

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

Diplomová práce

Detekce vybraných aktivit diabetického pacienta 1. typu

Místo této strany bude
zadání práce.

Prohlášení

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

V Plzni dne 31. května 2021

Bc. David Pivovar

Abstract

The text of the abstract (in English). It contains the English translation of the thesis title and a short description of the thesis.

Abstrakt

Text abstraktu (česky). Obsahuje krátkou anotaci (cca 10 řádek) v češtině. Budete ji potřebovat i při vyplňování údajů o bakalářské práci ve STAGu. Český i anglický abstrakt by měly být na stejné stránce a měly by si obsahem co možná nejvíce odpovídat (samozřejmě není možný doslovný překlad!).

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	27
3.4	Metody využívající neuronové sítě	35
3.5	Porovnání jednotlivých metod	37
4	Analýza metod detekce fyzické aktivity	39
4.1	•	39
5	SmartCGMS	40
6	Detekce příjmu karbohydrátů	43
6.1	Příprava dat	44
6.2	Rekurentní neuronové sítě	46
6.2.1	Long short-term memory	47
6.2.2	Gated recurrent unit	48
6.2.3	Model	49
6.3	Lineární a kvadratická diskriminační analýza	51
6.3.1	Teorie	51
6.3.2	Množina prvků	52
6.3.3	Časové okno	52
6.4	Detekce hran průběhu intersticiální glukózy	53
6.5	Implementace do SmartCGMS	55

7	Detekce fyzické aktivity	58
8	Závěr	59
	Přehled zkratk	60
	Literatura	61
	Příloha A: Adresářová struktura	65
	Příloha B: Návod na spuštění	66

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.

V mé práci se budu zabývat možnostmi detekce příjmu karbohydrátů a fyzické aktivity z dat senzoru CGMS. Následně tyto metody implementuji 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. [20, 26]

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. [20, 26]

Celosvětově je Světovou zdravotnickou organizací evidováno více než 422 milionů lidí s diabetem [30].

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 upozorování 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, sníží riziko komplikací spojených s diabetem. [20]

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 jídle	5-7,5 (pod 9) mmol/l
krvní 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 [19]

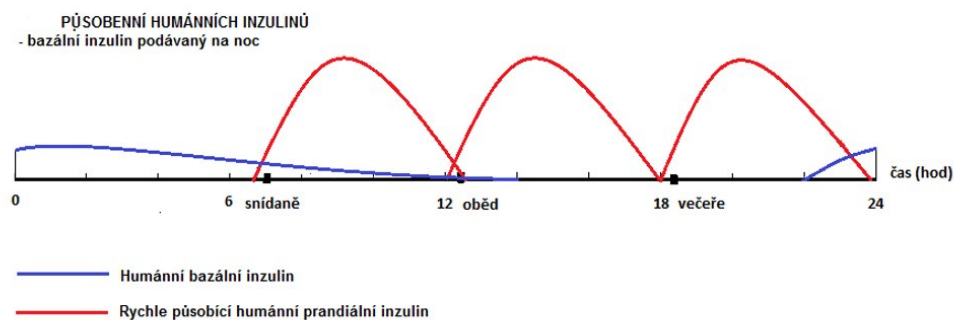
2.2.1 Inzulinové režimy

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. [16, 22]

Obrázek 2.1: Působení inzulinu



Zdroj: www.cukrovka.cz [19]

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. [18, 22]

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.

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. [17]

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. [18]

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. Takovým systémem je například 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 [27] 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 [28].

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. [28] 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 :

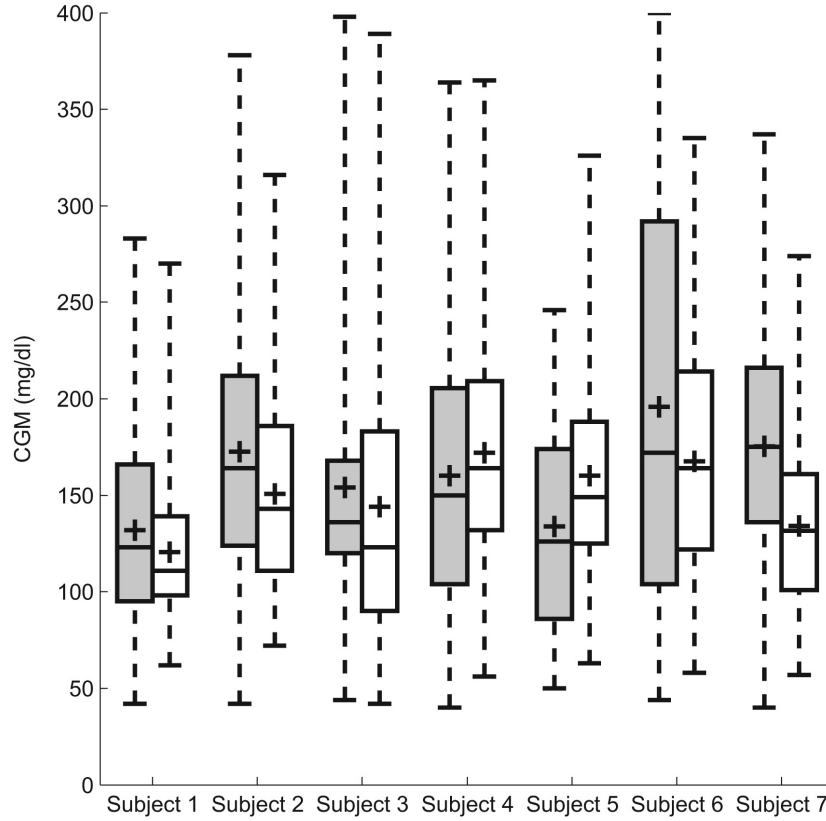
$$R_a(t) = \frac{C(t)}{V\tau^2}te^{-\frac{t}{\tau}}$$

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.

Na obrázku 3.1 je porovnání obou způsobů. Experiment ukázal, že výsledky jsou statisticky ekvivaletní (hladina významnosti $p=0.05$).

Obrázek 3.1: Porovnání řízení aplikace inzulinu pacientem a algoritmem pro detekci karbohydrátů

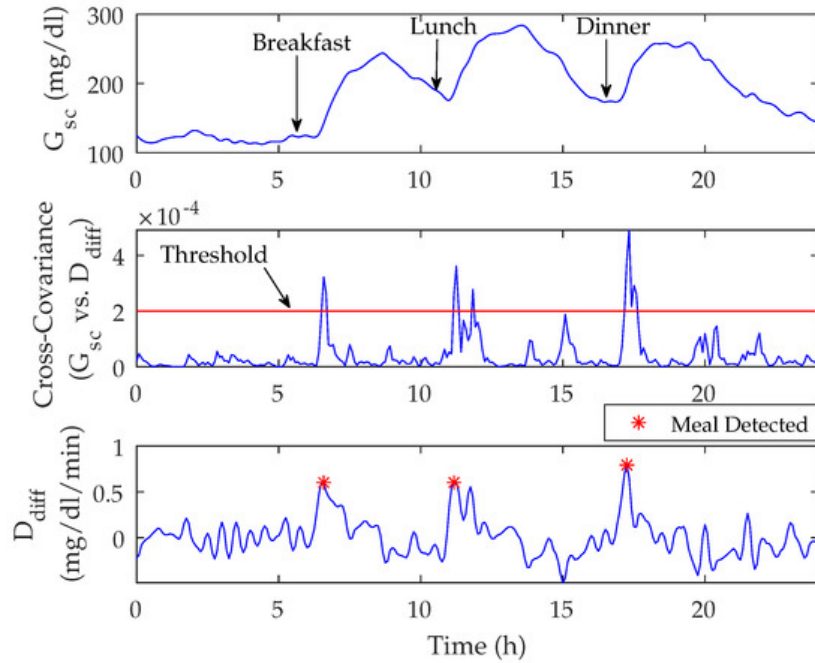


Zdroj: Real-time insulin bolusing for unannounced meals with artificial pancreas [28]

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

Autoři Ramkisson et al. [21] 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 3.2: 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 [21]

Úspěšnost detekce jednotlivých oken je v tabulce na obrázku 3.3, zpoždění detekce vzhledem k množství přijatých karbohydrátů v tabulce na obrázku 3.4.

