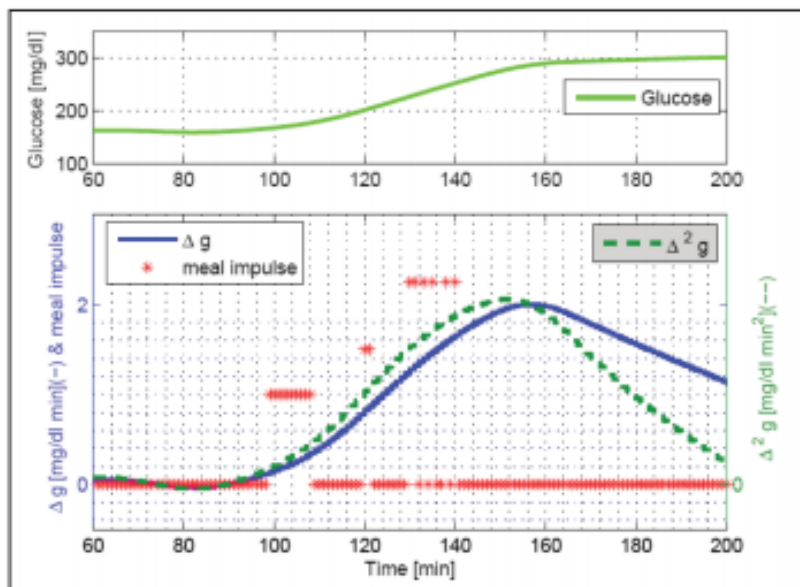
Hyunjin et al. [11] použili hlasovací schéma pro detekci karbohydrátů. Algoritmus je založen na kontinuálním pozorování první a druhé derivace koncentrace glukózy, které při splnění kritérii vyše sérii impulsů (až 15 impulsů během 30ti minut). Tato metoda je založena na vzorkování po jedné minutě.

Impulsy jsou vyslány při překročení thresholdu. Pro nastavení thresholdu se vychází z toho, že jedno jídlo způsobí nárůst rychlosti výskytu glukózy v krvi dG (první derivace) 0-2 mg/dl/min, druhé derivace d_2G 0-0,02 mg/dl/min². Thresholdy jsou proto nastaveny na 0, 0.5, 1.25, 1.8 pro dG a 0, 0.005, 0.0125, 0.018 pro d_2G . Jelikož je mezi dG a d_2G zpoždění, další impuls je vyslán pokud se hodnoty kříží. Vyslání impulsů je na obrázku 3.7.

Tyto impulsy jsou poté zesíleny a převedeny na gramy karbohydrátů (ve studii počítají se 4 g na impuls), viz obrázek 3.8.

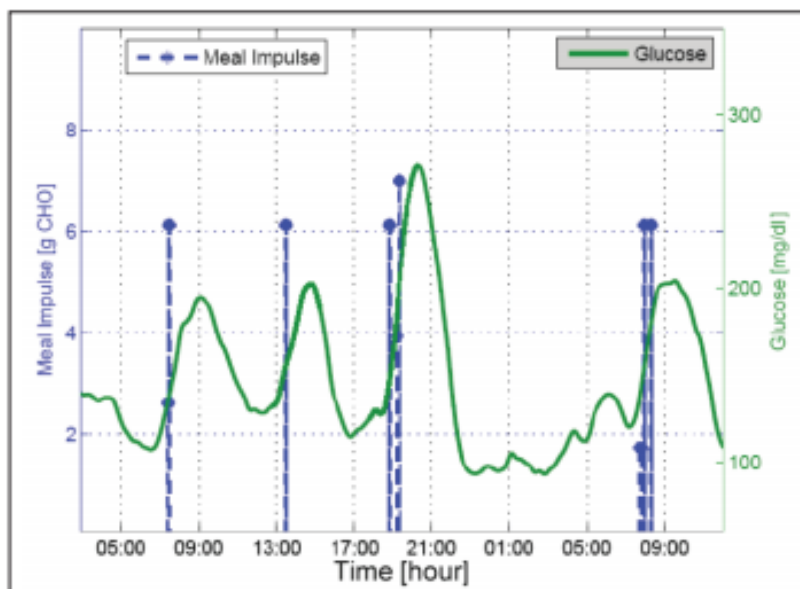
Počet impulsů, časové okno, thresholdy a množství karbohydrátů na impuls lze individuálně změnit podle diabetického profilu pacienta.

Obrázek 3.7: Impulsy



Zdroj: *A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator [11]*

Obrázek 3.8: Predikované karbohydráty



Zdroj: *A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator [11]*

Algoritmus detekuje 656 jídel z 800 (82 %) a 144 falešně pozitivních

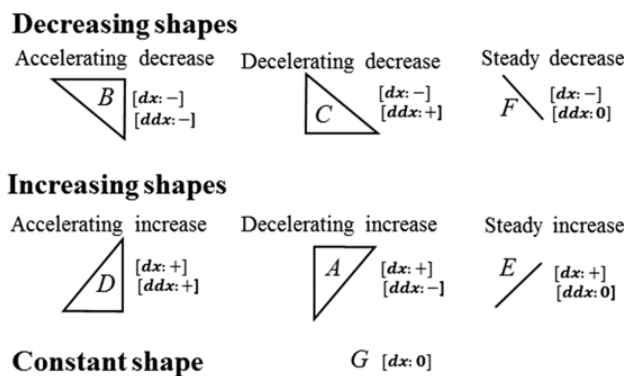
detekci. Průměrný odhad jídla je 36,41 g karbohydrátů. Zpoždění detekce je průměrně 31 minut [11].

3.3.3 Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data

V této studii Samadi et al. [27] použili waveletové filtrování pro odstranění šumu z dat naměřených CGM. Waveletové filtrování rozdělí frekvenční obsah vstupních dat do několika pásem, kde se jinou měrou projevuje šum a užitečná složka [12]. Zvolený parametr míry filtrování musí být dostatečně velký, aby odfiltroval šum, ale ne příliš, aby se neodfiltrovaly ostré nárůsty způsobené příjmem karbohydrátů.

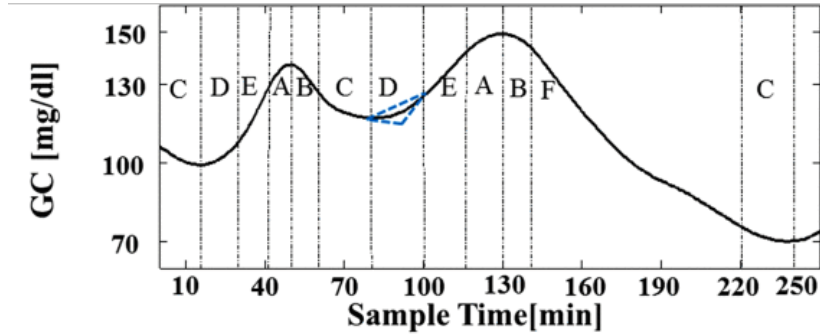
Pro extrakci vlastností je použita trojúhelníková kvalitativní reprezentace. V kvalitativních metodách extrakce dat jsou časová data konvertována na sekvenci kvalitativních proměnných. Pro trojúhelníkovou kvalitativní reprezentaci jsou kvalitativní proměnné trojúhelníky (viz obrázek 3.9). Trojúhelníky dělí data na epizody (obrázek 3.10) jejichž hraniční body jsou buďto extrém ($dG/dt = 0$) nebo inflexní bod ($d_2G/dt = 0$). Epizody se nepřekrývají a každé dvě sousední jsou rozdílné. Hodnoty dx a ddx jsou vypočteny z numerické derivace a mapovány na $+/0/-$.

Obrázek 3.9: Kvalitativní proměnné trojúhelníkové kvalitativní reprezentace



Zdroj: Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data [27]

Obrázek 3.10: Data rozdělená na epizody



Zdroj: Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data [27]

Pro tuto studii byla definice epizod modifikována tak, že velikost epizody je pevně dána a obsahuje určitý počet vzorků. Epizody se překrývají a sousední mohou být stejné. Výpočet dx a ddx udávající tvar trojúhelníku je:

$$d^1G(n) = G(n) - G(n - l_e)$$

$$d^2G(n) = G'(n) - G'(n - l_e)$$

$$G'(n) = G(n) - G(n - 1)$$

Do trojúhelníkové reprezentace se přidá neurčitost (fuzzy logika) tak, že epizodu popisuje více tvarů trojúhelníků (definují se další způsoby pro výpočet d^1x a d^2x).

Pro detekci příjmu karbohydrátů se každému trojúhelníku přiřadí váha odpovídající tvaru. Vynásobením vah s vektorem fuzzy trojúhelníkové reprezentace vznikne proměnná “*increase of glucose trend index*” I_{gt} pro každý vzorek (váhy jsou nastaveny tak, že hodnota I_{gt} je v intervalu $\langle -3; 3 \rangle$). Pakliže hodnota I_{gt} (nebo součet I_{gt} v časovém okně) překročí hranici 2,5, je detekován příjem.

Úspěšnost detekce byla u dospělých 87 % (falešně pozitivní 21 %) a u dětí 93 % (falešně pozitivní 3 %). Zpoždění detekce je přibližně půl hodiny [27].

3.3.4 Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1

Kölle et al. [14] použili Moving horizon estimator pro odhad glukózy a lineární diskriminační analýzu pro klasifikaci.

Porovnávány jsou 4 metody (2 na základě klasifikace a 2 metody pro porovnání za použití thresholdu):

1. **Klasifikace odhadu Ra horizontů**

Odhady rychlosti výskytu glukózy Ra horizontů jsou predikovány na základě Bergmanova modelu. Na tyto Ra je použita Lineární diskriminační analýza (LDA).

2. **Klasifikace CGM horizontů**

Klasifikace je na hrubých vyhlazených datech z CGM modulu. Také v tomto případě je použita LDA.

3. **Threshold pro aktuální Ra odha**

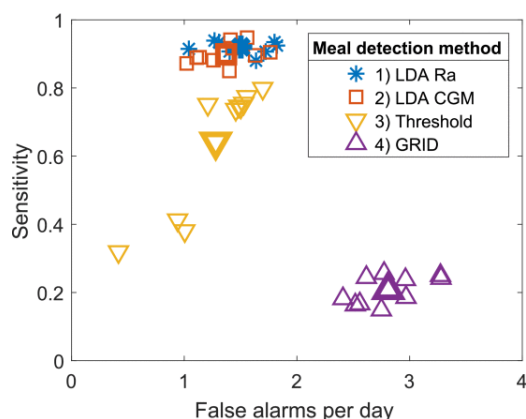
Detekce přijatých karbohydrátů je pokud Ra překročí daný threshold (viz kapitola 3.2.1).

4. **GRID algoritmus**

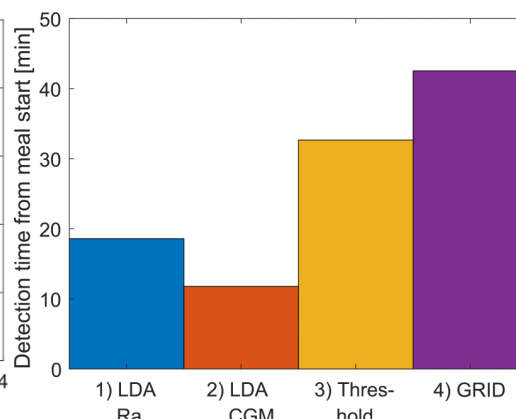
Threshold pro naměřené hodnoty glukózy a jejich rychlost změny.

Na obrázku 3.11 je výkonnost jednotlivých metod při provedeném experimentu. Je vidět, že metody využívající LDA si vedly výrazně lépe. Jejich úspěšnost se pohybuje kolem 88 %. Také v případě času potřebného k detekci překonaly metody založené na thresholdu. Srovnání zpoždění u jednotlivých metod je na obrázku 3.12.

Obrázek 3.11: Úspěšnost detekce



Obrázek 3.12: Zpoždění detekce



Zdroj: *Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1* [14]

3.4 Metody využívající neuronové sítě

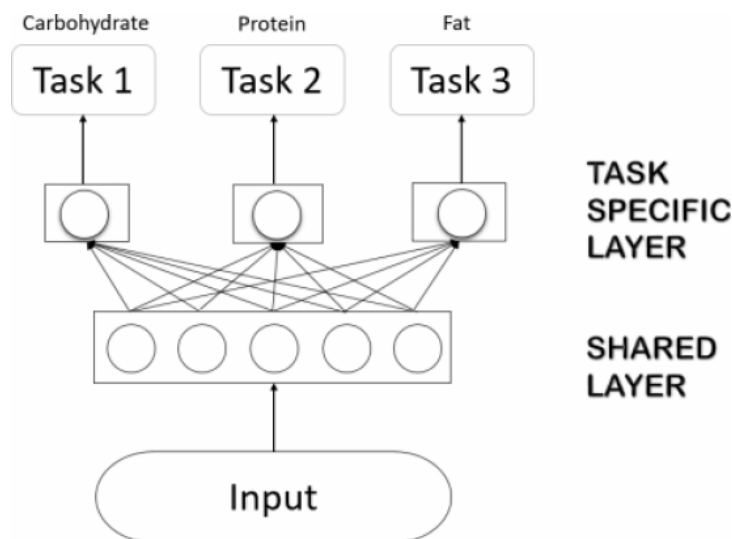
3.4.1 Predicting the meal macronutrient composition from continuous glucose monitors

V této studii Huo et al. [10] použili multitaskovou neurální síť pro určení složení (karbohydráty, proteiny, tuky) v přijatém jídle.

Pro zachycení charakteristických vlastností z dat CGM byl vypočten určitý integrál (oblast pod křivkou) pro 5 časových bodů (klidový stav nalačno, zvýšená glykémie, střední hodnota při návratu do klidového stavu, pokles glukózy a konečný stav).

Multitasková neuronová síť obsahuje jednu společnou vrstvu pro vstup (vypočítané integrály) a výstupní vrstvu (viz obrázek 3.13). Aktivační funkce pro vstupní vrstvu je Rectified Linear Units (ReLU), pro výstupní lineární funkce, ztrátová funkce je Huberova ztrátová funkce. Trénování neuronové sítě bylo na 1000 epochách.

Obrázek 3.13: Neuronová síť



Zdroj: Predicting the meal macronutrient composition from continuous glucose monitors [10]

Trénování neuronové sítě a následná cross-validace byla provedena pro dva scénáře. První je natrénování neuronové sítě na datech několika pacientů a validace na datech jiného pacienta. Druhý způsob spočívá v natrénování neuronové sítě na části dat jediného pacienta (vynechání jednoho jídla) a validace na zbylých datech. Z experimentu vyšel lépe druhý způsob zohledňující metabolismus konkrétního jedince se střední kvadratickou chybou pro karbohydráty 0,39. Kvadratická chyba karbohydrátů pro první způsob je 0,45 [10].

Metoda počítá s poměrně přesným příjmem jídla v časovém rozmezí 8 hodin. Jako vstup pro neuronovou síť pak jsou vypočtená data z tohoto celého časového okna. Při sběru dat byla také vyloučená fyzická aktivita, která ovlivňuje hladinu glukózy. Z těchto důvodů není metoda vhodná pro řízení dávkování inzulínu v reálném čase.

3.4.2 Continuous glucose monitoring prediction

V této studii Jose et al. [13] implementovali 3 různé algoritmy pro vývoj glykémie po jídle. Studie se nezabývá vlastní detekcí příjmu karbohydrátů.

Seasonal Autoregressive Integrated Moving Average (SARIMA) je regresní model pro časové řady. SARIMA počítá s trendem a sezónními parametry, které musí být před použitím nastaveny. Pro nastavení těchto para-

metrů byla použita autokorelace a částečná autokorelace.

V druhé metodě je použit Kalmanův filtr pro predikci vývoje hladiny glukózy pacienta.

V poslední metodě jsou implementovány dvě rekurentní neuronové sítě, konkrétně Long Short Term Memory model (LSMT). Pro druhou neuronovou síť slouží výstup LSMT jako vstup do další vrstvy spolu s konstantami. Těchto 7 konstant jsou první, druhý peak a střední hodnota Fouriroy transformace, střední hodnota a rozptyl Waveletové transformace a diskretní Waveletové transformace provedenými nad daty CGM. Model byl natrénován na 10 epoch.