Obrázek 3.3: Úspěšnost detekce

Sensitivity (%)	Δ Glucose (mg/dL)	Detection Time (min)	TP	FP	FN	FP/day
Highest Sensitivity (window = 3; threshold = 0.000039)						
92 \pm 3	6 \pm 13	31 \pm 16	1375 \pm 38	719 \pm 111	126 \pm 38	1 \pm 1
92 (87, 96)	5 (-8, 17)	25 (10, 60)	1374 (1306, 1433)	716 (561, 897)	126 (67, 194)	1 (0, 3)
Trade-Off (window = 9; threshold = 0.00019)						
82 \pm 4	19 \pm 9	38 \pm 14	1232 \pm 56	100 \pm 23	268 \pm 56	0.2 \pm 0.5
83 (76, 87)	19 (6, 31)	35 (25, 65)	1242 (1139, 1310)	96 (72, 154)	259 (190, 361)	0 (0, 1)
Low False Positive (window = 12; threshold = 0.00082)						
54 \pm 9	42 \pm 18	43 \pm 11	813 \pm 130	11 \pm 9	687 \pm 130	0.02 \pm 0.2
52 (44, 71)	45 (10, 56)	40 (30, 65)	776 (663, 1060)	10 (5, 34)	724 (440, 837)	0 (0, 0)

Zdroj: Unannounced Meals in the Artificial Pancreas: Detection Using Continuous Glucose Monitoring [21]

Obrázek 3.4: Zpoždění detekce

			Carbohydrates (grams)		
			20–40	40–80	80–120
High Sensitivity window = 3 threshold = 0.000039	Detection Time (min)	Mean	39 ± 22	31 ± 16	25 ± 13
		Median	30 (10, 90)	30 (10, 60)	25 (0, 50)
	Δ Glucose (mg/dL)	Mean	5 ± 14	5 ± 10	7 ± 15
		Median	5 (-13, 23)	5 (-7, 16)	6 (-6, 15)
	Sensitivity (%)	Mean	74 ± 6	93 ± 3	99 ± 1
		Median	72 (67, 87)	94 (87, 96)	99 (97, 100)
Trade-Off window = 9 threshold = 0.00019	Detection Time (min)	Mean	46 ± 18	40 ± 16	32 ± 15
		Median	45 (25, 80)	35 (25, 65)	30 (0, 55)
	Δ Glucose (mg/dL)	Mean	18 ± 11	17 ± 9	20 ± 9
		Median	19 (0, 32)	18 (0, 28)	20 (8, 32)
	Sensitivity (%)	Mean	49 ± 9	84 ± 4	96 ± 2
		Median	47 (39, 64)	86 (76, 89)	97 (94, 98)
Low False Positive window = 12 threshold = 0.00082	Detection Time (min)	Mean	45 ± 11	44 ± 8	41 ± 11
		Median	50 (14, 71)	45 (30, 65)	40 (30, 60)
	Δ Glucose (mg/dL)	Mean	29 ± 42	42 ± 18	42 ± 16
		Median	43 (-46, 57)	45 (12, 57)	44 (16, 56)
	Sensitivity (%)	Mean	8 ± 7	49 ± 12	83 ± 7
		Median	6 (2, 26)	45 (36, 71)	82 (73, 93)

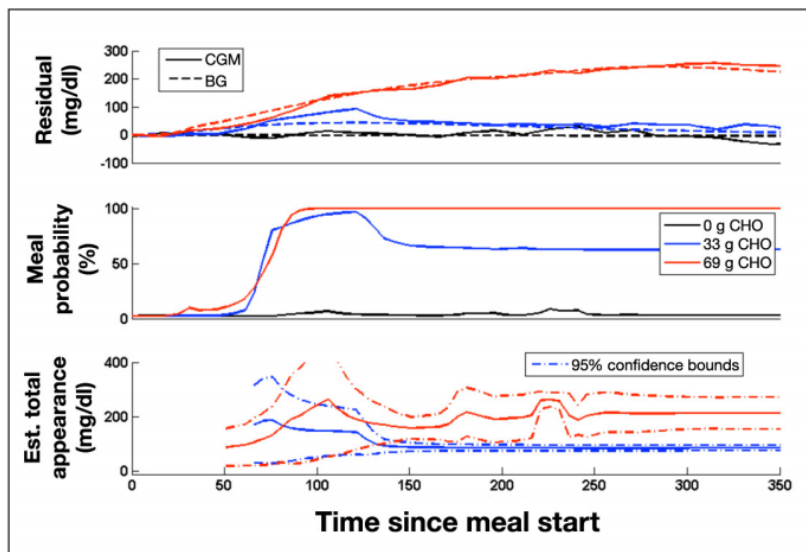
Zdroj: Unannounced Meals in the Artificial Pancreas: Detection Using Continuous Glucose Monitoring [21]

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

V této práci Fraser et al. [4] použili inzulino-glukózový model, který udává rychlost změny intersticiální glukózy závislé na působení inzulinu 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 3.5 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 %.

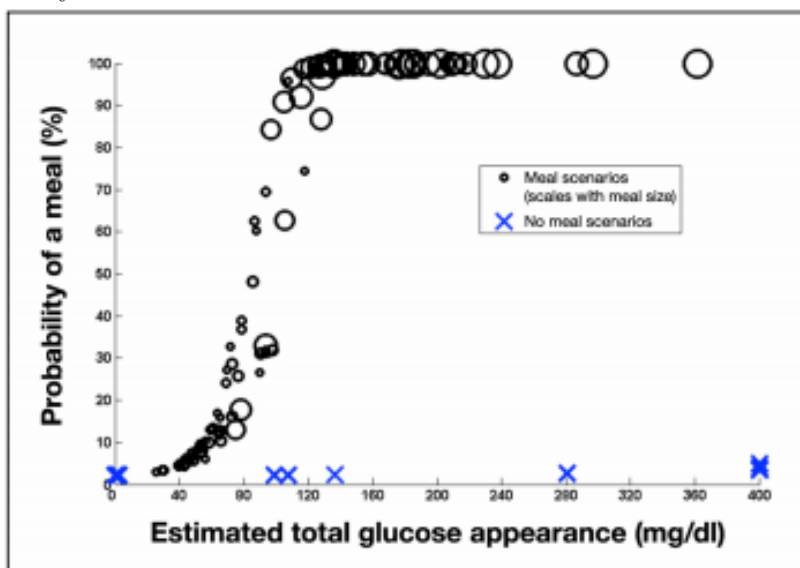
Obrázek 3.5: Detekce karbohydrátů v čase



Zdroj: Probabilistic Evolving Meal Detection and Estimation of Meal Total Glucose Appearance [4]

Experiment byl proveden pro 99 scénářů. Pro 11 scénářů bez jídla nebyly žádné falešně pozitivní detekce. Pro scénáře, kdy byly přijaté karbohydráty, byl algoritmus úspěšný podle množství přijatých karbohydrátů, jak je vidět na obrázku 3.6.

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



Zdroj: Probabilistic Evolving Meal Detection and Estimation of Meal Total Glucose Appearance [4]

3.2.4 An Unannounced Meal Detection Module for Artificial Pancreas Control Systems

Fathi et al. [3] 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 3.7.

Obrázek 3.7: 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* [3]

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

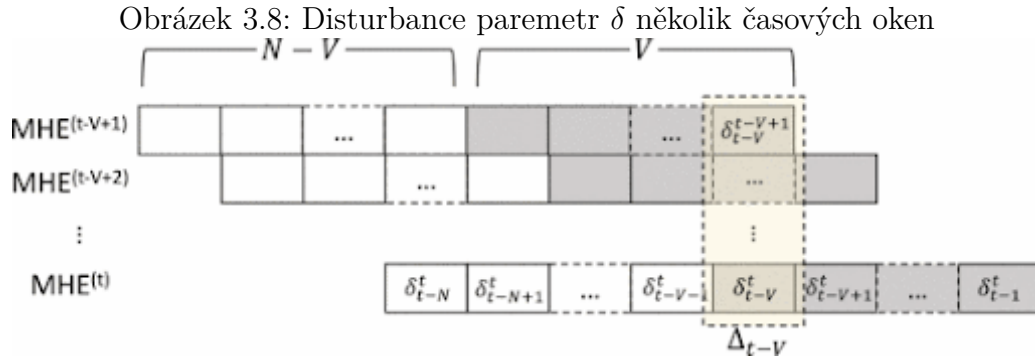
Committed Moving Horizon Estimation (CMHE) [1] umožňuje predikovat čas a velikost nehlášených jídel. Metoda je založena na Moving Horizon Estimation (MHE). MHE je metoda odhadů omezených stavů nelineárního diskrétního modelu, která predikuje sekvenci stavových proměnných a poruch δ (disturbance) tak, 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. Odhad v čase $t+N$ zase nezohledňuje vývoj glukózy v minulosti a zároveň je k dispozici s velkým zpožděním. Tento problém řeší CMHE, který nebere jednu disturbance hodnotu, ale kompromis mezi predikovanými a naměřenými hodnotami. Výsledná hodnota disturbance parametru v čase Δ_{t-V} je vypočítána jako vážený průměr odhadů v daný čas z předchozích iterací včetně (obrázek 3.8):

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

kdy V je commitment level, $V < N$, δ je disturbance parametr, $W(i)$ jsou váhy prioritizující odhady mezi predikovanými a naměřenými hodnotami.