Pro výpočet chyby a ztrátové funkce u všech metod je použita střední absolutní chyba (MEA) a střední kvadratická chyba (RMSD). V tabulce 3.2 jsou výsledky jednotlivých metod.

Tabulka 3.2: Výsledky

Algorithm	Train		Test	
	MAE	RMSE	MAE	RMSE
SARIMA	3.87	6.90	3.45	4.16
KALMAN	3.21	4.36	3.08	3.79
RNN	0.32	0.56	0.22	0.47
Modified LSTM	11.0	20.15	5.43	7.05

Zdroj: Continuous glucose monitoring prediction [13]

3.5 Porovnání jednotlivých metod

V tabulce 3.3 je srovnání jednotlivých metod pro detekci karbohydrátů. Kri-
térii pro porovnání jednotlivých metod je přesnost detekce a zpoždění de-
tekce od doby příjmu karbohydrátů.

Tabulka 3.3: Porovnání jednotlivých metod

Studie	Zpoždění*	TP	FN	TN	FP
3.2.1 Turksoy et al. [35]	-	-	-	-	-
3.2.2 Ramkissooon et al. [25]	25-40	91.6%**	-	-	-
3.2.3 Fraser et al. [7]	-	-	-	11	0
3.2.4 Fathi et al. [5]	30-50	93.23%	6.77%	-	4.17%
3.2.5 Chen et al. [1]	20	88.5%	11.5%	-	2.6FP/den
3.3.1 Zheng et al. [39]	26-39	90%	10%	-	11.47%
3.3.2 Hyunjin et al. [11]	31 (avg.)	656 (82%)	144 (18%)	-	54 (6.75%)
3.3.3 Samadi et al. [27]	30 (avg.)	87%	13%	-	21%
3.3.4 Ra Kölle et al. [14]	18.59	92%	8%	-	1.5FP/den
3.3.4 CGM Kölle et al. [14]	11.78	90%	10%	-	1.37FP/den
3.4.1 Huo et al. [10]	8 hodin	0.39***			

* doba detekce karbohydrátů od přijetí jídla v minutách

** úspěšnost detekce jednotlivých oken v tabulce

*** RMSRE - střední kvadratická chyba

TP - true positive

TN - true negative

FP - false positive

FN - false negative

4 Analýza metod detekce fyzické aktivity

Fyzická aktivita výrazně snižuje koncentraci glukózy v krvi. Z toho důvodu je potřeba před zahájením cvičení snížit dávkování inzulínu, jinak hrozí riziko hypoglykémie. Systémy kontinuální monitorace glukózy jsou sice schopny signalizovat pokles hladiny glukózy, ale tato detekce nemusí nastat včas, nebo není dostatečná. Proto je třeba detekovat parametry fyzické aktivity jako takové.

Většina publikovaných metod využívá k detekci data ze senzorů pro měření srdeční aktivity, teploty kůže, vodivosti kůže a akcelerometrů pro snímání pozice a pohybu těla. Tyto senzory mohou být externí, nebo součástí CGMS.

4.1 Metody

4.1.1 The Impact of Accelerometer and Heart Rate Data on Hypoglycemia Mitigation in Type 1 Diabetes

Stenerson et al. [31] rozšířili existující Predictive low glucose suspend (PLGS) algoritmus. Tento algoritmus predikuje z dat CGM předpokládaný vývoj koncentrace glukózy v krvi a na základě této predikce pozastaví podávání inzulínu inzulinovou pumpou. Snižuje tak riziko hypoglykémie [6]. Pro měření akcelerace a srdečního rytmu použili senzor Zephyr BioHarness 3, k měření glukózy pak Dexcom G4 Platinum CGM nebo Medtronic Sof-Sensor CGM. Predikční okno algoritmu je 30 minut a hranice koncentrace glukózy pro pozastavení podání inzulínu je 80mg/dl. Algoritmus je následně rozšířen tak, aby pozastavil podání inzulínu v případě, že tepová frekvence překročí 90 tepů za minutu, nebo velikost vektoru dat akcelerometru překročí 0,1 a současně koncentrace glukózy v krvi je pod 180 mg/dl. Pro experiment bylo použito 11 061 měření CGM z čehož 336 bylo pod 70 mg/dl (hypoglykemický stav). Při simulaci bylo zjištěno, že použití samotného PLGS redukuje riziko hypoglykémie o 62 %. Při použití PLGS v kombinaci s daty akcelerometru

a srdečního tepu je redukce 76 % (viz výsledky v tabulce 4.1).

Tabulka 4.1: Počet hypoglykemických stavů

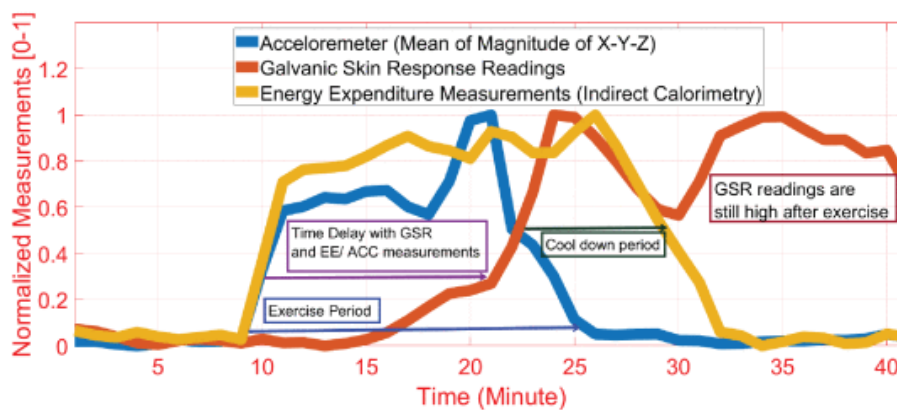
Intervention	Readings below 70 mg/dl
Actual (real life)	336
PLGS	127
PLGS + HRM	97
PLGS + accelerometer	86
PLGS + accelerometer + HRM	80

Zdroj: The Impact of Accelerometer and Heart Rate Data on Hypoglycemia Mitigation in Type 1 Diabetes [31]

4.1.2 Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions in Diabetes Management

Sevil et al. [30] využívají strojové učení pro klasifikaci fyzické aktivity. Empatica E4 wristband měří elektrodermální aktivitu, puls, teplotu kůže a akceleraci. Tato data senzor měří každou sekundu. V grafu na obrázku 4.1 je vidět, jak fyzická aktivita ovlivňuje měřená data.

Obrázek 4.1: Příklad naměřených dat



Zdroj: Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions in Diabetes Management [30]

Z naměřených dat elektrodermální aktivity, teploty a akcelerace byl odstraněn šum pomocí median filtru a Savitzky-Golay filtru. Hodnoty pulsu byli očištěny o šum pomocí Butterworthova band-pass filtru a waveletové dekompozice. Následně se z dat extrahovali charakteristické vlastnosti pro každý minutový interval. Použité parametry jsou uvedeny v tabulce 4.2. Celkem se extrahovalo 2216 příznaků, z čehož bylo použito 1730 příznaků.

Tabulka 4.2: Počet hypoglykemických stavů

Physical State Classification		
Feature Name	Time Delay (min)	p-value
Norm of 4th detail coefficient of Daubechies wavelet decomposition of ACC X-axis	0	≤ 0.01
Total energy response of BVP	1	≤ 0.01
Range for Magnitudes of ACC	0	≤ 0.01
Kurtosis of GSR	3	≤ 0.01
Standard Deviation of ACC / Std of GSR	0	≤ 0.01
Median of HR / Median of GSR	0	≤ 0.01
Std of ACC / Mean of ACC	0	≤ 0.01
Skewness of BVP	1	≤ 0.01
Minimum of HR	0	≤ 0.01
Mean of ST	0	≤ 0.01
p-values calculated using a two-sided t-test at $\alpha = 1\%$ significance level		
EE Estimation (Features with COSMED EE Measurements)		
Feature Name	Time Delay (min)	Correlation
Std of magnitude of ACC (X-Y-Z)	0	0.56
Mean of logarithm of magnitude of ACC (X-Y-Z)	0	0.55
Maximum of ACC Z-axis	0	0.55
Zero cross of BVP	0	0.49
Minimum of high frequency of ST (0.15-0.4 Hz)	0	0.49
Root-mean-square energy of GSR	4	0.48
Maximum of power spectral density estimate of HR	0	0.48
1st Quartile of HR	0	0.48
Norm of 1st detail coefficient of Daubechies wavelet decomposition of GSR	0	0.48
Summation of absolute differentiation of GSR	12	0.34

Zdroj: Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions in Diabetes Management [30]

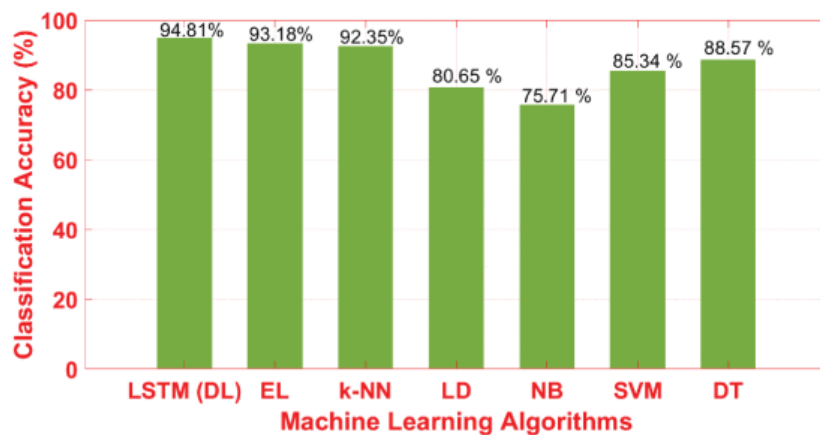
Vysoce korelovaná data mohou negativně ovlivnit proces strojového učení. Proto bylo použita metoda hlavních komponent (PCA) pro redukci dimenze. Následně byly použity tyto metody strojového učení:

- Long-Short Term Memory (LSTM)

- Ensemble Learning (EL)
- k-Nearest Neighbors (k-NN)
- Linear Discrimination/Regression (LD)
- Naive Bayes (NB)
- Support Vector Machine/Regression (SVM)
- Decision Tree (DT)

V grafu na obrázku 4.2 je přesnost jednotlivých metod. Nejlépe vychází LSTM s přesností 94,81 %.

Obrázek 4.2: Porovnání jednotlivých metod



Zdroj: Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions in Diabetes Management [30]

4.1.3 Sensor Monitoring of Physical Activity to Improve Glucose Management in Diabetic Patients: A Review

Autoři této studie Ding – Schumacher [3] zjišťovali korelaci mezi koncentrací glukózy v krvi a fyzickou aktivitou. Využili k tomu zařízení SenseWear® Pro Armband vyvinuté firmou BodyMedia (Pittsburgh, PA, USA) pro měření pohybových dat. Toto zařízení je upevněno kolem ruky a sbírá data z pěti typů senzorů. Sensory snímají pozici a pohyb ruky a těla, teplotu kůže a okolí, vodivost kůže a srdeční aktivitu.

Autoři následně použili regresní model pro odhad koncentrace glukózy a zjišťovali korelaci těchto odhadů s daty CGM. Pro experiment, kdy pacienti šli 60 minut po běžeckém pásu, byl korelační koeficient $r = 0,90$.

4.1.4 Black-box Model Identification of Physical Activity in Type-1 Diabetes Patients

V této práci Faccioli et al. [4] rozšířili Multi Input Single Output (MISO) black-box model predikce glukózy o data pohybového senzoru. Výsledný model je:

$$g(t) = [G_1(q^{-1}, \theta)G_2(q^{-1}, \theta)G_3(q^{-1}, \theta)] \begin{bmatrix} i(t) \\ m(t) \\ a(t) \end{bmatrix} + H(q^{-1}, \theta)e(t)$$

kde $g(t)$ je hodnota intersticiální glukózy, $i(t)$ je množství podaného inzulínu, $m(t)$ je množství přijatých karbohydrátů, $a(t)$ je fyzická aktivita a $G_1(q^{-1}, \theta), G_2(q^{-1}, \theta), G_3(q^{-1}, \theta), aH(q^{-1})$ jsou transfer funkce, kde $q^{-1}[g(t)] = g(t-1)$.

Následující hodnota intersticiální glukózy má rovnici:

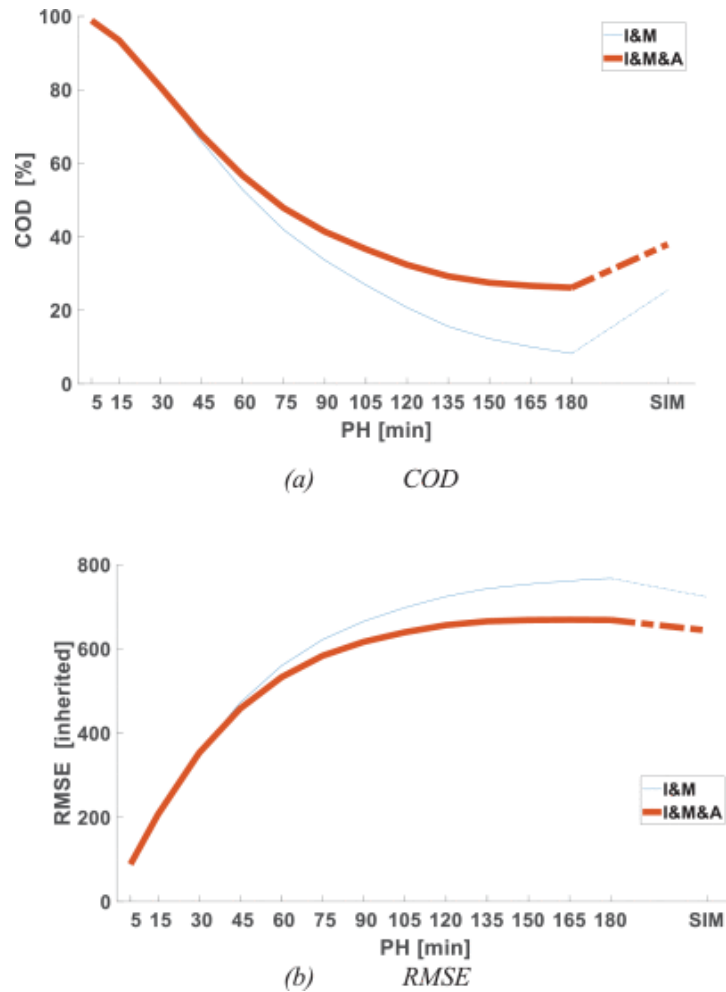
$$\begin{aligned} \hat{g}(t|t-1, \theta) &= H^{-1}(q^{-1}, \theta)[G_1(q^{-1}, \theta)G_2(q^{-1}, \theta)G_3(q^{-1}, \theta)] \begin{bmatrix} i(t) \\ m(t) \\ a(t) \end{bmatrix} \\ &+ [1 - H^{-1}(q^{-1}, \theta)]g(t) \end{aligned}$$

Parametr θ získáme minimalizováním chyby:

$$\hat{\theta} = \arg \min_{\theta} \sum_{t=1}^N (g(t) - \hat{g}(t|t-1, \theta))^2$$

Jako metrika pro vyhodnocení výsledků byly použita střední kvadratická chyba (RMSE) a Koeficient determinace (COD). Porovnání modelu rozšířeného o data fyzické aktivity (I&M&A) oproti základnímu modelu (I&M) v závislosti na délce fyzické aktivity (PH) jsou v grafech na obrázku 4.3. Z grafů je vidět, že pro krátké intervaly si jsou modely podobné, pro delší už I&M&A dosahuje lepších výsledků.