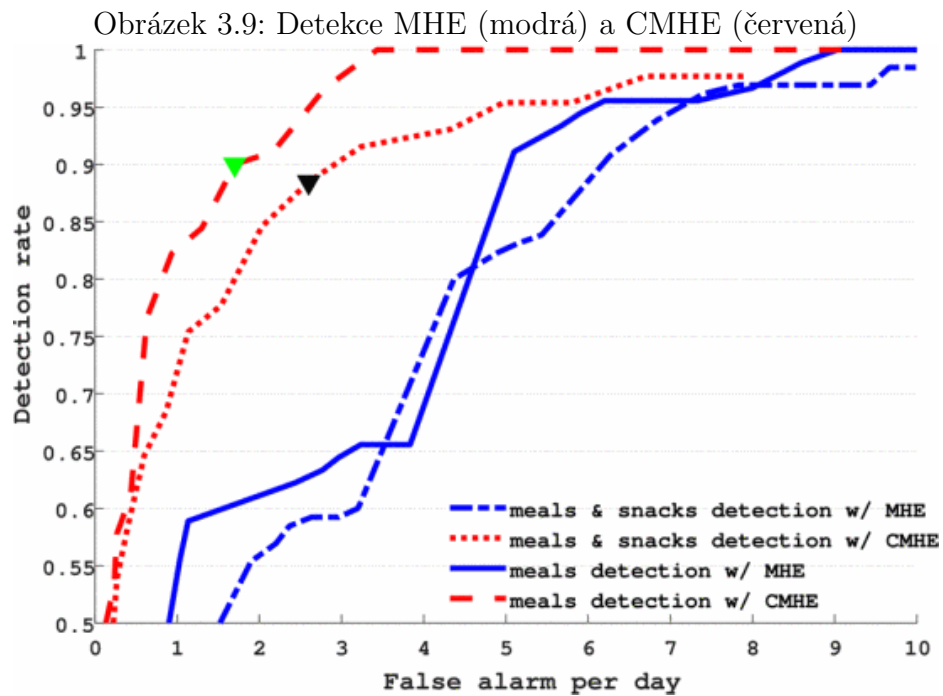


Zdroj: Committed Moving Horizon Estimation for Meal Detection and Estimation in Type 1 Diabetes [1]

CMHE instance ($\min C_t$) 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 ale spou 80 % času w .

V provedeném experimentu je měření prováděno každou minutu, časové okno $N = 180$, commitment level $V = 40$, $w = 5$ minut. Na obrázku 3.9 je srovnání MHE a CMHE v závislosti na nastaveném thresholdu. V tabulce 3.1 jsou výsledky experimentu pro optimálně nastavený threshold (černý trojúhelník). Metoda dosahuje 100 % úspěšnost detekce pro hlavní jídla, celkových 88.5 % je způsobeno menšími jídly, pro jejichž lepší detekci by

musel být snížen threshold. To by ale vedlo ke zvýšení falešných detekcí, kterých je v průměru 2.6 za den. Průměrná doba detekce je 20 minut [1].



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]

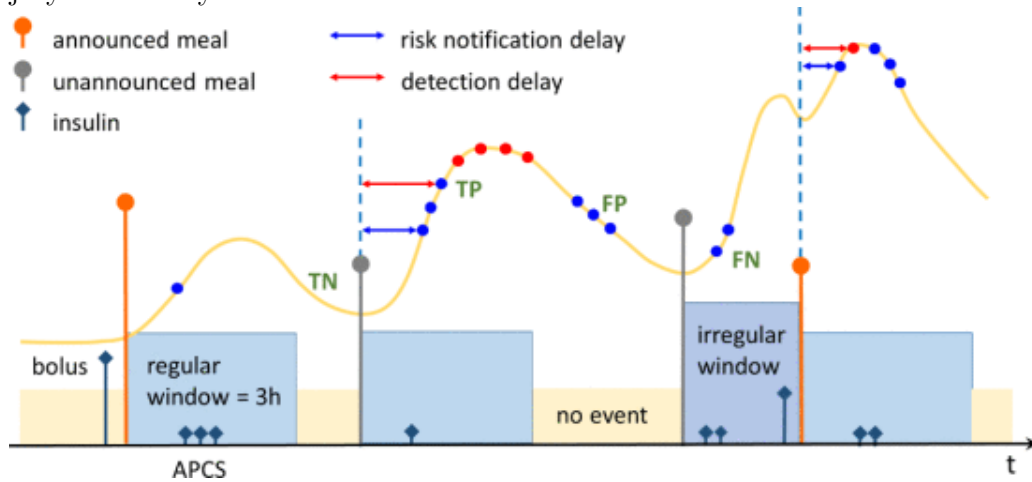
3.3 Data-driven metody

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 [14].

Vlastnosti, které slouží jako vstup pro Extended isolation forest [32] jsou derivace naměřené glukózy v plazmě \mathbf{dGp} , zadané karbohydráty (pokud byly zadány) a podaný inzulin. Pro výpočet derivace glukózy v čase t $\mathbf{dGp}(t)$ je potřeba znát naměřenou glukózu v čase $t+1$ $\mathbf{Gp}(t+1)$. 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 3.10. Notifikace je vznešena pokud se objeví jedna anomálie s větší mírou rizika (červené), nebo 3 anomálie s menší mírou rizika (modré).

Obrázek 3.10: Úspěšnost detekce vzhledem hladině glukózy a množství přijatých karbohydrátů



Zdroj: *Unannounced Meal Detection for Artificial Pancreas Systems Using Extended Isolation Forest* [32]

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 [32].

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

Hyunjin et al. [8] použili hlasovací schéma pro detekci karbohydrátů. Algoritmus je založen na kontinuálním pozorování první a druhé derivace koncentrace glukózy, které při splnění kritérií vyše sérii impulsů (až 15 impulsů během 30ti minut). Tato metoda je založena na vzorkování po jedné minutě.

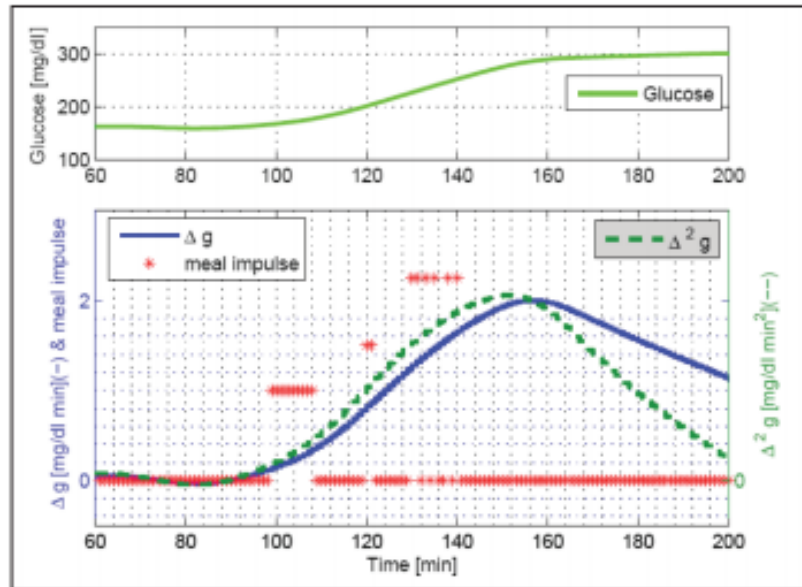
Impulsy jsou vyslány při překročení thresholdu. Pro nastavení thresholdu se vychází z toho, že jedno jídlo způsobí nárůst rychlosti výskytu glukózy v krvi dG (první derivace) 0-2 mg/dl/min, druhé derivace d_2G 0-0,02 mg/dl/min². Thresholdy jsou proto nastaveny na 0, 0.5, 1.25, 1.8 pro dG a 0, 0.005,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

kud se hodnoty kříží. Vyslání impulsů je na obrázku 3.11.

Tyto impulsy jsou poté zesíleny a převedeny na gramy karbohydrátů (ve studii počítají se 4 g na impuls), viz obrázek 3.12.

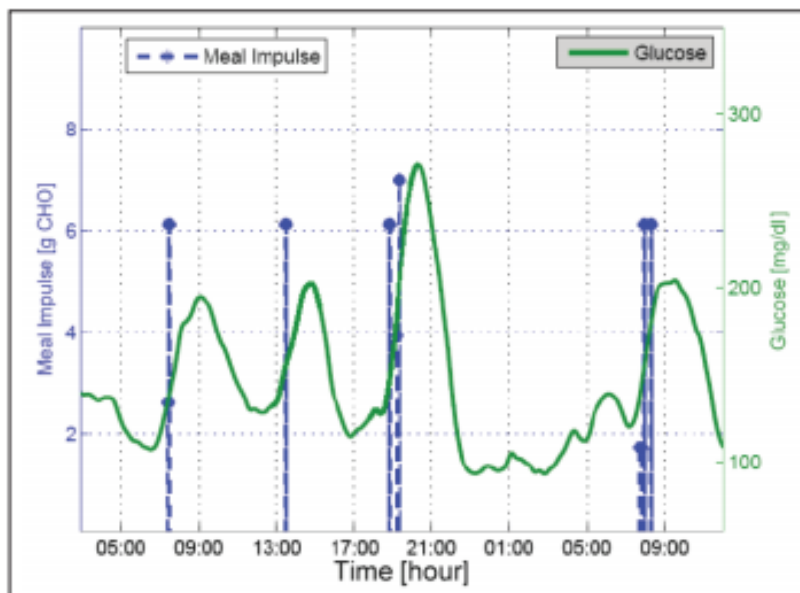
Počet impulsů, časové okno, thresholdy a množství karbohydrátů na impuls lze individuálně změnit podle diabetického profilu pacienta.

Obrázek 3.11: Impulsy



Zdroj: A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator [8]

Obrázek 3.12: Predikované karbohydráty



Zdroj: A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator [8]

V tabulce 3.2 jsou uvedeny výsledky provedeného testování. Algoritmus dosahuje úspěšnost 82 % a průměrný odhad jídla je 36,41 g karbohydrátů. Zpoždění detekce je průměrně 31 minut [8].

Tabulka 3.2: Výsledky

Table 4. Performance Analysis of New MSE on 100 Adolescents and 100 Adults	
Total number of meals	800
Total number of meal detections (true +)	656 (82.00%)
Total number of false positives	54 (6.75%)
Total number of false negatives	144 (18.00%)
Average detection time after meal onsets	31 [minutes]
Mean magnitude of false positively estimated meals	36.41 [g CHO]
Mean of min glucose of 0–90 min after false positive meals	119.84 [mg/dl]
Min of min glucose of 0–90 min after false positive meals	75.67 [mg/dl]
Max of min glucose of 0–90 min after false positive meals	180.73 [mg/dl]

Zdroj: A Closed-Loop Artificial Pancreas Using Model Predictive Control and a Sliding Meal Size Estimator [8]

3.3.3 Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data

V této studii Samadi et al. [23] 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 [9]. Zvolený parametr míry filtrování musí být dostatečně velký, aby odfiltroval šum, ale ne příliš, aby se neodfiltrovaly ostré nárůsty způsobené příjmem karbohydrátů.

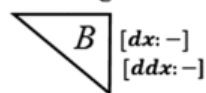
Pro extrakci vlastností je použita trojúhelníková kvalitativní reprezentace. V kvalitativních metodách extrakce dat jsou časová data konvertována na sekvenci kvalitativních proměnných. Pro trojúhelníkovou kvalitativní reprezentaci jsou kvalitativní proměnné trojúhelníky (viz obrázek 3.13). Trojúhelníky dělí data na epizody (obrázek 3.14) 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 $+/-$.

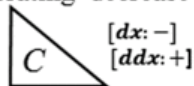
Obrázek 3.13: Kvalitativní proměnné trojúhelníkové kvalitativní reprezentace

Decreasing shapes

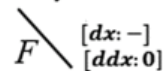
Accelerating decrease



Decelerating decrease

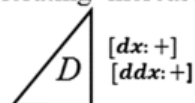


Steady decrease

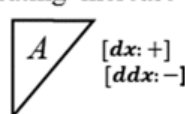


Increasing shapes

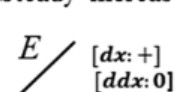
Accelerating increase



Decelerating increase



Steady increase

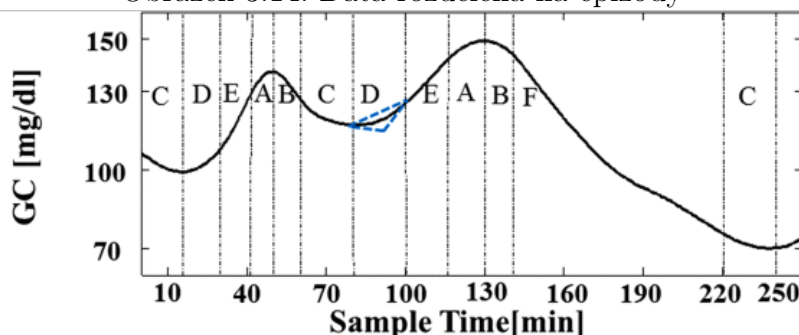


Constant shape

G [dx: 0]

Zdroj: Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data [23]

Obrázek 3.14: Data rozdělená na epizody



Zdroj: Meal Detection and Carbohydrate Estimation Using Continuous Glucose Sensor Data [23]

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”* Igt pro každý vzorek (váhy jsou nastaveny tak, že hodnota Igt je v intervalu $\langle -3; 3 \rangle$). Pakliže hodnota Igt (nebo součet Igt 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 [23].

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

Kölle et al. [11] 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).

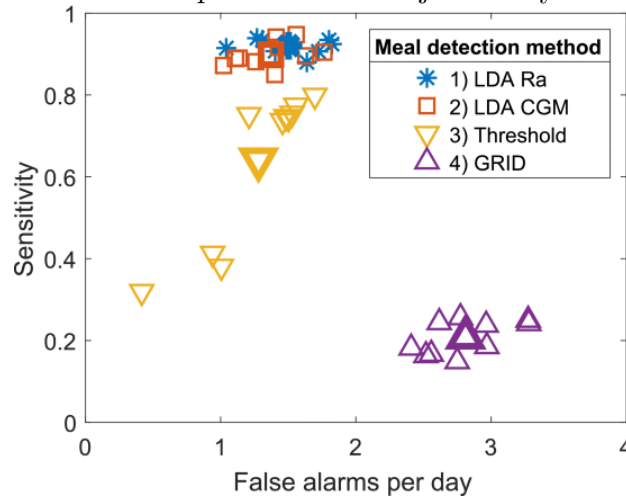
4. GRID algoritmus

Threshold pro naměřené hodnoty glukózy a jejich rychlost změny.

Na obrázku 3.15 je výkonnost jednotlivých metod při provedeném experimentu. Je vidět, že metody využívající LDA si vedly výrazně lépe. Jejich

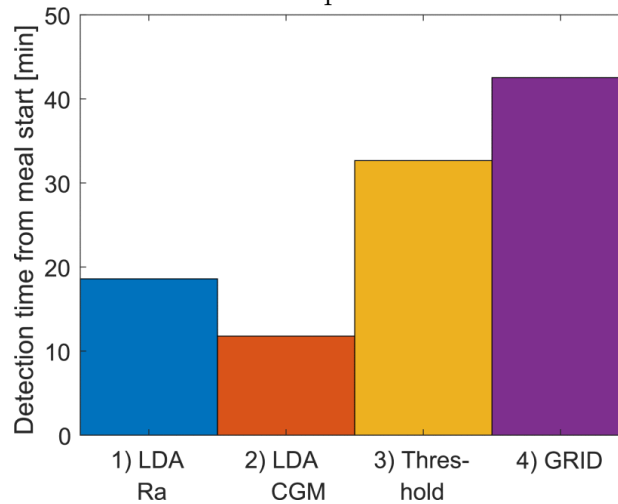
úspěšnost se pohybuje kolem 88 %. Také v případě času potřebného k detekci překonaly metody založené na thresholdu. Srovnání zpoždění u jednotlivých metod je na obrázku 3.16.