Obrázek 4.3: Porovnání základního modelu (I&M) s modelem rozšířeným o data fyzické aktivity (I&M&A)



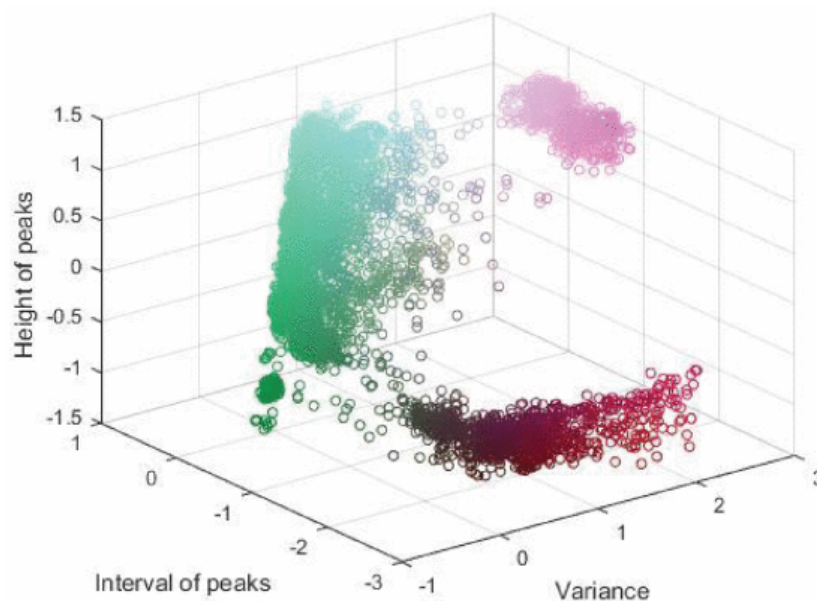
Zdroj: Physical Activity and Psychological Stress Detection and Assessment of Their Effects on Glucose Concentration Predictions in Diabetes Management [30]

4.1.5 A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors

Qi et al. [24] implementovali algoritmus pro detekci fyzické aktivity za pomoci dvou pohybových senzorů. Data akcelerace z obou senzorů jsou nejprve vyhlazena pomocí Savitzky-Golay filtru. Hodnoty tepu jsou získány z EKG senzoru.

V první fázi rozdělí rozdělí aktivity na aerobní cvičení a cvičení se závažím. K tomu použili one-class support vector machine (OC-SVM) klasifikátor. Threshold pro zařazení cvičení do dané kategorie je určen třemi vlastnostmi a to vzdáleností mezi vrcholy, výškou vrcholů a variací z dat akcelerometrů (rozložení je na obrázku 4.4). Přesnost klasifikace je 85 %.

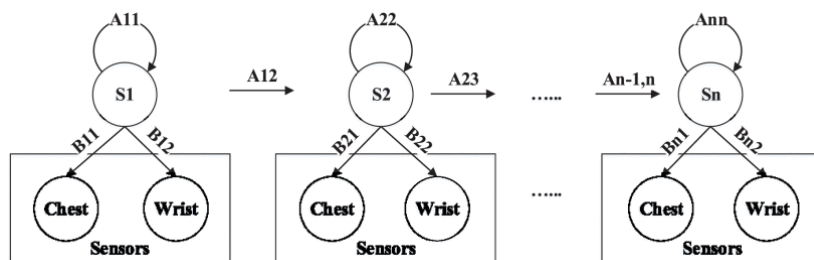
Obrázek 4.4: Rozložení dat akcelerometru podle vzdálenosti mezi vrcholy (ružová), výškou vrcholů (zelená) a variací (červená)



Zdroj: A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors [24]

Následně skrytý Markovův model (HMM) jsou použity pro klasifikaci posilovacích cvičení a neuronové sítě pro klasifikaci aerobních cvičení. U HMM se vychází z toho, že pro určitý cvik se opakuje několik po sobě jdoucích stavů. Pro klasifikaci byl vytvořen soubor stavů a přechodů mezi nimi. Struktura HMM je na obrázku 4.5.

Obrázek 4.5: Struktura skrytého Markovova modelu



Zdroj: *A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors [24]*

Před klasifikací aerobních cvičení se redukuje dimenze dat pomocí metody hlavních komponent (PCA). Následně se použije 3 vrstvá neuronová síť. Vstupem vstupní vrstvy jsou data akcelerometru ze zápěstí, skrytá vrstva má 18 a výstupní 9, což odpovídá počtu aerobních aktivit.

Přesnost detekce jednotlivých typů aktivit je mezi 86 % - 92 %.

4.1.6 Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor

Tapia et al. [33] použili pro detekci fyzické aktivity 5 tříosých akcelerometrů a monitor srdečního tepu. Data ze senzorů byly interpolovány cubic spline a rozděleny do posuvného okna po 4,2 sekundách. Pro každé okno pak byla spočítaná plocha pod křivkou, variance, střední hodnota, entropie a korelace pro každou osu. Pro detekci byl použit Naivní Bayesův klasifikátor.

Pro klasifikátor natrénovaný na specifických datech každého jedince byla úspěšnost detekce 94,6 %. Pro klasifikátor natrénovaný na obecných datech byla úspěšnost detekce pouze 56,3 %. Podle této studie má srdeční tep pouze malý vliv (1,2 %) na úspěšnost detekce.

4.2 Porovnání metod

Všechny zkoumané metody využívají k detekci fyzické aktivity alespoň jeden dodatečný senzor. Tyto senzory poskytují data akcelerace ve třech osách. Tato data jsou následně filtrována a jsou získány charakteristické vlastnosti. Část metod pracuje také s daty srdečního rytmu.

Pro klasifikaci dat je následně použita některá z metod strojového učení (kromě metody v kapitole 4.1.1, kde se používají pouze thresholdy). Obecně lepších výsledků dosahují metody s tříosým akcelerometrem umístěným na zápěstí. Nejlepší výsledky dosahuje LSTM neuronová síť v kapitole 4.1.2. Porovnání metod detekce fyzické aktivity je v tabulce 4.3.

Tabulka 4.3: Porovnání metod detekce

Studie	Úspěšnost detekce
4.1.2 Sevil et al. [30] LSTM	94,81 %
4.1.2 Sevil et al. [30] EL	93,18 %
4.1.2 Sevil et al. [30] k-NN	92,35 %
4.1.2 Sevil et al. [30] LD	80,65 %
4.1.2 Sevil et al. [30] NB	75,71 %
4.1.2 Sevil et al. [30] SVM	85,34 %
4.1.2 Sevil et al. [30] DT	88,57 %
4.1.5 Qi et al. [24]	86 % - 92 %
4.1.6 Tapia et al. [33] specifická data	94,6 %
4.1.6 Tapia et al. [33] obecná data	56,3 %

Metoda 4.1.1 Stenerson et al. [31] není zařazena do porovnání protože autoři neudali počet detekovaných aktivit, ale pouze procentuální snížení počtu hypoglykemických stavů o 76 %.

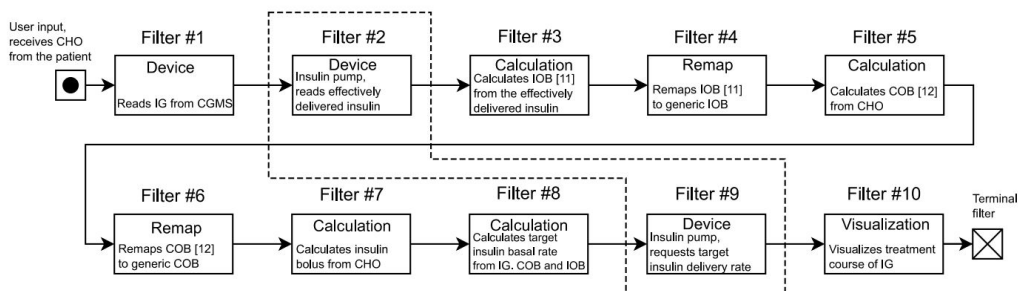
5 SmartCGMS

SmartCGMS je systém kontinuální monitorace glukózy nové generace vyvíjený na Katedře informatiky a výpočetní techniky na Fakultě aplikovaných věd Západočeské univerzity v Plzni. Umožňuje kontinuální monitoraci glukózy, modelování dynamiky glukózy a řízení inzulinové pumpy.

Aplikace je napsaná v jazyce C++ a je zkompilována pro x86-64 systémy MS Windows, macOS, Debian GNU/Linux a pro armeabi-v7a a arm64-v8a [15]. Skládá se z filtrů, modelů, metrik a solverů.

Filtr je základní stavební prvek aplikace, který poskytuje určitou funkcionalitu. Ty mají jednotné rozhraní a jsou poskládány lineárně za sebou (viz příklad konfigurace na obrázku 5.1). To umožňuje vysokou modularitu a zaměnitelnost filtrů (například filtr, který čte data ze senzoru CGMS, se může nahradit filtrem, který přečte data uložená v logu, bez nutnosti změn ve zbytku konfigurace). Filtry spolu komunikují předáváním událostí. Událost obsahuje typ události (level, info, parameters), id filtru, který událost vytvořil, typ signálu, hodnotu, časový segment a čas vytvoření události. Nejčastější je událost typu level, která v sobě má hodnotu naměřeného nebo vypočteného signálu. Události se zpracovávají postupně tak, jak přichází. Filtr může událost poslat dál beze změny, událost upravit, nebo událost zahodit a případně vygenerovat nový signál.

Obrázek 5.1: Příklad konfigurace filtrů SmartCGMS



Zdroj: SmartCGMS as an Environment for an Insulin-Pump Development with FDA-Accepted In-Silico Pre-Clinical Trials [36]

Model umožňuje načtení více hodnot a jejich následné zpracování dle nastavených parametrů. Narozdíl od filtru model zpracovává události dávkově. Je tak vhodný pokud potřebujeme pracovat s více signály najednou nebo

predikovat hodnoty do budoucna. SmartCGMS obsahuje modely dynamiky glukózy Bergmanův model vylepšený Hovorkovo modelem a Koutného difúzním modelem, SimGlucose s vynulovaným parametrem ξ u UVA/Padova S2013, T1DMS a DMMS.R [16]. Výpočet modelu je spuštěn filtrem *Signal Generator* nebo filtrem *Calculated Signal*, pokud chceme parametry modelu dynamicky spočítat pomocí solveru.

Metrika je druh filtru, který umožňuje spočítat metriky jako je například střední chyba, směrodatná odchylka nebo plocha pod křivkou [15].

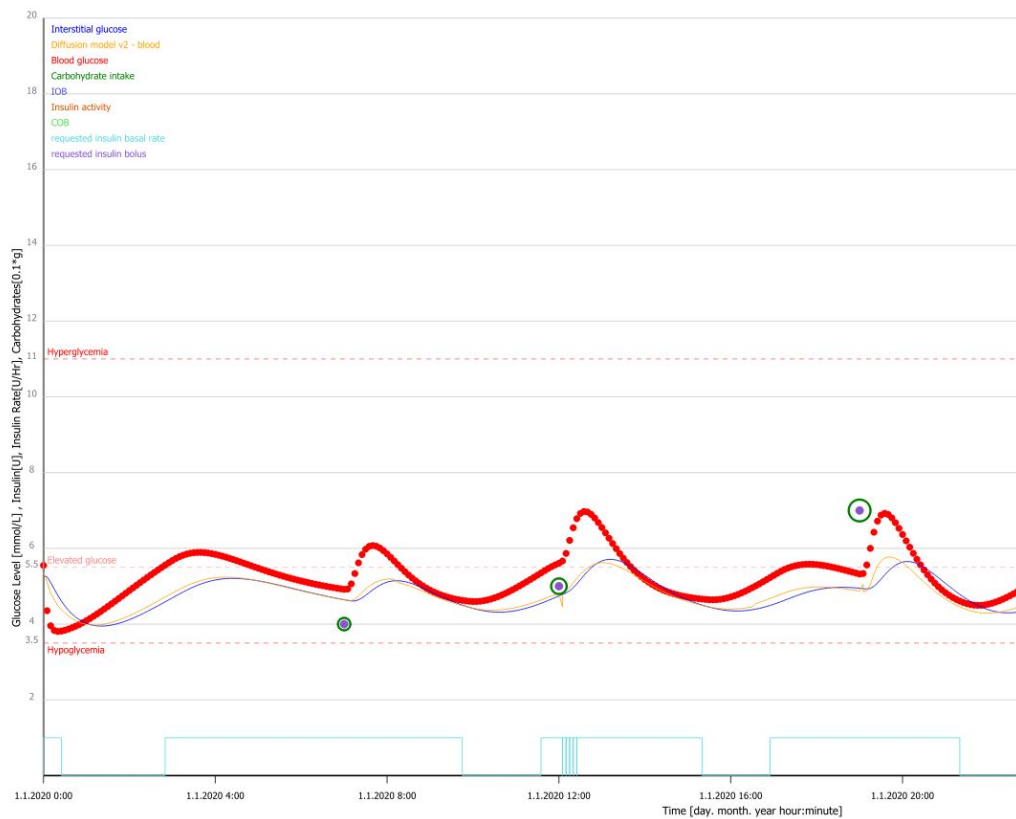
Pro určení parametrů modelu lze využít **solver**. Ten na základě metrik signálů modelu spočítá jeho optimální parametry. Příkladem solveru je Meta-differential evolution, Pathfinder, Sequential brute force scan, Particle swarm optimization, Spiral optimization a další.

Vytvořené filtry, modely a solvery se zkompilují jako dynamická knihovna, která se načte při spuštění SmartCGMS. Aplikace také umožňuje integraci skriptů v jazyce Matlab. Skripty je nutné definovat v souboru *matlab_manifest.xml*.

SmartCGMS lze spustit v příkazové řádce (**console3.exe**) nebo v grafickém uživatelském rozhraní (**gpredict3.exe**). V záložce *Filters* si uživatel nastaví jednotlivé filtry a jejich parametry. V záložce *Simulation* pak spustí běh programu a může prohlížet grafy a logy s výstupy (příklad grafického výstupu je na obrázku 5.2). Konfiguraci lze uložit a poté opětovně načíst.

SmartCGMS je dostupný na <https://diabetes.zcu.cz/smartcgms>.

Obrázek 5.2: Příklad výstupu SmartCGMS



Zdroj: SmartCGMS

6 Návrh metody detekce karbohydrátů

Příjem karbohydrátů způsobuje zvýšení hladiny koncentrace glukózy v krvi, které je nutné kompenzovat podáním inzulínu. Včasnou detekcí může systém upravit množství podávaného inzulínu tak, aby nedošlo k hyperglykémii.

Detekce karbohydrátů je na základě dat intersticiální glukózy měřené senzorem CGMS a její derivace. Pro detekci byly navrženy a testovány tři metody:

- Long short-term memory neuronová síť (kapitola 6.1)
- Lineární a kvadratická diskriminační analýza (kapitola 6.2)
- Detekce hran pomocí thresholdů 1. difference hodnot intersticiální glukózy (kapitola 6.3)

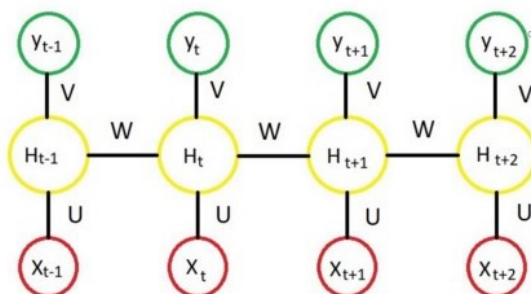
6.1 Rekurentní neuronové sítě

Rekurentní neuronové sítě jsou vhodné pro predikci sekvenčních dat (například časové řady). Vyznačují se tím, že si předávají informaci o předchozím stavu (aktivaci skryté vrstvy):

$$H_t = \sigma(W_h H_{t-1} + W_x X_t + b)$$

kde W jsou váhy a b je bias. Neurony, které takto uchovávají svůj stav, se nazývají buňky.

Obrázek 6.1: Rekurentní neuronová síť



Zdroj: <https://medium.com/>

Riziko RNN spočívá v náchylnosti vzniku jevu označovaného jako mizející nebo explodující gradient vznikající při zpětné propagaci, kdy hodnoty exponenciálně klesají nebo rostou. To je z důvodu, že se chyba propaguje přes všechny iterace. Tento problém eliminuje long short-term memory nebo gated recurrent unit neuronová síť.

6.1.1 Long short-term memory

Long short-term memory neuronová síť (LSTM) je speciální případ rekurentní neuronové sítě. V každé iteraci si buňka uchovává dodatečnou informaci nazvanou *memory*.