Obrázek 3.15: Úspěšnost detekce jednotlivých metod



Zdroj: Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1 [11]

Obrázek 3.16: Zpoždění detekce



Zdroj: Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1 [11]

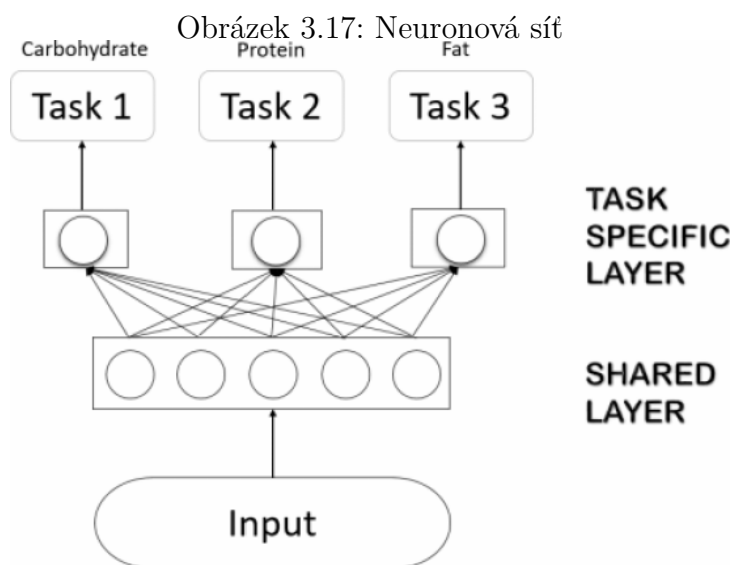
3.4 Metody využívající neuronové sítě

3.4.1 Predicting the meal macronutrient composition from continuous glucose monitors

V této studii Huo et al. [7] 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 3.17). 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.



Zdroj: Predicting the meal macronutrient composition from continuous glucose monitors [7]

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 zohled-

ňující metabolismus konkrétního jedince se střední kvadratickou chybou pro karbohydráty 0,39. Kvadratická chyba karbohydrátů pro první způsob je 0,45 [7].

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. [10] 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 parametrů 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 3.3 jsou výsledky jednotlivých metod.

Tabulka 3.3: 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 [10]

3.5 Porovnání jednotlivých metod

V tabulce 3.4 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 3.4: Porovnání jednotlivých metod

Studie	Zpoždění*	TP	FN	TN	FP
3.2.1 Turksoy et al. [28]	-	-	-	-	-
3.2.2 Ramkissoon et al. [21]	25-40	91.6%**	-	-	-
3.2.3 Fraser et al. [4]	-	-	-	11	0
3.2.4 Fathi et al. [3]	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. [32]	26-39	90%	10%	-	11.47%
3.3.2 Hyunjin et al. [8]	31 (avg.)	656 (82%)	144 (18%)	-	54 (6.75%)
3.3.3 Samadi et al. [23]	30 (avg.)	87%	13%	-	21%
3.3.4 Ra Kölle et al. [11]	18.59	92%	8%	-	1.5FP/den
3.3.4 CGM Kölle et al. [11]	11.78	90%	10%	-	1.37FP/den
3.4.1 Huo et al. [7]	8 hodin	0.39***			

* doba detekce karbohydrátů od příjetí 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 •

Autoři této metody využívají zařízení SenseWear® Pro Armband vyvinuté firmou BodyMedia (Pittsburgh, PA, USA). Toto zařízení je upevněno kolem ruky a sbírá data z pěti typů senzorů. Senzory snímají pozici a pohyb ruky a těla, teplotu kůže a okolí, vodivost kůže a srdeční aktivitu.

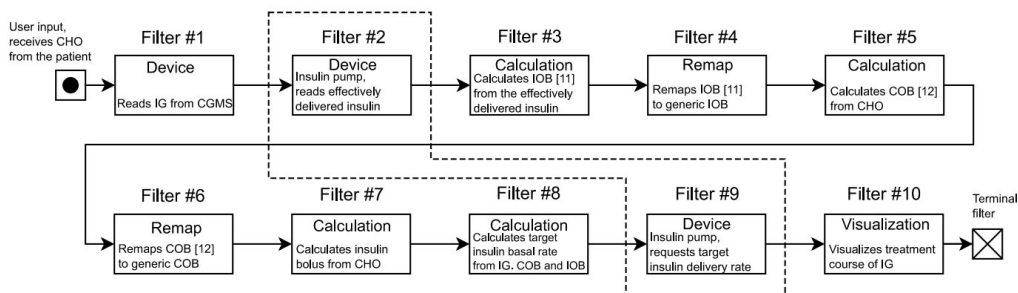
5 SmartCGMS

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 [12]. 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 [29]

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 [13]. 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 [12].

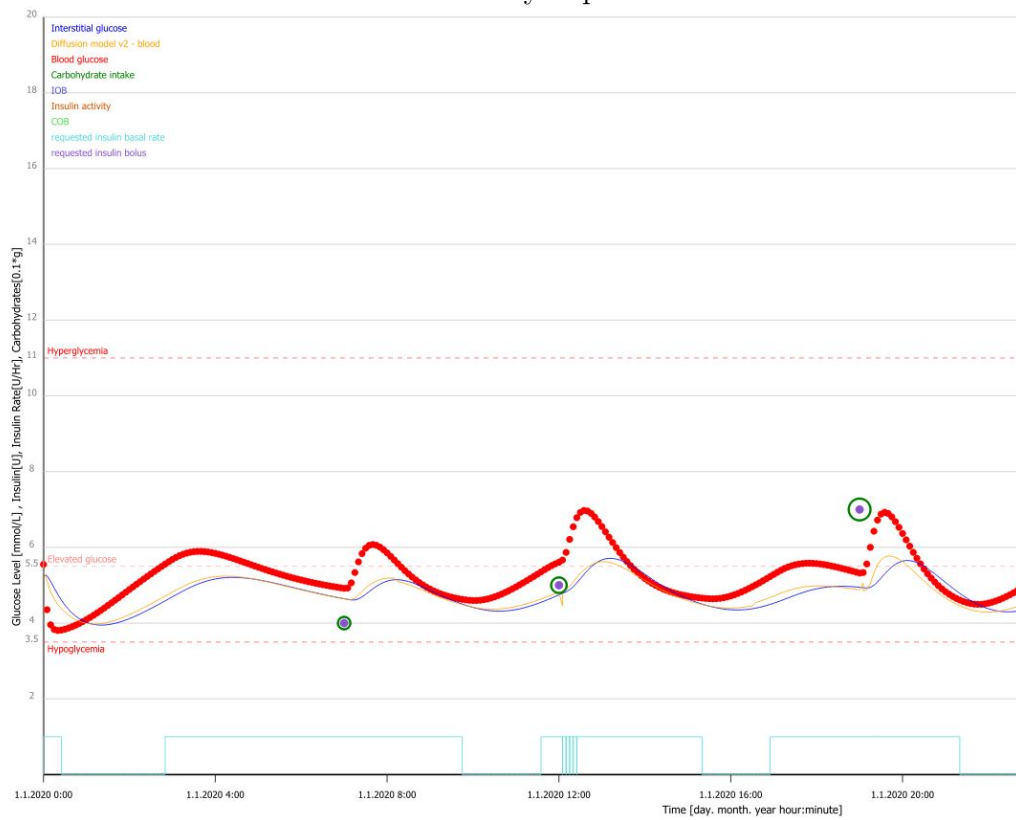
Pro určení parametrů modelu lze využít **solver**. Ten na základě metrik signálů modelu spočítá jeho optimální parametry. Příkladem solveru je Meta-differential evolution, Pathfinder, Sequential brute force scan, Particle swarm optimization, Spiral optimization a další.

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 Detekce příjmu karbohydrátů

Pro detekci příjmu karbohydrátů jsem měl k dispozici anonymizovaná data ze senzoru CGMS. Data obsahují naměřené a zadané hodnoty:

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

Pro detekci karbohydrátů jsem implementoval tyto 3 metody:

- Long short-term memory neuronová síť (kapitola 6.2)
- Lineární a kvadratická diskriminační analýza (kapitola 6.3)
- Detekce hran pomocí thresholdů 1. difference hodnot intersticiální glukózy (kapitola 6.4)

Programovací jazyk pro zpracování dat ze senzoru, návrh a vyhodnocení jednotlivých metod jsem zvolil Python. Python jsem zvolil pro jeho knihovny NumPy, pandas a SciPy pro zpracování a analýzu dat, matplotlib pro vykreslení dat a knihovny TensorFlow a Keras pro práci s neuronovými sítěmi. Přestože je Python interpretovaný jazyk a tudíž pomalý při zpracování velkého množství dat, tyto knihovny jsou implementovány v jazyce C

a optimalizovány pro vysoký výkon. Doba zpracování dat je tak srovnatelná s kompilovanými jazyky.

Implementace do SmartCGMS je v jazyce C++.

6.1 Příprava dat

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 6.1

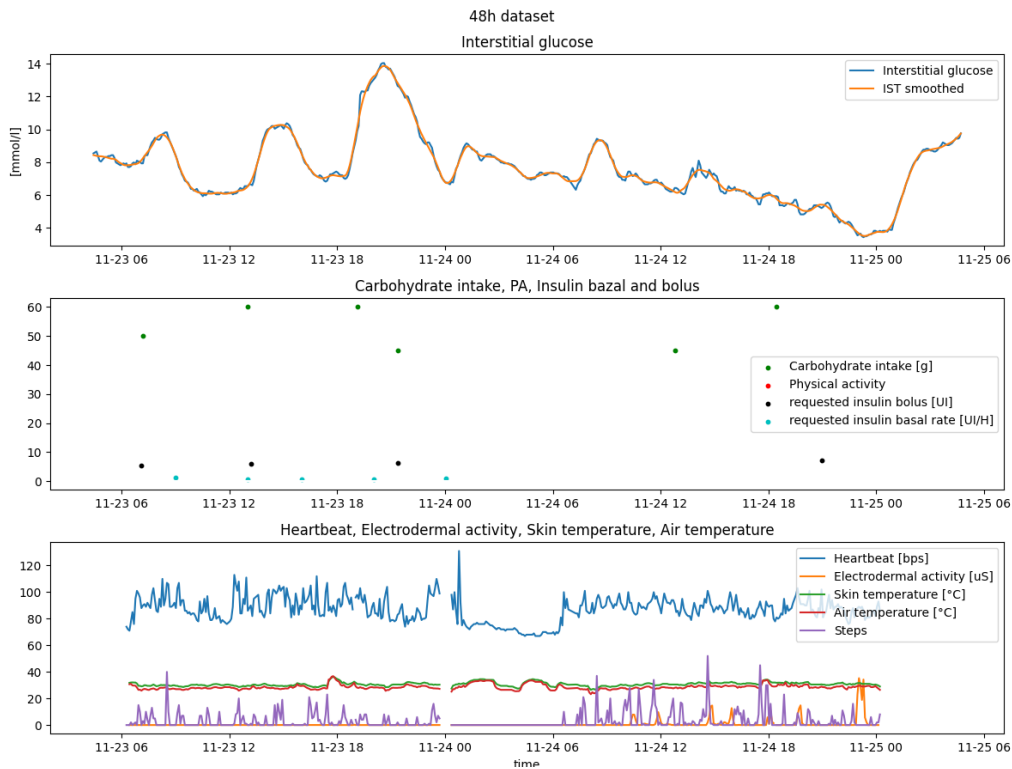
Tabulka 6.1: Signály ze CGMS

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...

Informaci pro následnou detekci v sobě nesou sloupce Device Time (čas měření), Signal (typ signálu) a Info (hodnota). Ty jsem extrahoval do dvou-rozměrné tabulky, kde řádky jsou čas měření a sloupce jednotlivé typy signálů. Data intersticiální glukózy jsem interpoloval Akima spline [25], z níž jsem získal chybějící hodnoty a derivace 1. 2. a 3. řádu. Pro vyhlazení hodnot intersticiální glukózy jsem použil Savitzky-Golay filtr [15] řádu 3 s velikostí okna 21. Jelikož různé typy signálu nejsou měřeny ve stejný okamžik, řádky jsem seskupil podle sloupce intersticiální glukózy, která je měřena v pětiminutových intervalech.

V grafech na obrázku 6.1 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 zbylé měřené hodnoty.

Obrázek 6.1: Data ze CGMS



Spuštění

Script pro transponaci dat se spustí zavoláním funkce:

```
load_log(patientID)
```

ze souboru `load_log.py`, kdy *patientID* je ID pacienta. Logy musí být umístěny v adresáři `data/`. Pro spuštění transponace dat z více logů se zavolá funkce:

```
load_log_all(patientIDs)
```

kdy *patientIDs* je pole ID logů, které chceme transponovat. Transponovaná data jsou uložena do .csv souboru *patientID-transposed.csv*.

Načtení a modifikace transponovaných dat se spustí voláním:

```
load_data(patientID, from_file, fill_missing, smooth,  
          derivation, norm, verbose, graphs, analyze)
```

ze souboru `load_data.py`, potažmo

```
load_data_all(patientIDs, from_file, fill_missing, smooth,  
              derivation, norm, verbose, graphs, analyze)
```

pro načtení několika souborů. *PatientID* je ID pacienta, *from_file* určuje zda se načtou již modifikovaná data nebo se spustí úprava transponovaných dat z logu. Parametr *fill_missing* určuje zda se mají nahradit chybějící hodnoty měření (" nenahrazuje se, 'akima' pro interpolaci Akima spline nebo 'mean' pro průměrnou hodnotu). Parametr *smooth* určuje způsob filtrace ('savgol' pro Savitzky-Golay filter). Parametr *derivation* určuje způsob výpočtu derivací ('akima' pro derivace Akima spline, 'gradient' pro numpy gradient, 'manual' pro výpočet difference). Parametr *norm* říká, zda chceme data normalizovat ('std' normalizace směrodatnou odchylkou na interval <-1;1>, 'min-max' standardizace rozpětím na interval <0;1>). Parametr *verbose* je pro vypisování průběhu zpracování dat, *graph* zobrazí modifikovaná data v grafech a *analyze* spustí analýzu dat z knihovny *sweetviz*. Modifikovaná data jsou uložena do .csv souboru *patientID-modified.csv*.

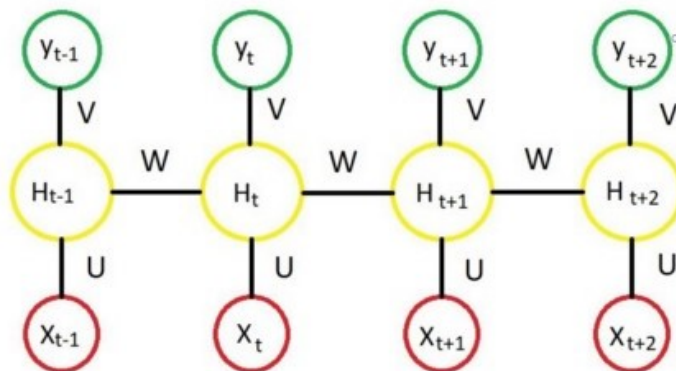
6.2 Rekurentní neuronové sítě

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. Riziko RNN spočívá v náchylnosti vzniku jevu označovaného jako mizející nebo explodující gradient vznikajícího 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.

Obrázek 6.2: Rekurentní neuronová síť



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

6.2.1 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 [31]:

$$\begin{aligned} f_t &= \sigma(X_t \otimes U_f + H_{t-1} \otimes W_f) \\ \bar{C}_t &= \tanh(X_t \otimes U_c + H_{t-1} \otimes W_c) \\ I_t &= \sigma(X_t \otimes U_i + H_{t-1} \otimes W_i) \\ O_t &= \sigma(X_t \otimes U_o + H_{t-1} \otimes W_o) \end{aligned}$$

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

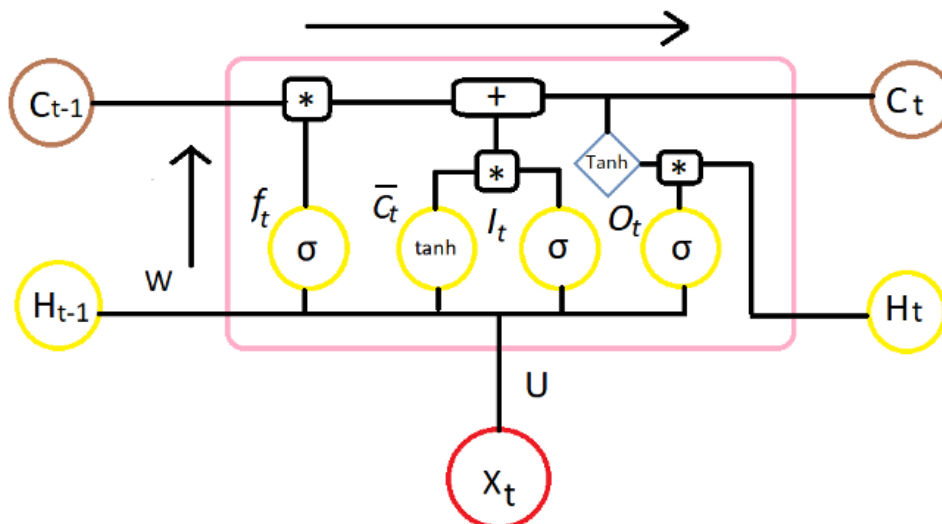
$$C_t = f_t \otimes C_{t-1} \oplus I_t \otimes \bar{C}_t$$

a výstupní stav H_t :

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

Diagram LSTM buňky je na obrázku 6.3.