Nejprve se ze vstupního vektoru X_t a stavu předchozí buňky H_{t-1} spočte Forget gate f_t , Candidate layer \bar{C}_t Input gate I_t a Output gate O_t [38]:

$$f_t = \sigma(X_t \otimes U_f + H_{t-1} \otimes W_f)$$

$$\bar{C}_t = \tanh(X_t \otimes U_c + H_{t-1} \otimes W_c)$$

$$I_t = \sigma(X_t \otimes U_i + H_{t-1} \otimes W_i)$$

$$O_t = \sigma(X_t \otimes U_o + H_{t-1} \otimes W_o)$$

kde W a U jsou váhové vektory. Následně se spočte *memory* prvek C_t :

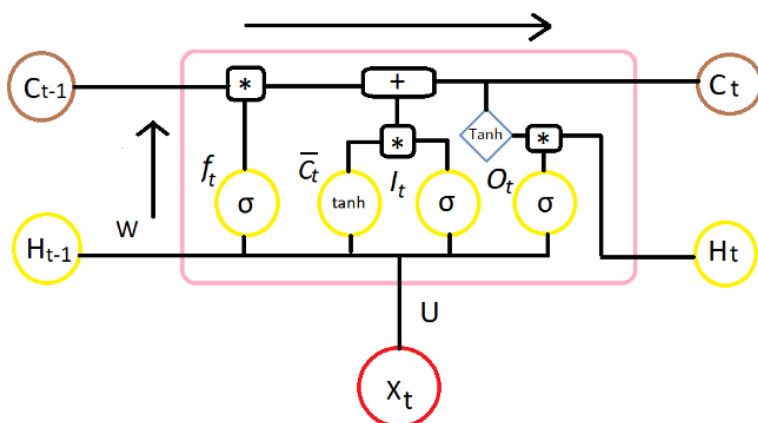
$$C_t = f_t \otimes C_{t-1} \oplus I_t \otimes \bar{C}_t$$

a výstupní stav H_t :

$$H_t = O_t \otimes \tanh(C_t)$$

Diagram LSTM buňky je na obrázku 6.2.

Obrázek 6.2: LSTM buňka



Zdroj: <https://medium.com/>

6.1.2 Gated recurrent unit

Gated recurrent unit neuronová síť (GRU) využívá Update gate z_t a Reset gate r_t . Ty určují, jaká část informace bude propagována na výstup. Mohou tak vytěžit relevantní informace. Update a reset gate se spočte [2]:

$$z_t = \sigma(W_z x_t + U_z h_{t-1})$$
$$r_t = \sigma(W_r x_t + U_r h_{t-1})$$

kde W a U jsou váhové vektory. Následně se spočte relevantí informace z předchozího stavu \bar{h}_t :

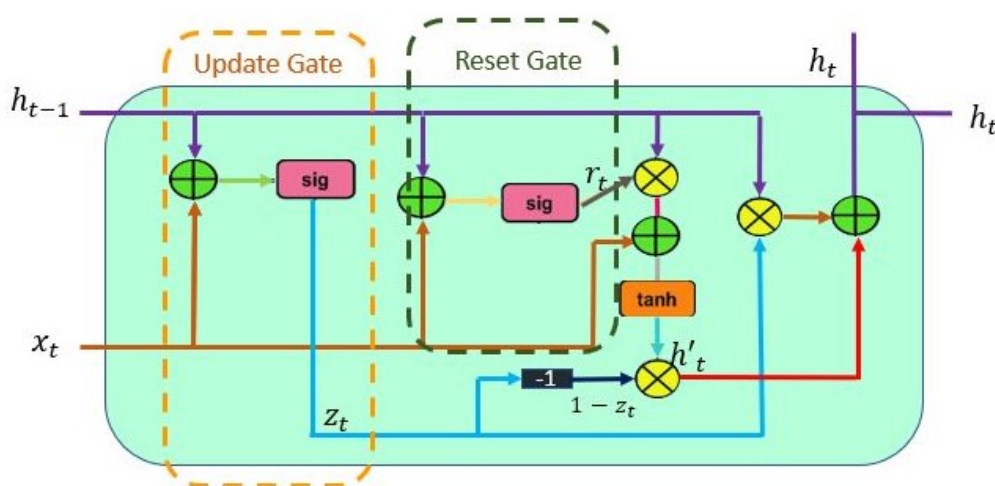
$$\bar{h}_t = \tanh(W_h x_t + U_h(r_t \otimes h_{t-1}))$$

a výstupní stav h_t :

$$h_t = z_t \otimes h_{t-1} + (1 - z_t) \otimes \bar{h}_t$$

Diagram GRU buňky je na obrázku 6.3.

Obrázek 6.3: GRU buňka



Zdroj: <https://www.pluralsight.com/>

6.1.3 Model

Pro detekci karbohydrátů byl sestaven sekvenční keras model¹ se čtyřmi vrstvami:

- Obousměrná LSTM nebo GRU

¹Keras je nadstavba nad knihovnou TensorFlow pro neuronové sítě

- Dropout(0,5)
- Dense(128)
- Dense(1)

První vrstvou je rekurentní neuronová síť. Ta může být buď LSTM nebo GRU. Tato vrstva má 128 neuronů, dropout 0,2 a aktivační funkcí je hyperbolický tangens. Vstupem vrstvy je časové okno N vstupních prvků velikosti W . Tvar vstupních dat je $W \times N$:

$$\begin{bmatrix} X_1^1 & X_1^2 & \dots & X_1^N \\ X_2^1 & X_2^2 & \dots & X_2^N \\ \dots & \dots & \dots & \dots \\ X_W^1 & X_W^2 & \dots & X_W^N \end{bmatrix}$$

Dropout vrstva s $rate=0,2$ náhodně nastaví některé vstupy na nulu v poměru 0,2. Nevynulované vstupy jsou škálovány $1/(1-rate)$, takže součet všech vstupů je nezměněn. To zabrání přetrénování neuronové sítě. Dropout vrstva je aplikována pouze při trénování sítě. První Dense vrstva (propojení neuronů formou každý s každým) má 128 neuronů, aktivační funkce je *relu*. Druhá Dense vrstva s jedním neuronem je výstupní.

6.2 Lineární a kvadratická diskriminační analýza

Diskriminační analýza byla vybrána vzhledem k vysoké úspěšnosti a malému zpoždění, kterého dosáhli Kölle et al. [14] v článku **Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1** (kapitola 3.3.4).

Diskriminační analýza slouží k rozdělení prvků do konečného počtu tříd na základě lineární kombinace charakteristických prvků. Pro klasifikaci potřebujeme znát posterioiry tříd $\mathbf{P}(\mathbf{Y}|\mathbf{X})$. Pravděpodobnostní model $\mathbf{P}(\mathbf{X}|\mathbf{y}=\mathbf{k})$ pro každou třídu \mathbf{k} lineární a kvadratické diskriminační analýzy je dán vícerozměrnou Gaussovo distribucí [9]:

$$P(\mathbf{X}|\mathbf{y} = \mathbf{k}) = \frac{1}{(2\pi)^{p/2} |\Sigma_{\mathbf{k}}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{x}-\mu_{\mathbf{k}})^T \Sigma_{\mathbf{k}}^{-1}(\mathbf{x}-\mu_{\mathbf{k}})}$$

kde \mathbf{p} je počet prvků. Pro predikci je pak použito Bayesovo rozhodovací pravidlo [9]:

$$P(y = k|x) = \frac{P(x|y=k)P(y=k)}{P(x)} = \frac{P(x|y=k)P(y=k)}{\sum_l P(x|y=l)P(y=l)}$$

kdy hledáme třídu s maximálním posteriorem.

Logaritmus posterioru pro kvadratickou diskriminační analýzu (QDA) je [9]:

$$P(y = k|x) = -\frac{1}{2}\log|\Sigma_k| - \frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k) + \log P(y = k)$$

Predikovaná třída je ta, která má maximální logaritmický posterior.

Lineární diskriminační analýza (LDA) je speciální případ QDA, kdy předpokládáme, že třídy mají stejnou kovarianční matici. Logaritmus posterioru je pak vyjádřen [28]:

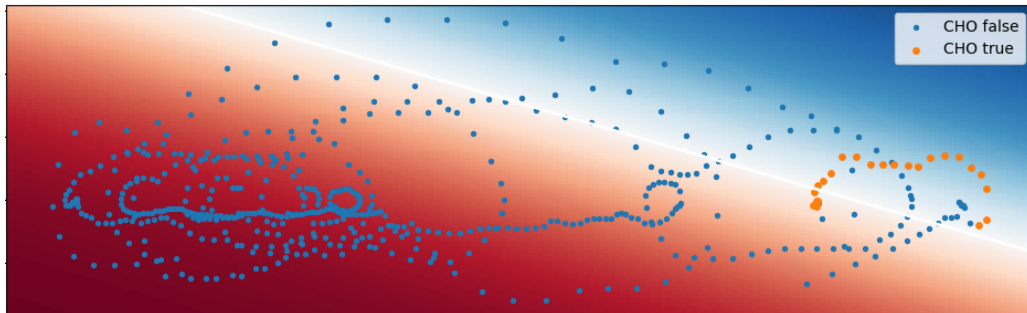
$$P(y = k|x) = -\frac{1}{2}(x - \mu_k)^T \Sigma^{-1}(x - \mu_k) + \log P(y = k)$$

Vstupem diskriminační analýzy může být buď množina prvků pro daný časový okamžik nebo časové okno jednoho prvku. Predikované třídy jsou příjem/nepříjem karbohydrátů. Ty dostaneme převedením sloupce karbohydrátů na booleanovskou proměnnou (0 je false, cokoli větší než 0 je true).

6.2.1 Množina prvků

Nejčastějším vstupem diskriminační analýzy je množina charakteristických prvků X_i . V tomto případě se diskriminační analýza snaží rozdělit prostor tak, aby odlišila jednotlivé třídy (viz obrázek 6.4).

Obrázek 6.4: Rozdělení prostoru lineární diskriminační analýzou dvou prvků



6.2.2 Časové okno

V případě časového okna může být na vstupu pouze jediný prvek. To je dáno tím, že vstupem diskriminační analýzy je jedno-dimenzionální pole dat. Vstupní data mají tvar:

$$\begin{bmatrix} X_{i-W} & X_{i-W-1} & \dots & X_{i-1} & X_i \end{bmatrix}$$

kde W je velikost časového okna. Časové okno jako vstupní data použili i Kölle et al. [14].

Při experimentální implementaci diskriminační analýzy se nepodařilo dosáhnout stejných výsledků, jako autoři původního článku. Nejlepších výsledků dosahovala kvadratická diskriminační analýza, jejímž vstupem bylo dvouhodinové časové okno. Z testovacích 150 zadaných jídel jich detekovala 54 % a falešně pozitivních detekcí bylo 155. Rozdílné výsledky mohou být dány tím, že původní metoda byla otestována pouze s daty predikovanými na základě Bergmanova modelu, ale ne s reálnými daty pacientů.

6.3 Detekce hran průběhu intersticiální glukózy

Tato metoda detekuje vzestupné a klesající hrany průběhu intersticiální glukózy pomocí thresholdů první difference měřených hodnot intersticiální glukózy.

Derivace funkce $f'(x)$ v bodě x je směrnici tečny funkce $f(x)$ v daném bodě. Její hodnota je rovna tangens úhlu α , který svírá tečna s osou x . Pakliže je funkce $f(x)$ v bodě x rostoucí, $tg(\alpha)$ je kladný. Pokud je $f(x)$ klesající, $tg(\alpha)$ je záporný. Velikost derivace v bodě x pak udává velikost změny $f(x)$, neboli říká, jak strmě funkce stoupá či klesá. Derivace funkce $f(x)$ v bodě x_1 můžeme vyjádřit jako:

$$f'(x) = \frac{d}{dx} f(x) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

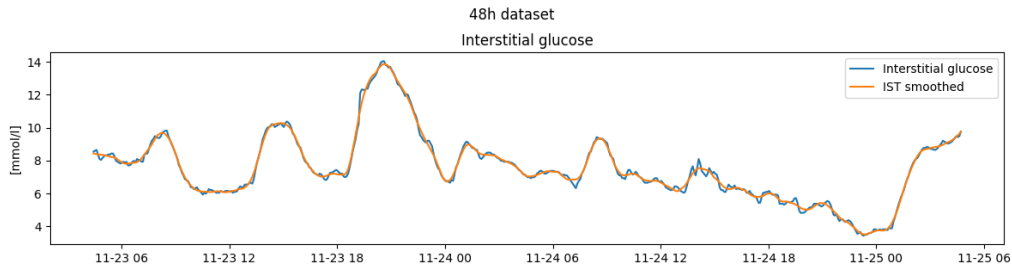
Jelikož data ze senzoru CGMS jsou diskrétní, můžeme derivaci nahradit rovnicí první difference. Pro data intersticiální glukózy tak dostáváme vztah:

$$\Delta IST = \frac{IST_t - IST_{t-1}}{\Delta t}$$

Naměřená data intersticiální glukózy mohou být zatížena šumem a případně i dočasnými výchyly hodnot glukózy. Každá taková výchylka pak má značný vliv na hodnotu derivace. Z toho důvodu je potřeba data před výpočtem vyhladit. Pro vyhlazení dat byl zvolen digitální Savitzky-Golay

filtr. Každá podmnožina $2m + 1$ prvků je vzorkována na polynom stupně p ($p \leq 2m$) ve smyslu nejmenších čtverců [18]. Pro vyhlazení dat intersticiální glukózy byl zvolen polynom stupně 3 a velikost podmnožiny 21. Vyhlazená data vidíme v grafu na obrázku 6.5.

Obrázek 6.5: Vyhlazená data intersticiální glukózy



Následně se data ohodnotí. Pokud ΔIST překročí určitý threshold th_i , přiřadí se váha w_i . Takových dvojic thresholdů a vah může být libovolné množství. Experimentálně byla zjištěna kombinace thresholdů $th = [0.0125, 0.018]$ a vah $w = [2.25, 3]$.

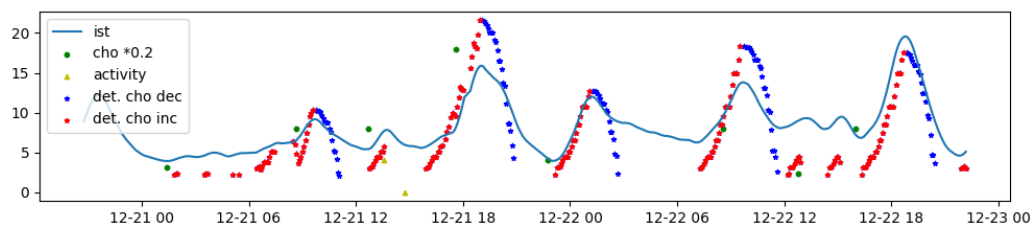
Samotné ohodnocení podle thresholdů zachytí jakýkoli jednorázový větší výkyv v datech intersticiální glukózy. Jelikož příjem karbohydrátů se projevuje zvýšením glykémie v řádu desítek minut až hodin, chceme znát vývoj křivky intersticiální glukózy v čase. Z toho důvodu pro kladně ohodnocená data zvýšíme ohodnocení v případě, že předchozí hodnoty vykazují vzrůstající trend po dobu dvou hodin nazpět (tj. 24 hodnot, data jsou vzorkována po pěti minutách). Získáme tak aktivační funkci pro rostoucí hrany.

```
if activation[i] > 2:
    for j in range(24):
        if activation[i-j] >= 2+0.2*j:
            activation[i] += 0.1*j
```

Pro klesající hrany použijeme stejný postup, ale se záporným ohodnocením.

V grafu na obrázku 6.6 je ukázka detekce hran. Červeně je aktivace vzestupné hrany, modře aktivace sestupné hrany posunutá. Počátek sestupné hrany je posunut na úroveň poslední hodnoty aktivace vzestupné hrany. Mezi vzestupnou a sestupnou hranou je vždy několik hodnot, kdy byla hladina intersticiální glukózy vyrovnaná.