Obrázek 6.3: LSTM buňka



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

6.2.2 Gated recurrent unit

Gated recurrent unit neuronová síť (GRU) využívá Update gate z_t a Reset gate r_t . Ty určují která 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 relevantní informace z předchozího stavu \bar{h}_t :

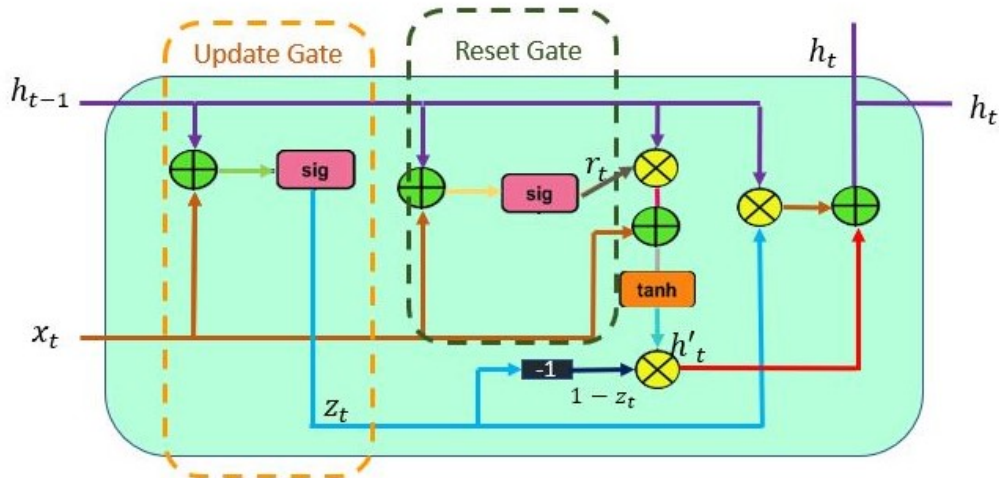
$$\bar{h}_t = \tanh(W_h x_t + U_h(r_t \otimes h_{t-1}))$$

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 6.4.

Obrázek 6.4: GRU buňka



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

6.2.3 Model

Pro detekci karbohydrátů jsem sestavil sekvenční keras¹ model se čtyřmi vrstvami:

- Obousměrná LSTM nebo GRU
- 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 a dropout 0,2. 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

¹Keras je nadstavba nad knihovnou TensorFlow pro neuronové sítě

vrstva je aplikována pouze při trénování sítě. První Dense vrstva (propojení neuronů formou každý s každým) má 128 neuronů, aktivací funkce je *relu*. Druhá Dense vrstva s jedním neuronem je výstupní.

Jako ztrátová funkce je použita střední kvadratická chyba (MeanSquareError), optimalizační algoritmus je Adam. Z datasetu je 80 % dat použito pro trénování a 20 % pro validaci. Data jsou do neuronové sítě dávkována v dávkách o velikosti 64. Jelikož se jedná o časovou řadu, data se nepromíchávají.

Spuštění

Trénování neuronové sítě se spustí voláním funkce:

```
lstm(df, headers, label, type, width, epochs)
```

ze souboru `cho_detection.py`, kde *df* je pandas dataframe modifikovaných dat, *headers* je pole názvů sloupců vstupních dat, *label* je predikovaná třída, *type* je LSTM nebo GRU, *width* je velikost okna a *epochs* je počet epoch, kolikrát má proběhnout trénovací cyklus.

Natrénovaný model se uloží do souboru `model/[patientID]_keras_model.h5` a následně se exportuje do souboru `model/[patientID]_fdeep_model.h5`, který používá C++ knihovna `frugally-deep` (viz kapitola 6.5).

Výsledky

Pro detekci karbohydrátů jsem jako vstupní prvky zvolil vyhlazená data intersticiální glukózy, jejich derivaci 1. řádu a minuty normalizované na interval $<0;1>$. Velikost okna je 24 (tj. 2 hodiny). Predikovaná třída je množství zadaných karbohydrátů. Tyto hodnoty jsou pro účel trénování rozkopírovány po dobu dvou hodin od času zadání. Neuronová síť je trénovaná ve 200 epochách.

Výstupem jsou predikované odhady karbohydrátů. Pro detekci příjmu karbohydrátů určíme `threshold`. Pokud množství karbohydrátů predikovaných neuronovou sítí je větší než tento `threshold`, je detekováno jídlo. `Threshold` je pro každého pacienta jiný.

//TODO

6.3 Lineární a kvadratická diskriminační analýza

Diskriminační analýzu jsem se rozhodl implementovat vzhledem k vysoké úspěšnosti a malému zpoždění, které dosáhli Kölle et al. [11] v článku **Pattern Recognition Reveals Characteristic Postprandial Glucose Changes: Non-Individualized Meal Detection in Diabetes Mellitus Type 1** (kapitola 3.3.4).

6.3.1 Teorie

Diskriminační analýza slouží k rozdělení prvků do konečného počtu tříd na základě lineární kombinace charakteristických prvků. Pro klasifikaci potřebujeme znát posteriory 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 [6]:

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

kde \mathbf{p} je počet prvků. Pro predikci je pak použito Bayesovo rozhodovací pravidlo [6]:

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

kdy hledáme třídu s maximálním posteriorem.

Logaritmus posterioru pro kvadratickou diskriminační analýzu (QDA) je [6]:

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

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 [24]:

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

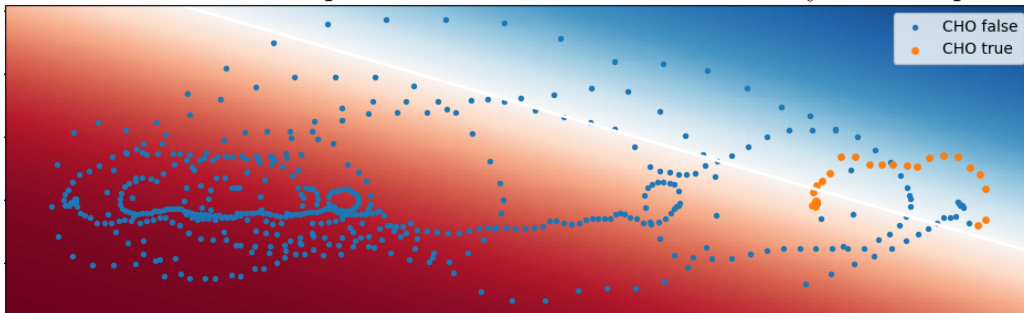
Vstupem diskriminační analýzy může být buď množina prvků pro daný časový okamžik nebo časové okno jednoho prvku. Predikované třídy jsou příjem/nepříjem karbohydrátů. Ty dostaneme převedením sloupce karbo-

hydrátů na booleanovskou proměnnou (0 je false, cokoli větší než 0 je true).

6.3.2 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 6.5).

Obrázek 6.5: Rozdělení prostoru lineární diskriminační analýzou dvou prvků



Spuštění

Trénování LDA a QDA pro množinu prvků se spustí voláním funkce:

```
lda(df, headers)
```

ze souboru `cho_detection.py`, kde *df* je pandas dataframe modifikovaných dat, *headers* je pole názvu sloupců vstupních dat.

Výsledky

```
//TODO
```

6.3.3 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.

Spuštění

Trénování LDA a QDA pro časové okno se spustí voláním funkce:

```
lda_window(df, header, width)
```

ze souboru `cho_detection.py`, kde `df` je pandas dataframe modifikovaných dat, `header` je název sloupce vstupních dat a `width` je velikost okna.