Obrázek 6.6: Detekované hrany průběhu intersticiální glukózy



Jídlo je detekováno, pakliže aktivační funkce vzestupné hrany je větší než zvolený threshold. Nízký threshold detekuje většinu jídel, ale také může detekovat výkyvy nesouvisející s příjmem karbohydrátů. Naopak vysoce zvolený threshold nebude detekovat menší jídla. Řešením je použití více thresholdů pro stanovení pravděpodobnosti jídla.

7 Návrh metody detekce fyzické aktivity

Při fyzické zátěži dochází k poklesu cukru v krvi. Její detekcí a snížením množství podávaného bazálu inzulínu lze snížit riziko hypoglykémie.

Senzor CGMS měří buď srdeční tep, počet kroků a elektrodermální aktivitu v pětiminutových intervalech nebo akceleraci a elektrodermální aktivitu v minutových intervalech (viz kapitola 8.2). Tyto hodnoty se zvyšují zvláště u aerobních cvičení. Při posilovacích cvičeních nemusí být tak markantní.

Srdeční tep se v klidovém stavu pohybuje mezi 60 - 80 tepy za minutu, při zátěži pak mezi 100 - 180 tepy za minutu dle typu zátěže. U trénovaných sportovců jsou tyto hodnoty nižší. Srdeční tep zvyšují i další vlivy jako je stres nebo vypití kofeinového nápoje.

Fyzická aktivita je často určitá forma pohybu. Ten je měřen buď akcelerací nebo počtem kroků. U aerobních cvičení budou změny v pohybu značné. U posilování, kdy je člověk většinou na místě, budou zanedbatelné. Vyšší hodnoty budou také při běžné chůzi.

Elektrodermální aktivita je kožně galvanická reakce, tj. udává vodivost kůže. Při cvičení dochází k pocení, čímž se elektrodermální aktivita pokožky zvyšuje. K tomu dochází až se zpožděním. Zvýšení může být opět v důsledku více vlivů.

Navržený algoritmus detekuje fyzickou aktivitu na základě těchto ukazatelů. Pokud hodnota překročí daný threshold, je detekována fyzická aktivita. Použit se může buď aktuální hodnota nebo průměr za určité časové okno. Thresholdy jsou individuální pro každého pacienta (především srdeční tep se u každého liší). Pro snížení falešně pozitivních detekcí je vhodné vytvořit kombinace ukazatelů, které se budou podílet na detekci. Aby byla fyzická aktivita detekována, musí být hodnota všech ukazatelů nad daným thresholdem. Tím se eliminují vlivy, které působí pouze na jeden ukazatel.

Jelikož při fyzické aktivitě dochází k poklesu hladiny glukózy, může se detekce zpřesnit detekcí sestupné hrany. K tomu lze použít algoritmus *Detekce hran průběhu intersticiální glukózy* navržený v kapitole 6.3. Pro detekci sestupné hrany jsou thresholdy a váhy záporné. Detekci sestupné hrany není možné použít samostatně, protože k poklesu krevního cukru dochází i v důsledku podání vyššího množství inzulínu.

8 Implementace

8.1 Použité nástroje

Při návrhu metod byl použit jazyk Python 3.7, který je vhodný pro analýzu dat, zvláště pak jeho knihovny NumPy, Pandas a SciPy pro zpracování a analýzu dat a matplotlib pro vykreslení dat. Dále jsou použity knihovny TensorFlow a Keras pro práci s neuronovými sítěmi a knihovnu Scikit-learn pro strojové učení.

Implementace modulů do SmartCGMS je v jazyce C++. Použití Keras modelu neuronové sítě je za pomoci knihovny frugally-deep[8].

8.2 Dataset BGLP

Dataset BGLP je veřejně dostupný soubor dat, který obsahuje anonymizovaná data ze senzoru CGMS. Dataset obsahuje naměřené a zadané hodnoty:

- Glukóza v krvi (BG)
- Intersticiální glukóza (IST)
- Bazální množství inzulínu
- Bolus inzulínu
- Příjem karbohydrátů (CHO)
- Fyzická aktivita
- Kvalita spánku
- Elektrodermální aktivita
- Teplota kůže
- Teplota okolí
- Srdeční tep
- Počet kroků
- Akcelerace

Intersticiální glukóza je měřena v pětiminutových intervalech. Glukóza v krvi, podávaný inzulín (bolus i bazál), karbohydráty, fyzická aktivita a

kvalita spánku jsou zadávány pacientem. Senzor měří buď srdeční tep a počet kroků, nebo akceleraci. Pacient měl zároveň senzor Empatica Embrace nebo Basis Peak. V případě Empatica Embrace se měří velikost akcelerace v minutových intervalech. Basis Peak měří srdeční tep a počet kroků v pětiminutových intervalech Elektrodermální aktivita a teplota kůže a okolí se měří ve stejném intervalu podle toho, zda se měří tep nebo akcelerace.

CGMS senzor posílá data ve formě signálů, které mají strukturu:

- Logical Clock
- Device Time
- Event Code
- Signal
- Info
- Segment Id
- Event Code Id
- Device Id
- Signal Id

Příklad signálů ze senzoru CGMS je v tabulce 8.1. Informaci pro následnou detekci v sobě nesou sloupce Device Time (čas měření), Signal (typ signálu) a Info (hodnota).

Tabulka 8.1: Signály ze CGMS

Logical Clock	Device Time	Event Code	Signal	Info	Segment Id	Event Code Id	Device Id	Signal Id
1	2022-01-02 00:...	Time_Segment...			1	9	{172EA814-9DF...	{00000000-000...
2	2022-01-02 00:...	Level	Interstitial glucose	11.8769	1	2	{78DF0982-070...	{3034568D-F49...
4	2022-01-02 00:...	Level	requested insulin ba...	0.7	1	2	{78DF0982-070...	{B5897BBD-1E3...
5	2022-01-02 00:...	Level	Heartbeat	82	1	2	{78DF0982-070...	{6DFCFD02-C4...
6	2022-01-02 00:...	Level	Steps	11	1	2	{78DF0982-070...	{D0D33895-FBE...
7	2022-01-02 00:...	Level	Electrodermal activity	9.5e-05	1	2	{78DF0982-070...	{ACFF91A6-D9...
8	2022-01-02 00:...	Level	Air temperature	26.5	1	2	{78DF0982-070...	{1AB2ABF6-E7...
9	2022-01-02 00:...	Level	Skin temperature	29.7	1	2	{78DF0982-070...	{6FB4AA36-C0...
10	2022-01-02 00:...	Level	Interstitial glucose	12.0434	1	2	{78DF0982-070...	{3034568D-F49...
13	2022-01-02 00:...	Level	Heartbeat	81	1	2	{78DF0982-070...	{6DFCFD02-C4...
14	2022-01-02 00:...	Level	Steps	3	1	2	{78DF0982-070...	{D0D33895-FBE...
15	2022-01-02 00:...	Level	Electrodermal activity	0.000128	1	2	{78DF0982-070...	{ACFF91A6-D9...
16	2022-01-02 00:...	Level	Air temperature	26.7	1	2	{78DF0982-070...	{1AB2ABF6-E7...
17	2022-01-02 00:...	Level	Skin temperature	29.6	1	2	{78DF0982-070...	{6FB4AA36-C0...
18	2022-01-02 00:...	Level	Interstitial glucose	12.0434	1	2	{78DF0982-070...	{3034568D-F49...
21	2022-01-02 00:...	Level	Heartbeat	78	1	2	{78DF0982-070...	{6DFCFD02-C4...

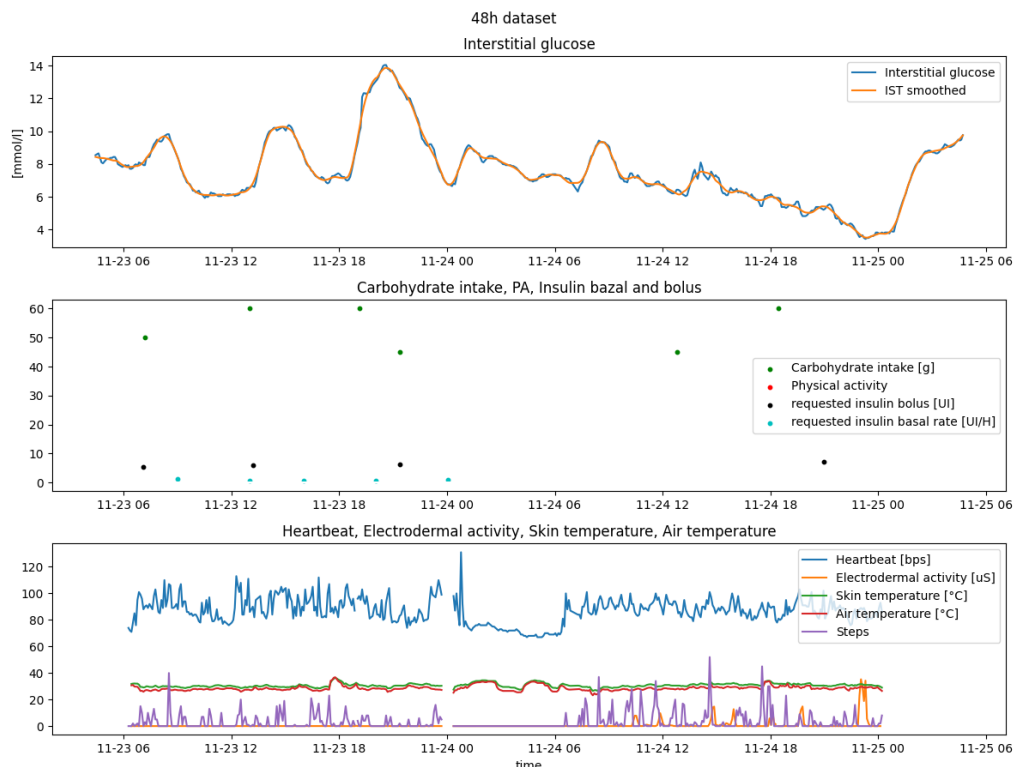
8.2.1 Analýza dat

Pro analýzu dat bylo vytvořeno několik skriptů v Pythonu. V první řadě bylo nutné data ve formě signálů transformovat do dvourozměrné tabulky, kdy řádky jsou čas měření (Device Time), sloupce typ signálu (Signal). Jelikož různé typy signálů nejsou měřeny ve stejný okamžik, lze řádky seskupit podle zvoleného signálu.

Data lze modifikovat. Chybějící data lze nahradit interpolací Akima spline[29] nebo střední hodnotou. Vyhlazení dat je pomocí Savitzky-Golay filtru. Data lze normalizovat na interval $<0,1>$ nebo $<-1,1>$. Pro výpočet derivace 1., 2. a 3. řádu se používá Akima spline nebo B-spline. Případně lze spočítat 1. a 2. diferenci.

V grafech na obrázku 8.1 jsou transformovaná data naměřená senzorem CGMS za 48 hodin. Na prvním grafu jsou hodnoty Intersticiální glukózy a její vyhlazení Savitzky-Golay filtrem. Druhý graf znázorňuje zadanou bazální dávku inzulínu, bolusy, příjem karbohydrátů a fyzickou aktivitu. Na posledním grafu jsou hodnoty srdečního tepu, počtu kroků, elektrodermální aktivity, teploty kůže a teploty okolí.

Obrázek 8.1: Data ze CGMS



Další skripty slouží pro analýzu metod detekce karbohydrátů a fyzické aktivity. Použití skriptů je popsáno v příloze B.

8.3 Filtr pro SmartCGMS

Algoritmy jsou do SmartCGMS implementovány v podobě filtrů. Ty implementují rozhraní *scgms::IFilter* a *refcnt::IReferenced*. Při vytvoření instance filtru se volá metoda *Configure*, která slouží pro nastavení filtru (typicky přečtení a nastavení konfiguračních parametrů). Metoda *Execute* je volána pokaždé, když přijde signál od předchozího filtru. Tato metoda vykonává požadovanou funkcionalitu. Původní událost se na závěr zpravidla posílá dalšímu filtru. Současně se musí zaregistrovat descriptor nového filtru, který definuje ID filtru, název a konfigurační parametry, a případně i descriptor nového signálu. Vytvořená dynamická knihovna musí exportovat funkce *do_create_filter*, *do_get_filter_descriptors* a *do_get_signal_descriptors*.

Data mohou být rozdělena do časových segmentů dle měření. V takovém případě filtry pracují s každým segmentem zvlášť. Rozdělení do segmentů je realizováno pomocí mapy, kdy klíčem je ID segmentu a hodnotou datová struktura pro daný filtr. Na konci segmentu se údaj z mapy vymaže.

Implementované algoritmy pracují s daty v určitém časovém okně. Pro tyto účely byl vytvořen vlastní spojový seznam na principu klouzavého okénka **swl**. Ten dědí od *std::deque*, která umožňuje vkládání prvků z obou stran i indexaci. Konstruktor má jeden parametr určující velikost okna. Při vložení nadlimitního prvku se odebere prvek z druhého konce seznamu.

Pro vyhlazení dat je implementován **Savitzky-Golay** filtr. Tomuto filtru se nastavuje typ signálu, který má vyhladit, velikost okna a stupeň polynomu. Výstupem je **IST smoothed** signál.

Pro vyhodnocení úspěšnosti detekce je implementován **Evaluate** filtr. Tento filtr neřeší časové segmenty. Nastavit lze zkoumaný a referenční signál, časové okno pro detekci, cooldown pro započtení falešně pozitivního výsledku a minimální počet referenčních aktivit za den. Filtr započítává dvě úrovně detekovaného signálu. Nižší (1) pro detekované aktivity a vyšší (2) pro potvrzení. Filtr zaznamenává výsledky pro každý den. Pokud počet referenčního signálu je menší, než minimální počet, den se nezapočte do celkového součtu.

Možnosti nastavení všech filtrů a spuštění detekce je popsáno v příloze B.

8.3.1 Detekce příjmu karbohydrátů

CHO detection filtr počítá aktivační funkci a detekci karbohydrátů. To je implementováno rekurentní neuronovou sítí nebo detekcí hran průběhu intersticiální glukózy tak, jak je popsáno v kapitole 6.3. Filtr posílá dva signály, **Activation** a **CHO probability**. **Activation** signál je výstup použitého algoritmu a **CHO probability** udává detekovaný příjem (1 - nižší pravděpodobnost, 2 - vyšší pravděpodobnost).

V případě detekce hran jsou nastaveny 2 thresholdy. Nižší pro brzkou detekci, která ale může detekovat výkyvy nesouvisející s příjmem jídla. Vyšší threshold je potvrzovací, kdy existuje vysoká pravděpodobnost příjmu karbohydrátů. Detekce je pouze pro vzestupnou hranu. Potvrzení může být i pomocí rekurentní neuronové sítě. Příklad konfigurace thresholdů je v konfiguračním souboru `setup_th.ini`. V případě detekce hran musí tomuto filtru předcházet Savitzky-Golay filtr pro vyhlazení dat intersticiální glukózy.

Rekurentní neuronová síť

Rekurentní neuronová síť využívá natrénovaného keras modelu. Ten je možné použít v C++ pomocí knihovny `frugally-deep` [8].

Neuronová síť je trénovaná na trénovacích datech dataset BGLP. Před natrénováním modelu jsou data ve formě signálů transformována a seskupena podle sloupce intersticiální glukózy, která je měřena v pětiminutových intervalech. Data intersticiální glukózy jsou interpolována Akima spline, z níž se získají chybějící hodnoty a derivaci 1. řádu.

Skript `train_rnn.py <type> <option> [IDs]` spustí transformaci a trénování dat. Vstupními parametry jsou typ rekurentní neuronové sítě a ID trénovacích dat pacientů. Model může být natrénován individuálně pro každého pacienta nebo na celém souboru dat. Vstupní data modelu jsou hodnoty intersticiální glukózy, jejich první derivace a čas signálu v minutách. Díky časovému údaji neuronová síť zahrne do predikce určité denní návyky pacienta. Časové okno je velikosti 24 (tj. 2 hodiny). Z datasetu je 80 % dat použito pro trénování a 20 % pro validaci. Data jsou do neuronové sítě dávána v dávkách o velikosti 64. Jelikož se jedná o časovou řadu, data se nepromíchávají. Jako ztrátová funkce je použita střední kvadratická chyba (`MeanSquaredError`), optimalizační algoritmus je Adam. Neuronová síť je trénovaná ve 100 epochách.

Natrénovaný model je nutné konvertovat pomocí skriptu `keras_export/`

`convert_model.py`, který je součástí `frugally-deep` knihovny. Podporované sítě jsou LSTM i GRU. SmartCGMS filtr následně tento model načte při konfiguraci. Příklad konfigurace s GRU je v konfiguračním souboru `setup_gru.ini`.

8.3.2 Detekce fyzické aktivity

Filtr **PA detection** detekuje fyzickou aktivitu na základě vybraných ukazatelů nebo jejich kombinace. Detekce probíhá z poslední hodnoty naměřené senzorem, nebo průměru více hodnot dle zvoleného časového okna. Pro každý ukazatel lze zadat `threshold`.

Pro detekci sestupné hrany je nutné před tento filtr zařadit *Savitzky-Golay* filtr pro vyhlazení dat. Nastavení `thresholdů` a `vah` je stejné jako u detekce karbohydrátů.

V případě, že se nepoužívá potvrzení detekcí sestupné hrany, posílá filtr signál **PA detected** o hodnotě 2. Při použití detekce sestupné hrany má tento signál hodnotu 1, pokud byla aktivita detekována pouze na základě ukazatelů, a hodnotu 2, pokud byla detekce potvrzena sestupnou hranou.

Příklad konfigurace s měřeným srdečním tepem a počtem kroků je v konfiguračním souboru `setup_bpm.ini`. Příklad konfigurace s akcelerací a potvrzováním pomocí detekce sestupné hrany je v konfiguračním souboru `setup_acc.ini`.

8.3.3 Možnosti rozšíření

Thresholdy a váhy u detekce hran průběhu intersticiální glukózy a thresholdy u detekce fyzické aktivity jsou zadávány jako parametry modelu. To umožňuje implementaci solveru, který by byl schopný optimalizovat thresholdy tak, aby detekce dosahovala nejlepších možných výsledků.

Další možností je propojení se senzorem akcelerace ve třech osách. Takovým mohou být například chytré hodinky nebo fitness náramek. Z dat 3-osého akcelerometru lze přesněji určit typ pohybu a tudíž přesněji rozlišit fyzickou aktivitu od běžného pohybu.

9 Výsledky

Detekce příjmu karbohydrátů a fyzické aktivity je testována na testovacích datech datasetu BGLP.

Pravdivě pozitivní detekce (TP) jsou hodnoty, kdy je aktivita (příjem jídla nebo fyzická aktivita) detekována (hodnota detekce alespoň 1) do dvou hodin od jejího zadání pacientem. Pokud je hodnota detekce 2, je aktivita brána jako potvrzená. Pakliže aktivita není detekována do dvou hodin od zadání, je výsledek falešně negativní (FN). V případě detekce hodnoty 2, kdy aktivita není zadána, je výsledek falešně pozitivní (FP). Zároveň je počítáno zpoždění detekce od zadání aktivity pacientem. Zároveň je možné nastavit minimální počet referenčních signálů za den, aby byla statistika daného dne zaznamenána do celkových výsledků měření.

Je spočítaná citlivost detekce (TPR), neboli míra pravdivě pozitivních detekcí vůči celkovému počtu zaznamenaných referenčních hodnot. Přesnost (PPV) je poměr pravdivě pozitivních detekcí vůči celkovému počtu detekcí. FNR je míra falešně negativních výsledků vůči celkovému počtu zaznamenaných referenčních hodnot. FDR je míra falešně pozitivních výsledků vůči celkovému počtu detekcí. Dále je spočteno F1 skóre, které udává harmonický průměr citlivosti a přesnosti. F1 skóre se spočítá jako:

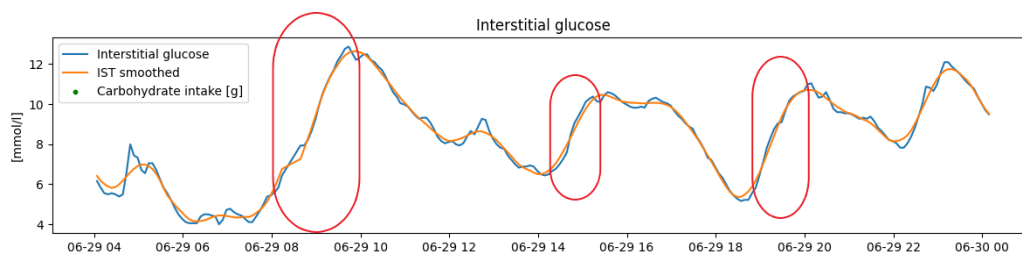
$$F_1 = 2 * \frac{PPV * TPR}{PPV + TPR}$$

Počet pravdivě negativních výsledků nelze spolehlivě určit, protože změny v měřených datech přetrvávají i po ukončení aktivity. Proto není ani možné spočítat specificitu.

9.1 Detekce karbohydrátů

Algoritmy detekce karbohydrátů byly testovány na datech jedenácti pacientů. Měření testovacích dat u každého pacienta probíhalo po dobu 10 - 11 dnů. V součtu pacienti zaznamenali 340 jídel za 109 dnů. Diabetičtí pacienti musí pravidelně jíst (uvádí se 5-6x denně). Proto se statistika z každého dne zaznamenaná byly-li zadány alespoň 3 jídla. V opačném případě je vysoká pravděpodobnost, že pacient některé jídlo nezaznamenal, což by ovlivnilo výsledky. Příklad takových dat je na obrázku 9.1. Z grafu je patrné zvýšení koncentrace glukózy v době snídaně, oběda a večeře, ale zadané karbohydráty chybí.

Obrázek 9.1: Data bez zadaných CHO



Testované algoritmy jsou detekce hran průběhu intersticiální glukózy, kdy potvrzení je nastaveným vyšším thresholdem nebo rekurentní neuronovou sítí, a samostatná GRU neuronová síť.

V tabulkách 10.1 až 10.5 v příloze A jsou výsledky detekce pro jednotlivé pacienty. V tabulce 9.1 jsou průměrné výsledky.

Tabulka 9.1: Výsledky detekce karbohydrátů

	Threshold	RNN (individual)	Th+RNN (individual)	RNN	Th+RNN
CHO	340	340	340	340	340
TP	289	141	303	169	299
TP/pac.	26,27	12,82	27,55	15,36	27,18
TP potvr.	167	-	87	169	116
Potvrzeno	47 %	-	24 %	49 %	36 %
FN	51	199	37	171	41
FN/pac.	4,64	18,09	3,36	15,55	3,73
FP	112	172	78	188	89
FP/pac.	10,18	15,64	7,09	17,09	8,09
Zpoždění*	27,54	21,23	22,61	21,04	22,00
TPR	85 %	41 %	89 %	50 %	88 %
PPV	72 %	45 %	80 %	47 %	77 %
FNR	15 %	59 %	11 %	50 %	12 %
FDR	28 %	55 %	20 %	53 %	23 %
F1 skóre	78 %	43 %	84 %	48 %	82 %

* průměrná doba detekce karbohydrátů od příjetí jídla v minutách

CHO - počet jídel TPR - citlivost - míra pravdivě pozitivních

TP - pravdivě pozitivní PPV - přesnost

FN - falešně negativní FNR - míra falešně negativních výsledků

FP - falešně pozitivní FDR - míra falešně pozitivních výsledků

Všechny formy detekce hran průběhu intersticiální glukózy dosahovaly citlivosti nad 85 %. Nejlepších výsledků (89 %) dosahuje detekce hran s

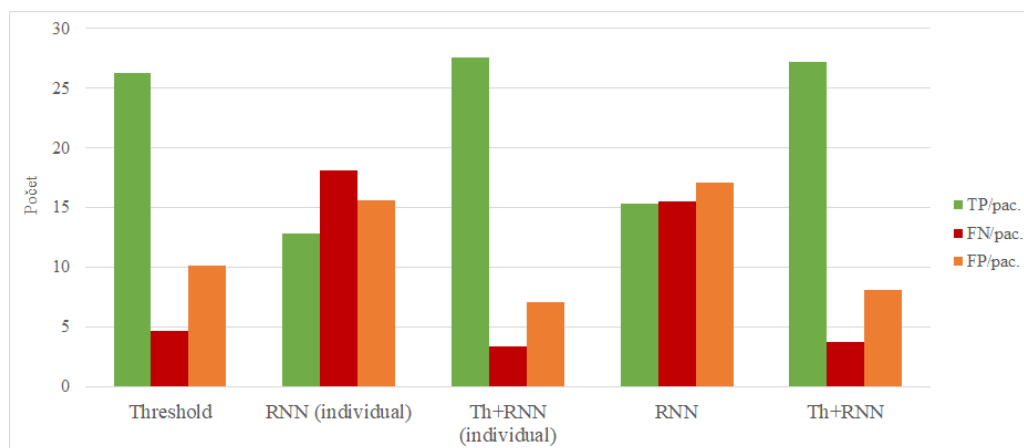
potvrzením GRU neuronovou sítí. U některých pacientů dosahovala citlivost až 95 %. Tato metoda má i nejnižší počet falešně pozitivních detekcí. Naopak samostatná neuronová síť dosahovala úspěšné detekce jen kolem 50 %.

Zpoždění v průměru 22 minut je srovnatelné, nebo i lepší, než výsledky zkoumaných algoritmů v kapitole 3.

Algoritmy vykazují značný počet falešně pozitivních detekcí. V datech se vyskytují oblasti, kdy jídlo nebylo pacientem zaznamenáno, nebo bylo zaznamenané se zpožděním. Také prudké výkyvy koncentrace cukru v krvi spojené s jinou činností mohou způsobit falešnou detekci.

V grafu na obrázku 9.2 je porovnání počtu pravdivě pozitivních, falešně negativních a falešně pozitivních detekcí na jednoho pacienta.

Obrázek 9.2: Porovnání metod detekce karbohydrátů



9.2 Detekce fyzické aktivity

Pro detekci fyzické aktivity jsou použity hodnoty srdečního tepu, počtu kroků, akcelerace, elektrodermální aktivity a jejich kombinace. Thresholdy jsou různé pro každý ukazatel a pacienta. Threshold pro srdeční tep se pohybuje od 100 do 120 tepů za minutu, počet kroků od 12 do 45 za 5 minut, velikost akcelerace od 1,1 do 1,5 m/s_2 a elektrodermální aktivita od 1 do 15 siemensů.

Pět pacientů mělo měření srdeční tep, počet kroků a elektrodermální aktivitu. Čtyři pacienti měli měřenou akceleraci a elektrodermální aktivitu. Pro tři pacienty nejsou dostupná testovací data se zadanou fyzickou aktivitou, proto nebyli zařazeni do testů.

V tabulkách 10.7 až 10.12 v příloze A jsou výsledky detekce u pacientů s měřeným srdečním tepem. V tabulkách 10.14 až 10.16 jsou výsledky detekce u pacientů s měřenou akcelerací. V tabulkách jsou i velikosti nastavených thresholdů pro jednotlivé ukazatele. V tabulce 9.2 jsou souhrnné výsledky pro srdeční tep a počet kroků, v tabulce 9.3 pro akceleraci.

Tabulka 9.2: Výsledky detekce fyzické aktivity
- srdeční tep, počet kroků, elektrodermální aktivita

	Srdeční tep	Kroky	Elektroderm. aktivita	Srdeční tep Kroky	Srdeční tep Elektroderm. aktivita	Kroky Elektroderm. aktivita
Fyz. aktivita	19	19	19	19	19	19
TP	16	17	13	15	8	10
TP/pac.	3,20	3,40	2,60	3,00	1,60	2,00
FN	3	2	6	4	11	9
FN/pac.	0,60	0,40	1,20	0,80	2,20	1,80
FP	122	73	8	28	1	2
FP/pac.	24,40	14,60	1,60	5,60	0,20	0,40
Zpoždění*	10,22	5,26	18,62	10,42	12,09	12,40
TPR	84 %	89 %	68 %	79 %	42 %	53 %
PPV	12 %	19 %	62 %	35 %	89 %	83 %
FNR	16 %	11 %	32 %	21 %	58 %	47 %
FDR	88 %	81 %	38 %	65 %	11 %	17 %
F1 skóre	20 %	31 %	65 %	48 %	57 %	65 %

Tabulka 9.3: Výsledky detekce fyzické aktivity
- akcelerace, elektrodermální aktivita

	Akcelerace	Elektrodermální aktivita	Akcelerace + Elektrodermální aktivita
Fyz. aktivita	17	17	17
TP	16	10	6
TP/pac.	4,00	2,50	1,50
FN	1	7	11
FN/pac.	0,25	1,75	2,75
FP	64	43	10
FP/pac.	16,00	10,75	2,50
Zpoždění*	25,46	22,98	24,85
TPR	94 %	59 %	35 %
PPV	20 %	19 %	38 %
FNR	6 %	41 %	65 %
FDR	80 %	81 %	63 %
F1 skóre	33 %	29 %	36 %

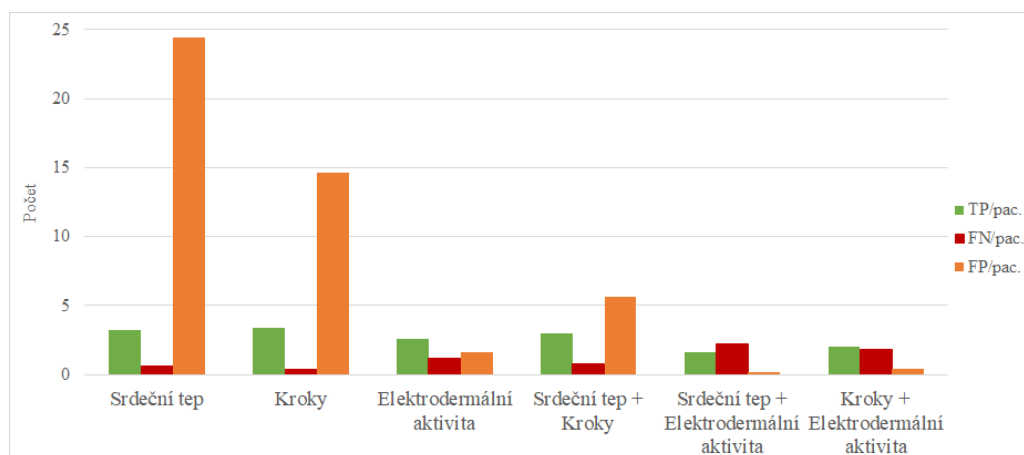
Dobré citlivosti dosahovala detekce při použití srdečního tepu, kroků a akcelerace, přičemž akcelerace dosahovala citlivosti 94 %. Nejnižší zpoždění mělo měření počtu kroků. Nejmenší citlivost má elektrodermální aktivita.

Všechny algoritmy mají vysoký počet falešně pozitivních detekcí. To je dáno tím, že měřená data jsou ovlivněna mnoha faktory, které nemusí s intenzivní fyzickou aktivitou souviset. Může se také jednat o lehčí aktivitu, kterou pacient nezadal. Kombinací více ukazatelů se podařilo značně snížit počet falešně pozitivních detekcí. Nejlépe pak vychází kombinace srdečního tepu a počtu kroků.