Výsledky

```
//TODO
```

6.4 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. Pakliže je funkce $f(x)$ v bodě x rostoucí, je její směrnice tečny v bodě x kladná. Pokud je $f(x)$ klesající, její směrnice tečny je v daném místě záporná. 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:

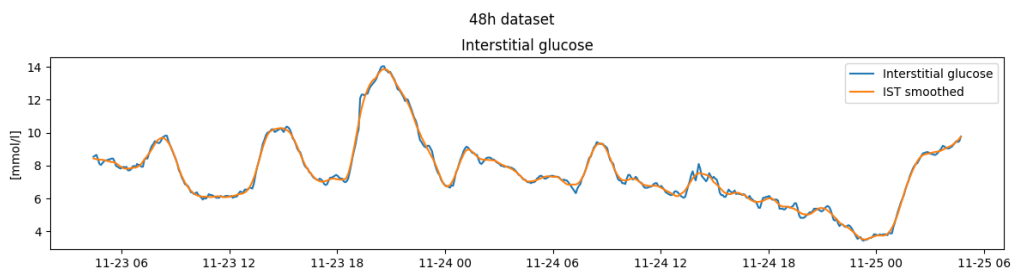
$$f'(x) = \frac{d}{dx} f(x) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

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 výchylkami 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 [15]. Pro vyhlazení dat intersticiální glukózy jsem zvolil polynom stupně 3 a velikost podmnožiny 21. Vyhlazená data vidíme v grafu na obrázku 6.6.

Obrázek 6.6: 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.

Spuštění

Detekce hran průběhu intersticiální glukózy se spustí voláním funkce:

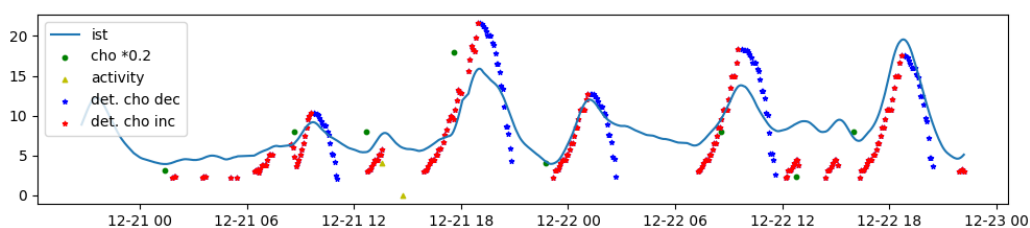
```
threshold(df, th, weight)
```

ze souboru `cho_detection.py`, kde df je pandas dataframe modifikovaných dat (`load_data` musí být spuštěno s parametry `smooth='savgol'`, `derivation='difference'`, `norm=''`), th je pole thresholdů a $weight$ je pole vah o stejné délce. Funkce vrací hodnoty aktivační funkce.

Výsledky

V grafu na obrázku 6.7 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 6.7: 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.

//TODO

6.5 Implementace do SmartCGMS

Do SmartCGMS jsem implementoval filtr pro detekci karbohydrátů pomocí detekce hran průběhu intersticiální glukózy a pomocí rekurentních neuronových sítí. Dále bylo potřeba implementovat Savitzky-Golay filtr pro vyhlazení dat a evaluační filtr. Vytvořené filtry jsou zkompileovány do knihovny `cho_detection.dll`.

Filtry 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.

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

Savitzky-Golay filtr počítá vyhlazená data intersticiální glukózy. Výstupem je **IST smoothed** signál.

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

Rekurentní neuronová síť využívá natrénovaného keras modelu. Ten je možné použít v C++ pomocí knihovny `frugally-deep` [5]. Model je nutné konvertovat pomocí skriptu `keras_export/convert_model.py`, který je součástí `frugally-deep` knihovny. Podporované sítě jsou LSTM i GRU. Příklad konfigurace s GRU je v konfiguračním souboru `setup_gru.ini` (viz příloha B).

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` (viz příloha B).

Evaluate filtr vyhodnocuje úspěšnost detekce. Pravdivě pozitivní (TP) jsou hodnoty, kdy je jídlo detekováno (hodnota detekce alespoň 1) do dvou hodin od jeho zadání pacientem. Pokud je hodnota detekce 2, je jídlo bráno jako potvrzeno. Pakliže jídlo není detekováno do dvou hodin od zadání, je výsledek falešně negativní (FN). V případě detekce hodnoty 2, kdy jídlo není zadáno, výsledek je falešně pozitivní (FP). Zároveň je počítáno zpoždění detekce od zadání jídla pacientem. Jelikož v datech nemusí být některá jídla zadána, statistika každého dne se započítá pouze tehdy, kdy byly zadány alespoň 3 jídla. Evaluate filtr neřeší časové segmenty.

Výsledky

Algoritmy byly testovány na datech jedenácti pacientů. Výsledky pro jednotlivé pacienty jsou v souboru *results.txt*. V tabulce 6.2 jsou souhrnné výsledky.

Tabulka 6.2: Výsledky

	Detekce hran	Detekce hran (+GRU)	GRU
Jídel	340		
Úspěšnost	85 %		
TP	289		
TP/pacient	26,27		
TP potvrzeno	167		
Potvrzeno	57,8 %		
FN	45		
FN/pacient	4,09		
FP	112		
FP/pacient	10,18		
Zpoždění*	27,54		

* průměrná doba detekce karbohydrátů od příjetí jídla v minutách

TP - true positive

FN - false negative

FP - false positive

Průměrná úspěšnost detekce jídla je 85 % s průměrným zpožděním 27,54 minut. Všechny algoritmy vykazují vysoký počet falešně pozitivních výsledků. To může být dáno častými výkyvy glukózy u pacienta nebo tím, že pacient jídlo nezadal, nebo ho zadal pozdě.

7 Detekce fyzické aktivity

8 Závěr

Přehled zkratek

CGM	kontinuální monitorace glukózy
CGMS	systém kontinuální monitorace glukózy
BG	glukóza v krvi
IST	intersticiální glukóza
CHO	karbohydráty
RNN	rekurentní neuronová síť
LSTM	Long short-term memory neuronová síť
GRU	Gated recurrent unit neuronová síť
ReLU	Rectified linear units funkce
LDA	lineární diskriminační analýza
QDA	kvadratická diskriminační analýza
CMHE	Committed moving horizon estimation
TP	pravdivě pozitivní
TN	pravdivě negativní
FN	falešně negativní
FP	falešně pozitivní
MAE	střední absolutní chyba
RMSE	střední kvadratické chyba

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] 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>.
- [4] 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>.
- [5] *Frugally Deep* [online]. 2021. [cit.]. Dostupné z: <https://github.com/Dobiasd/frugally-deep>.
- [6] HASTIE T., F. J. T. R. *The Elements of Statistical Learning*. Springer, 2008. ISBN 978-0-387-84858-7.
- [7] 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>.
- [8] 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>.
- [9] 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.
- [10] JOSE, J. A. – WAGGONER, T. – MANIKANDAN, S. Continuous Glucose Monitoring Prediction. 2021. Dostupné z: <https://arxiv.org/ftp/arxiv/papers/2101/2101.02557.pdf>.

- [11] 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>.
- [12] 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.
- [13] KOUTNY, T. – UBL, M. *SmartCGMS* [online]. Medicínská informatika, KIV, FAV, ZČU, 2015. [cit. 2021/05/05]. Dostupné z: <https://diabetes.zcu.cz/smartcgms>.
- [14] 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>.
- [15] 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>.
- [16] 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>.
- [17] PELIKÁNOVÁ TEREZIE A BARTOŠ VLADIMÍR. *Praktická diabetologie*. Maxdorf, 2010. ISBN 978-80-7345-216-2.
- [18] PERUŠIČOVÁ JINDRA. *Diabetes mellitus 1. typu*. Geum, 2007. ISBN 978-80-86256-49-8.
- [19] 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/>.

- [20] PSOTTOVÁ JANA. *Praktický průvodce cukrovkou III*. Maxdorf s. r. o, 2019. ISBN 978-80-7345-630-6.
- [21] 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>.
- [22] RYBKA JAROSLAV. *Diabetes mellitus - komplikace a přidružená onemocnění: diagnostické a léčebné postupy*. Grada, 2007. ISBN 9788024716718.
- [23] 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>.
- [24] *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.
- [25] *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>.
- [26] 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>.
- [27] 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.
- [28] 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>.

- [29] 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.
- [30] *Diabetes* [online]. World Health Organization, 2021. [cit. 2021/04/11]. Dostupné z: <https://www.who.int/health-topics/diabetes>.
- [31] 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.
- [32] 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: Adresářová struktura

Příloha B: Návod na spuštění