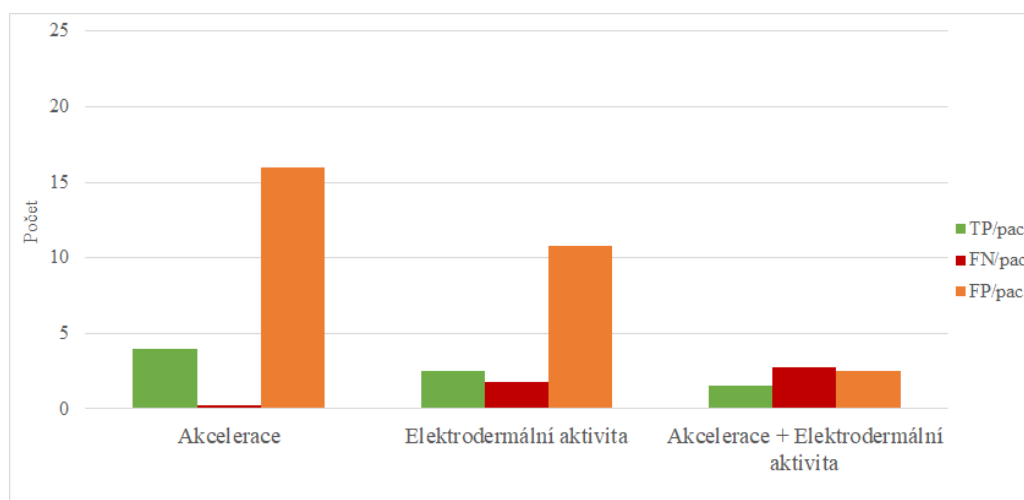
Detekce sestupné hrany neměla vliv na výsledky detekcí. Z toho vyplývá, že ve všech případech, kdy byla fyzická aktivita detekována hodnota intersticiální glukózy klesala.

V grafech na obrázcích 9.3 a 9.4 je porovnání počtu pravdivě pozitivních, falešně negativních a falešně pozitivních detekcí na jednoho pacienta.

Obrázek 9.3: Porovnání metod detekce fyzické aktivity
- srdeční tep, počet kroků, elektrodermální aktivita



Obrázek 9.4: Porovnání metod detekce fyzické aktivity
- akcelerace, elektrodermální aktivita



10 Závěr

V rámci práce byly navrženy a implementovány metody detekce příjmu karbohydrátů pomocí rekurentních neuronových sítí a detekce hran průběhu intersticiální glukózy. Nejlepších výsledků dosahovala kombinace detekce hran a rekurentní neuronové sítě, kdy citlivost detekce byla 89 %. Zpoždění detekce 22,61 minut je srovnatelné, nebo i lepší, než výsledky zkoumaných algoritmů v kapitole 3.

Metoda detekce průběhu intersticiální glukózy by se dala za použití detekce vzestupné i sestupné hrany rozšířit pro detekci glykemického indexu.

Detekce fyzické aktivity je realizována na základě pohybových dat, srdečního tepu, elektrodermální aktivity a jejich kombinací. Nejvyšší citlivosti detekce bylo dosaženo při použití pohybových dat. Detekce sestupné hrany průběhu intersticiální glukózy neměla na výsledky vliv. Přesnost detekce fyzické aktivity by se dala zvýšit použitím 3-osého akcelerometru.

Metody detekce příjmu karbohydrátů i fyzické aktivity mají vyšší počet falešně pozitivních výsledků. To je dáno tím, že zkoumané ukazatele jsou ovlivněny mnoha externími faktory. V datech se také vyskytují oblasti, kdy pacient nezašel do aplikace poctivě všechny aktivity. U detekce karbohydrátů se množství falešně pozitivních detekcí podařilo snížit metodou potvrzování detekce vyšším thresholdem nebo neuronovou sítí. U detekce fyzické aktivity se falešně pozitivní výsledky značně snížily při použití kombinace ukazatelů.

Do aplikace SmartCGMS byly implementovány celkem čtyři filtry. Savitzky-Golay filtr pro vyhlazení dat, filtr detekce karbohydrátů umožňující detekci pomocí hran průběhu intersticiální glukózy, pomocí rekurentních neuronových sítí, nebo jejich kombinace, filtr detekce fyzické aktivity a evaluační filtr pro vyhodnocení výsledků.

Parametry filtrů jsou nastavitelné, pro ukázkou bylo vytvořeno několik konfigurací. U detekčních filtrů jsou thresholdy ve formě parametrů modelu a je možné aplikaci rozšířit o solver, který by našel neoptimálnější thresholdy pro každého pacienta.

Přehled zkratk

BG	glukóza v krvi
CGM	kontinuální monitorace glukózy
CGMS	systém kontinuální monitorace glukózy
CHO	karbohydráty
IST	intersticiální glukóza
CMHE	Committed moving horizon estimation
DT	Decision tree
EL	Ensemble learning
GRU	Gated recurrent unit neuronová síť
HMM	Hidden Markov model (Skrytý Markovův model)
k-NN	k-Nearest Nighbors
LDA	lineární diskriminační analýza
LSTM	Long short-term memory neuronová síť
NB	Naive Bayes
NN	neuronová síť
OC-SVM	one-class support vectore machine klasifikátor
PCA	metoda hlavních komponent
PLGS	Predictive low glucose suspend algoritmus
QDA	kvadratická diskriminační analýza
ReLU	Rectified linear units funkce
RNN	rekurentní neuronová síť
SVM	Support vectore machine
Th	Threshold
TP	pravdivě pozitivní
TN	pravdivě negativní
FN	falešně negativní
FP	falešně pozitivní
TPR	citlivost - míra pravdivě pozitivních
PPV	přesnost
FNR	míra falešně negativních výsledků
FDR	míra falešně pozitivních výsledků
MAE	střední absolutní chyba
RMSE	střední kvadratické chyby
COD	koefficient determinace

Literatura

- [1] CHEN, H. et al. Committed Moving Horizon Estimation for Meal Detection and Estimation in Type 1 Diabetes. 2019, s. 4765–4772. doi: 10.23919/ACC.2019.8814868. Dostupné z: <https://ieeexplore.ieee.org/document/8814868>.
- [2] DEY, R. – SALEM, F. M. Gate-variants of Gated Recurrent Unit (GRU) neural networks. 2017, s. 1597–1600. doi: 10.1109/MWSCAS.2017.8053243.
- [3] DING, S. – SCHUMACHER, M. Sensor Monitoring of Physical Activity to Improve Glucose Management in Diabetic Patients: A Review. *Sensors*. 2016, 16, 4. ISSN 1424-8220. doi: 10.3390/s16040589. Dostupné z: <https://www.mdpi.com/1424-8220/16/4/589>.
- [4] FACCIOLO, S. et al. Black-box Model Identification of Physical Activity in Type-1 Diabetes Patients. 2018, s. 3910–3913. doi: 10.1109/EMBC.2018.8513378.
- [5] FATHI, A. E. et al. An Unannounced Meal Detection Module for Artificial Pancreas Control Systems. 2019, s. 4130–4135. doi: 10.23919/ACC.2019.8814932. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/8814932>.
- [6] FORLENZA, G. P. et al. Predictive Low-Glucose Suspend Reduces Hypoglycemia in Adults, Adolescents, and Children With Type 1 Diabetes in an At-Home Randomized Crossover Study: Results of the PROLOG Trial. *Diabetes Care*. 2018, 41, 10, s. 2155–2161. ISSN 0149-5992. doi: 10.2337/dc18-0771. Dostupné z: <https://care.diabetesjournals.org/content/41/10/2155>.
- [7] FRASER, M. C. – GÜNTER, P. N. – BRUCE A., M. B. Probabilistic Evolving Meal Detection and Estimation of Meal Total Glucose Appearance. *Journal of Diabetes Science and Technology*. sep 2009, 3, 5, s. 1022–1030. Dostupné z: <https://journals.sagepub.com/doi/pdf/10.1177/193229680900300505>.
- [8] *Frugally Deep* [online]. 2021. [cit.]. Dostupné z: <https://github.com/Dobiasd/frugally-deep>.
- [9] HASTIE T., F. J. T. R. *The Elements of Statistical Learning*. Springer, 2008. ISBN 978-0-387-84858-7.

- [10] HUO, Z. et al. Predicting the meal macronutrient composition from continuous glucose monitors. 2019, s. 1–4. doi: 10.1109/BHI.2019.8834488. Dostupné z: <https://ieeexplore.ieee.org/document/8834488>.
- [11] HYUNJIN, L. P. et al. A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator. *Journal of Diabetes Science and Technology*. sep 2009, 3, 5, s. 1082–1090. Dostupné z: <https://journals.sagepub.com/doi/pdf/10.1177/193229680900300511>.
- [12] ING. LUKÁŠ SMITL. Vlnková filtrace elektrokardiogramů. 2013, s. 37. Dostupné z: https://www.vutbr.cz/www_base/zav_prace_soubor_verejne.php?file_id=69903.
- [13] JOSE, J. A. – WAGGONER, T. – MANIKANDAN, S. Continuous Glucose Monitoring Prediction. 2021. Dostupné z: <https://arxiv.org/ftp/arxiv/papers/2101/2101.02557.pdf>.
- [14] KÖLLE, K. et al. Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1. *IEEE Journal of Biomedical and Health Informatics*. 2020, 24, 2, s. 594–602. doi: 10.1109/JBHI.2019.2908897. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/8679995>.
- [15] KOUTNY, T. – UBL, M. SmartCGMS as a Testbed for a Blood-Glucose Level Prediction and/or Control Challenge with (an FDA-Accepted) Diabetic Patient Simulation. *Procedia Computer Science*. 2020, 177, s. 354–362. ISSN 1877-0509. doi: <https://doi.org/10.1016/j.procs.2020.10.048>. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1877050920323164>. The 11th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2020) / The 10th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH 2020) / Affiliated Workshops.
- [16] KOUTNY, T. – UBL, M. *SmartCGMS* [online]. Medicínská informatika, KIV, FAV, ZČU, 2015. [cit. 2021/05/05]. Dostupné z: <https://diabetes.zcu.cz/smartcgms>.
- [17] LIU, F. T. – TING, K. M. – ZHOU, Z. Isolation Forest. 2008, s. 413–422. doi: 10.1109/ICDM.2008.17. Dostupné z: <https://ieeexplore.ieee.org/document/4781136>.
- [18] LUO, J. – YING, K. – BAI, J. Savitzky–Golay smoothing and differentiation filter for even number data. *Signal Processing*. 2005, 85, 7, s. 1429–1434. ISSN 0165-1684. doi: <https://doi.org/10.1016/j.sigpro.2005.02.002>.

- Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0165168405000654>.
- [19] MUDR. PAVLÍNA PIŤHOVÁ. Inzulinové režimy z klinického pohledu. *Interní medicína pro praxi*. 2010, 12, 11, s. 531–534. Dostupné z: <http://www.internimedicina.cz/pdfs/int/2010/11/02.pdf>.
 - [20] PELIKÁNOVÁ TEREZIE A BARTOŠ VLADIMÍR. *Praktická diabetologie*. Maxdorf, 2010. ISBN 978-80-7345-216-2.
 - [21] PERUŠIČOVÁ JINDRA. *Diabetes mellitus 1. typu*. Geum, 2007. ISBN 978-80-86256-49-8.
 - [22] PROF. MUDR. FRANTIŠEK SAUDEK, DRSC. *Cukrovka* [online]. PANAX Co, s.r.o. a Diabetická asociace České republiky, z.s., 2017. [cit. 2021/04/11]. Dostupné z: <https://www.cukrovka.cz/>.
 - [23] PSOTTOVÁ JANA. *Praktický průvodce cukrovkou III*. Maxdorf s. r. o, 2019. ISBN 978-80-7345-630-6.
 - [24] QI, J. et al. A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors. *IEEE Internet of Things Journal*. 2019, 6, 2, s. 1384–1393. doi: 10.1109/JIOT.2018.2846359.
 - [25] RAMKISSOON, C. M. et al. Unannounced Meals in the Artificial Pancreas: Detection Using Continuous Glucose Monitoring. *Sensors*. 2018, 18, 3. ISSN 1424-8220. doi: 10.3390/s18030884. Dostupné z: <https://www.mdpi.com/1424-8220/18/3/884>.
 - [26] RYBKA JAROSLAV. *Diabetes mellitus - komplikace a přidružená onemocnění: diagnostické a léčebné postupy*. Grada, 2007. ISBN 9788024716718.
 - [27] SAMADI, S. et al. Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data. *IEEE Journal of Biomedical and Health Informatics*. 2017, 21, 3, s. 619–627. doi: 10.1109/JBHI.2017.2677953. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/7870660>.
 - [28] *Linear and Quadratic Discriminant Analysis* [online]. scikit-learn developers, 2007-2020. [cit. 2021/05/04]. Dostupné z: https://scikit-learn.org/stable/modules/lda_qda.html.
 - [29] *Akima interpolator* [online]. The SciPy community, 2021. [cit. 2021/04/19]. Dostupné z: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.interpolate.Akima1DInterpolator.html>.

- [30] SEVIL, M. et al. Determining Physical Activity Characteristics From Wristband Data for Use in Automated Insulin Delivery Systems. *IEEE Sensors Journal*. 2020, 20, 21, s. 12859–12870. doi: 10.1109/JSEN.2020.3000772.
- [31] STENERSON, M. et al. The Impact of Accelerometer and Heart Rate Data on Hypoglycemia Mitigation in Type 1 Diabetes. 2014, s. 64–69. doi: 10.1177/1932296813516208. Dostupné z: <https://ieeexplore.ieee.org/document/8814868>.
- [32] TAI N., WONG F.S. & WEN, L. . The role of gut microbiota in the development of type 1, type 2 diabetes mellitus and obesity. *Rev Endocr Metab Disord*. 2015, 16, s. 55–65. doi: 10.1007/s11154-015-9309-0. Dostupné z: <https://link.springer.com/article/10.1007/s11154-015-9309-0>.
- [33] TAPIA, E. M. et al. Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor. 2007, s. 37–40. doi: 10.1109/ISWC.2007.4373774.
- [34] TOWNSEND, C. – SERON, M. M. – GOODWIN, G. C. Characterisation of Optimal Responses to Pulse Inputs in the Bergman Minimal Model. *IFAC-PapersOnLine*. 2017, 50, 1, s. 15163–15168. ISSN 2405-8963. doi: <https://doi.org/10.1016/j.ifacol.2017.08.2263>. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S2405896317330562>. 20th IFAC World Congress.
- [35] TURKSOY, K. et al. Real-time insulin bolusing for unannounced meals with artificial pancreas. *Control Engineering Practice*. 2017, 59, s. 159–164. ISSN 0967-0661. doi: <https://doi.org/10.1016/j.conengprac.2016.08.001>. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S0967066116301642>.
- [36] UBL, M. – KOUTNY, T. SmartCGMS as an Environment for an Insulin-Pump Development with FDA-Accepted In-Silico Pre-Clinical Trials. *Procedia Computer Science*. 2019, 160, s. 322–329. ISSN 1877-0509. doi: <https://doi.org/10.1016/j.procs.2019.11.084>. Dostupné z: <https://www.sciencedirect.com/science/article/pii/S1877050919317843>. The 10th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN-2019) / The 9th International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare (ICTH-2019) / Affiliated Workshops.
- [37] *Diabetes* [online]. World Health Organization, 2021. [cit. 2021/04/11]. Dostupné z: <https://www.who.int/health-topics/diabetes>.

- [38] YU, Y. et al. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Computation*. 07 2019, 31, 7, s. 1235–1270. ISSN 0899-7667. doi: 10.1162/neco__a__01199. Dostupné z: https://doi.org/10.1162/neco_a_01199.
- [39] ZHENG, F. et al. Unannounced Meal Detection for Artificial Pancreas Systems Using Extended Isolation Forest. 2020, s. 5892–5895. doi: 10.1109/EMBC44109.2020.9176856. Dostupné z: <https://ieeexplore.ieee.org/abstract/document/9176856>.

Příloha A: Tabulky

Detekce karbohydrátů

Tabulka 10.1: Threshold

Pacient	540	544	552	559	563	570	575	584	588	591	596
CHO	22	40	12	23	24	29	42	13	35	41	59
TP	20	34	8	21	21	26	35	8	31	36	49
TP potvr.	9	29	3	12	14	14	22	4	18	17	25
Potvrzeno	41 %	73 %	25 %	52 %	58 %	48 %	52 %	31 %	51 %	41 %	42 %
FN	2	6	4	2	3	3	7	5	4	5	10
FP	9	5	1	10	16	15	11	10	11	16	8
Zpoždění*	35,67	22,75	7,41	33,87	26,5	43,24	21,62	33,69	23,97	27,34	26,93
TPR	91 %	85 %	67 %	91 %	88 %	90 %	83 %	62 %	89 %	88 %	83 %
PPV	69 %	87 %	89 %	68 %	57 %	63 %	76 %	44 %	74 %	69 %	86 %

Tabulka 10.2: GRU (individual)

Pacient	540	544	552	559	563	570	575	584	588	591	596
CHO	22	40	12	23	24	29	42	13	35	41	59
TP	15	27	7	0	15	11	30	2	0	6	28
FN	7	13	5	23	9	18	12	11	35	35	31
FP	5	33	3	2	39	17	31	1	1	13	27
Zpoždění*	38,25	44,08	33,67	-	13,04	24,90	23,19	9,77	0	7,27	18,14
TPR	68 %	68 %	58 %	0 %	63 %	38 %	71 %	15 %	0 %	15 %	47 %
PPV	75 %	45 %	70 %	0 %	28 %	39 %	49 %	67 %	0 %	32 %	51 %

Tabulka 10.3: Threshold + GRU (individual)

Pacient	540	544	552	559	563	570	575	584	588	591	596
CHO	22	40	12	23	24	29	42	13	35	41	59
TP	20	36	9	21	21	27	40	8	31	38	52
TP potvr.	10	13	2	0	13	6	20	0	0	2	21
Potvrzeno	45 %	33 %	17 %	0 %	54 %	21 %	48 %	0 %	0 %	5 %	36 %
FN	2	4	3	2	3	2	2	5	4	3	7
FP	3	17	2	1	22	3	13	0	0	6	11
Zpoždění*	27,26	15,00	16,12	33,89	10,25	40,76	17,40	17,54	23,97	27,51	19,05
TPR	91 %	90 %	75 %	91 %	88 %	93 %	95 %	62 %	89 %	93 %	88 %
PPV	87 %	68 %	82 %	95 %	49 %	90 %	75 %	100 %	100 %	86 %	83 %

Tabulka 10.4: GRU

Pacient	540	544	552	559	563	570	575	584	588	591	596
CHO	22	40	12	23	24	29	42	13	35	41	59
TP	20	20	7	9	4	7	25	7	9	28	33
TP potvr.	20	20	7	9	4	7	25	7	9	28	33
Potvrzeno	90 %	50 %	58 %	39 %	17 %	24 %	60 %	54 %	26 %	68 %	56 %
FN	2	20	5	14	20	22	17	6	26	13	26
FP	15	23	8	19	13	8	28	10	14	31	19
Zpoždění*	24,68	26,58	23,26	11,17	7,92	12,72	34,40	11,62	21,77	33,21	24,10
TPR	91 %	50 %	58 %	39 %	17 %	24 %	60 %	54 %	26 %	68 %	56 %
PPV	57 %	47 %	47 %	32 %	24 %	47 %	47 %	41 %	39 %	47 %	63 %

Tabulka 10.5: Threshold + GRU

Pacient	540	544	552	559	563	570	575	584	588	591	596
CHO	22	40	12	23	24	29	42	13	35	41	59
TP	21	34	9	21	21	27	38	8	31	37	52
TP potvr.	14	14	6	9	3	3	15	6	6	20	20
Potvrzeno	64 %	35 %	50 %	39 %	13 %	10 %	38 %	46 %	17 %	49 %	34 %
FN	1	6	3	2	3	2	4	5	4	4	7
FP	7	12	5	7	6	2	12	5	10	15	8
Zpoždění*	16,88	22,25	11,53	27,57	26,50	36,38	19,57	19,85	22,69	16,07	22,68
TPR	95 %	85 %	75 %	91 %	88 %	93 %	90 %	62 %	89 %	90 %	88 %
PPV	75 %	74 %	64 %	75 %	78 %	93 %	76 %	62 %	76 %	71 %	87 %

Tabulka 10.6: Souhrnné výsledky

	Threshold	RNN (individual)	Th+RNN (individual)	RNN	Th+RNN
CHO	340	340	340	340	340
TP	289	141	303	169	299
TP/pac.	26,27	12,82	27,55	15,36	27,18
TP potvr.	167	-	87	169	116
Potvrzeno	47 %	-	24 %	49 %	36 %
FN	51	199	37	171	41
FN/pac.	4,64	18,09	3,36	15,55	3,73
FP	112	172	78	188	89
FP/pac.	10,18	15,64	7,09	17,09	8,09
Zpoždění*	27,54	21,23	22,61	21,04	22,00
TPR	85 %	41 %	89 %	50 %	88 %
PPV	72 %	45 %	80 %	47 %	77 %
FNR	15 %	59 %	11 %	50 %	12 %
FDR	28 %	55 %	20 %	53 %	23 %
F1 skóre	78 %	43 %	84 %	48 %	82 %

Detekce fyzické aktivity

Srdeční tep a počet kroků

Tabulka 10.7: Srdeční tep

Pacient	559	563	570	588	591
Threshold	110	110	100	120	100
Fyz. aktivita	7	3	2	3	4
TP	6	2	2	2	4
FN	1	1	0	1	0
FP	28	59	29	4	2
Zpoždění*	21,57	3,34	13,50	7,67	5,00
TPR	86 %	67 %	100 %	67 %	100 %
PPV	18 %	3 %	6 %	33 %	67 %

Tabulka 10.8: Počet kroků

Pacient	559	563	570	588	591
Threshold	45	45	12	45	28
Fyz. aktivita	7	3	2	3	4
TP	7	2	2	3	3
FN	0	1	0	0	1
FP	7	12	43	2	9
Zpoždění*	2,86	3,00	13,50	6,67	0,25
TPR	100 %	67 %	100 %	100 %	75 %
PPV	50 %	14 %	4 %	60 %	25 %

Tabulka 10.9: Elektrodermální aktivita

Pacient	559	563	570	588	591
Threshold	15	10	10	3	10
Fyz. aktivita	7	3	2	3	4
TP	5	2	1	2	3
FN	2	1	1	1	1
FP	0	3	0	2	3
Zpoždění*	18,00	3,67	25,50	31,67	14,25
TPR	71 %	67 %	50 %	67 %	75 %
PPV	100 %	40 %	100 %	50 %	50 %

Tabulka 10.10: Srdeční tep + počet kroků

Pacient	559	563	570	588	591
Fyz. aktivita	7	3	2	3	4
TP	6	2	2	2	3
FN	1	1	0	1	1
FP	1	6	20	1	0
Zpoždění*	21,57	1,34	16,00	7,67	5,50
TPR	86 %	67 %	100 %	67 %	75 %
PPV	86 %	25 %	9 %	67 %	100 %

Tabulka 10.11: Srdeční tep + elektrodermální aktivita

Pacient	559	563	570	588	591
Fyz. aktivita	7	3	2	3	4
TP	3	2	1	0	2
FN	4	1	1	3	2
FP	0	1	0	0	0
Zpoždění*	19,14	3,33	25,50	0,00	12,50
TPR	43 %	67 %	50 %	-	50 %
PPV	100 %	67 %	100 %	-	100 %

Tabulka 10.12: Počet kroků + elektrodermální aktivita

Pacient	559	563	570	588	591
Fyz. aktivita	7	3	2	3	4
TP	5	2	1	0	2
FN	2	1	1	3	2
FP	0	1	0	0	1
Zpoždění*	18,00	5,00	27,50	0,00	11,50
TPR	71 %	67 %	50 %	-	50 %
PPV	100 %	67 %	100 %	-	67 %

Tabulka 10.13: Souhrnné výsledky

	Srdeční tep	Kroky	Elektroderm. aktivita	Srdeční tep Kroky	Srdeční tep Elektroderm. aktivita	Kroky Elektroderm. aktivita
Fyz. aktivita	19	19	19	19	19	19
TP	16	17	13	15	8	10
TP/pac.	3,20	3,40	2,60	3,00	1,60	2,00
FN	3	2	6	4	11	9
FN/pac.	0,60	0,40	1,20	0,80	2,20	1,80
FP	122	73	8	28	1	2
FP/pac.	24,40	14,60	1,60	5,60	0,20	0,40
Zpoždění*	10,22	5,26	18,62	10,42	12,09	12,40
TPR	84 %	89 %	68 %	79 %	42 %	53 %
PPV	12 %	19 %	62 %	35 %	89 %	83 %
FNR	16 %	11 %	32 %	21 %	58 %	47 %
FDR	88 %	81 %	38 %	65 %	11 %	17 %
F1 skóre	20 %	31 %	65 %	48 %	57 %	65 %

Akcelerace

Tabulka 10.14: Akcelerace

Pacient	544	552	584	596
Threshold	1,1	1,5	1,5	1,1
Fyz. aktivita	2	2	1	12
TP	2	2	1	11
FN	0	0	0	1
FP	31	1	4	28
Zpoždění*	33,50	12,50	3,00	52,83
TPR	100 %	100 %	100 %	92 %
PPV	6 %	67 %	20 %	28 %

Tabulka 10.15: Elektrodermální aktivita

Pacient	544	552	584	596
Threshold	15	2	35	1
Fyz. aktivita	2	2	1	12
TP	2	2	1	5
FN	0	0	0	7
FP	22	2	12	7
Zpoždění*	23,50	28,00	12,00	28,42
TPR	100 %	100 %	100 %	42 %
PPV	8 %	50 %	8 %	42 %

Tabulka 10.16: Akcelerace + elektrodermální aktivita

Pacient	544	552	584	596
Fyz. aktivita	2	2	1	12
TP	2	1	1	2
FN	0	1	0	10
FP	9	0	1	0
Zpoždění*	33,50	35,00	18,00	12,90
TPR	100 %	50 %	100 %	17 %
PPV	18 %	100 %	50 %	100 %

Tabulka 10.17: Souhrnné výsledky

	Akcelerace	Elektrodermální aktivita	Akcelerace + Elektrodermální aktivita
Fyz. aktivita	17	17	17
TP	16	10	6
TP/pac.	4,00	2,50	1,50
FN	1	7	11
FN/pac.	0,25	1,75	2,75
FP	64	43	10
FP/pac.	16,00	10,75	2,50
Zpoždění*	25,46	22,98	24,85
TPR	94 %	59 %	35 %
PPV	20 %	19 %	38 %
FNR	6 %	41 %	65 %
FDR	80 %	81 %	63 %
F1 skóre	33 %	29 %	36 %

Příloha B

Adresářová struktura

- Aplikace_a_knihovny
 - detection
 - * src - zdrojové soubory filtrů pro SmartCGMS
 - * CMakeLists.txt
 - lib - knihovny
 - scripts - python skripty
 - setup - konfigurační soubory
 - detection.dll - knihovna filtrů pro SmartCGMS
- Poster
- Text_prace
 - src - zdrojové tex soubory
 - dp.pdf
- Vstupni_data
 - model - natrénované keras modely GRU sítě
(vstupní data pacientů není možné přiložit z důvodu autorských práv)
- Vysledky - vysledky provedených experimentů

Knihovny, kompilace, spuštění

SmartCGMS je dostupný na <https://diabetes.zcu.cz/smartcgms>. Knihovna je kompilována v C++ 17. Pro kompilaci jsou nutné header soubory SmartCGMS a knihovny VisualLeakDetector, frugally-deep, Eigen, Functional-Plus a json. K dispozici je CMake skript. Vytvořená dynamická knihovna `detection.dll` se umístí do složky `filters`. Grafické rozhraní SmartCGMS se spustí programem `gpredict3.exe`, konzolová verze programem `console3.exe`.

Nastavení filtrů

Savitzky-Golay filtr

Filtr pro vyhlazení dat.

- Signal - zdrojový signál pro vyhlazení
- Window size - velikost okna Savitzky-Golay filtru
- Degre - stupeň polynomu

Filtr posílá vyhlazená data v signálu Savgol signal. V případě použití více filtrů je nutné výstupní signál přemapovat.

CHO detection

Filtr detekce příjmu karbohydrátů.

- Signal - detekovaný signál
- Window size - velikost klouzavého okénka
- Detect edges - detekce vzestupných hran
- Detect descending edges - detekce sestupných hran
- Rise threshold - threshold pro určení míry stoupání/klesání v čase
- Use RNN - použití rekurentní neuronové sítě
- RNN model file path - cesta k souboru s natrénovaným keras modelem převedeným do formátu pro frugally-deep
- RNN threshold - threshold detekce neuronovou sítí
- Thresholds
 - Threshold Low - threshold malé změny IST
 - Weight Low - váha malé změny IST
 - Threshold High - threshold velké změny IST
 - Weight High - váha velké změny IST

Filtr posílá aktivační funkce a detekované karohydráty. Příklad konfigurace detekce hran průběhu intersticiální glukózy je v souboru `setup/setup_th.ini`, příklad neuronové sítě v souboru `setup/setup_gru.ini`.

PA detection

Filtr detekce fyzické aktivity.

- Heartbeat - detekce podle srdečního tepu
- Steps - detekce podle počtu kroků
- Acceleration - detekce podle hodnoty akcelerace
- Electrodermal activity - detekce podle elektrodermální aktivity
v Mean - použití průměru za časové okno
- Window size - mean - velikost klouzavého okénka pro spočítání průměru (v případě velikosti okna 1 je průměr rovná aktuální hodnotě)
- Detect IST edges - potvrzení detekce sestupnou hranou dat IST
- Signal - detekovaný signal
- Window size - edges - velikost klouzavého okénka pro detekci hran
- Thresholds - threshold ukazatelů pro detekci a thresholdy a váhy pro detekci hran

Filtr posílá detekovanou fyzickou aktivitu. Příklad konfigurace s měřeným srdečním tepem a počtem kroků je v konfiguračním souboru `setup/setup_bpm.ini`. Příklad konfigurace s akcelerací a potvrzováním pomocí detekce sestupné hrany je v konfiguračním souboru `setup/setup_acc.ini`.

Evaluation

Filtr pro vyhodnocení výsledků.

- Reference signal - referenční signál
- Detected signal - detekovaný signál
- Max detection delay - maximální zpoždění, které může mít detekovaný signál oproti referenčnímu
- False positive cooldown - cooldown po detekci falešně pozitivního signálu, než je započítán další
- Late detection delay - čas před referenčním signálem, kdy se bude detekce počítat jako pravdivě pozitivní
- Min reference count - minimální počet referenčních signálů za den

Filtr na konci běhu simulace posílá info s naměřenými statistikami počtu referenčních signálů TP, potvrzené TP, FN, FP, zpoždění detekce a zpoždění potvrzení.

Python skripty

Skripty umožňují transformaci a modifikaci dat BGLP, trénování rekurentní neuronové sítě a analýzu metod detekce karbohydrátů a fyzické aktivity. Pro spuštění skriptů je nutné mít nainstalovaný Python 3.7 nebo 3.8. Python 3.9 není podporovaný z důvodu nepodporované kompatibility s knihovnou TensorFlow. Dále je nutné mít nainstalované tyto balíčky: numpy, pandas, scipy, sklearn, tensorflow, tabulate, matplotlib, sweetviz.

V souboru `examples.py` jsou příklady práce se skripty:

- `lg.load_log()` - transformace dat z logu do csv
- `ld.load_data()` - modifikace dat pro detekci karbohydrátů nebo fyzické aktivity
- `cho.rnn()` - trénování RNN
- `cho.lda()` - LDA/GDA
- `cho.threshold()` - detekce hran IST
- `pa.get_feature()` - spočtení vlastností (průměr, medián, směrodatná odchylka, rozdíl kvartilů)
- `pa.export_features()` - export vlastností do csv
- `pa.ML()` - test metod strojového učení pro detekci fyzické aktivity

Trénování neuronové sítě

Pro natrénování keras modelu rekurentní neuronové sítě je připraven skript `train_rnn.py`. Ten se spouští příkazem `python train_rnn.py<type> <options> [values]`, kdy `type` je typ neuronové `-gru` nebo `-lstm` (povinný parametr) a `options` jsou volitelné parametry:

`-f` název logu pacienta (* pro natrénování modelu na všech logách v adresáři)
`-i` adresář se vstupními logy (defaultně `data/`)
`-o` adresář pro uložení natrénovaného modelu (defaultně `model/`)
`-h` nápověda

Příklad natrénování GRU pro všechny pacienty: `python train_rnn.py -gru`

`-f * -i data/ -o model/` Příklad natrénování LSTM pro pacienta 575:

`python train_rnn.py -lstm -f 575-ws-training.log -i data/ -o model/`