

Prepoznavanje in napovedovanje hiperglikemij in hipoglikemij na neinvaziven način

Božidara Cvetkovič
Institut "Jožef Stefan", Odsek za
inteligentne sisteme
Jamova cesta 39
1000 Ljubljana, Slovenija
+386 1 4773498
boza.cvetkovic@ijs.si

Urška Pangerc
Institut "Jožef Stefan", Odsek za
inteligentne sisteme
Jamova cesta 39
1000 Ljubljana, Slovenija

Mitja Luštrek
Institut "Jožef Stefan", Odsek za
inteligentne sisteme
Jamova cesta 39
1000 Ljubljana, Slovenija
+386 1 4773380
mitja.lustrek@ijs.si

POVZETEK

V tem prispevku je predstavljen pristop strojnega učenja za razpoznavanje in napovedovanje abnormalnih stanj glukoze (hiperglikemija, hipoglikemija) pri bolnikih z diabetesom tipa I in tipa II. Algoritme strojnega učenja smo uporabili na podatkih prsnega merilnika elektrokardiogramskega signala, ki poleg delovanja srca spremlja tudi dihanje. Signal smo obdelali z algoritmom, ki izlušči pomembne parametre, katere smo uporabili pri grajenju modelov za klasifikacijo. Dosegli smo 84% točnost pri napovedovanju glikemij v primeru bolnikov z diabetesom tipa I in 88,5% pri bolnikih z diabetesom tipa II. Pri prepoznavanju smo dosegli 78,05% točnost v primeru bolnikov z diabetesom tipa I in 75,81% pri bolnikih z diabetesom tipa II.

Ključne besede

EKG, glukoza, hiperglikemija, hipoglikemija, napovedovanje, razpoznavanje.

1. UVOD

Diabetes je kronična bolezen, ki se pojavi bodisi kadar trebušna slinavka ne proizvaja zadosti inzulina ali kadar telo ne more učinkovito izrabiti proizvedenega inzulina. Diabetes tipa 1 je avtoimunska bolezen, v kateri imunski sistem uniči celice v trebušni slinavki, ki proizvajajo inzulin, diabetes tipa 2 pa se največkrat razvije zaradi odpornosti celic na inzulin, k čemur precej prispeva nezdrav način življenja. Po pričanju svetovne zdravstvene organizacije je število obolelih za diabetesom v letu 2014 preseglo 9% odrasle svetovne populacije [1] in še narašča.

Bolniki z diabetesom so primorani celo življenje spremljati in vzdrževati nivo glukoze na ustreznih ravni, saj je lahko povečanje ali zmanjšanje glukoze vzrok za vrsto bolezni, med katerimi je najbolj izpostavljen tveganje za obolenje srca in ožilja. Svetovna zdravstvena organizacija trdi, da je vzrok smrti vsakega drugega bolnika z diabetesom povezan z odpovedjo srca ali možgansko kapjo.

Za zmanjšanje tveganja smrti je pomembno poleg ravni glukoze spremljati tudi delovanje srca. V medicinski literaturi lahko najdemo, da je hipoglikemija povezana z zmanjšanjem srčnega utripa in da je hiperglikemija močno povezana s polarizacijo in depolarizacijo srčnega prekata, ki ga opisuje tako imenovani interval QT, ki ga lahko pridobimo iz signala elektrokardiograma (EKG). Sprememba intervala QT je prav tako povezana z aritmijo, ki poveča tveganje za nenadno smrt.

Dostopnost komercialnih naprav za zvezno spremljanje signala EKG lahko bistveno olajšajo prepoznavanje in napovedovanje tako anomalij pri delovanju srca, kot tudi anomalij zaradi spremembe ravni glukoze.

V tem prispevku bomo predstavili pristop za razpoznavanje in napovedovanje hipoglikemij in hiperglikemij na neinvaziven način in sicer z uporabo komercialnega prsnega merilnika EKG [2], ki poleg delovanja srca meri tudi dihanje.

2. SORODNO DELO

Večina raziskav se ukvarja z iskanjem relacij med hipoglikemijo ter boleznimi srca in ožilja in s tem povezano umrljivostjo. Raziskave temeljijo na povezavi med parametri, izračunanimi iz signala EKG, ter izmerjeno ravno glukoze. Raziskave na tem področju uporabljajo signale iz kliničnega merilnika EKG in rezultati temeljijo na meritvah, opravljenih pod nadzorovanimi kliničnimi pogoji.

Hanefeld et. al. [1] sistematično predstavijo dosežke na tem področju. V raziskavah zveznega spremljanja glukoze pri bolnikih z diabetesom tipa I in II so ugotovili, da pride do spremembe dolžine intervala QT v primeru hude hipoglikemije (tudi nočne), kar sproži anomalijo srčnega ritma in lahko doprinese k višjemu tveganju za smrt. Potrditev teze podaljšanja intervala QT v primeru hipoglikemije lahko najdemo tudi v drugih neodvisnih raziskavah [4][5].

Povezava med hiperglikemijo in parametri srčnega signala so raziskovali Singh et al. [6]. Ugotovili so, da se variabilnost srčnega utripa zmanjša v primeru hude hiperglikemije.

Detekcijo hiperglikemije in hipoglikemije iz preostalih parametrov, izluščenih iz signala EKG (pri bolnikih z diabetesom tipa I), so raziskovali Nguyen et al. [7]. Ugotovili so, da je hitrost srčnega utripa (narašča) povezana izključno s hipoglikemijo in sprememba intervala PR, ki ga dobimo iz signala RKG, izključno s hiperglikemijo.

Raziskave, kjer bi uporabili metode strojnega učenja pri napovedovanju ali odkrivanju stanja glukoze, zaenkrat ne vsebujejo informacij, pridobljenih iz signala EKG, ampak napovedujejo nivo glukoze glede na kompleksne dinamične modele, priučene iz zgodovinskih podatkov o ravni glukoze pri določenem bolniku [8].

3. ZBIRKA PODATKOV IN PRISTOP

3.1 Zbirka podatkov

Zbiranje podatkov je potekalo šest tednov na dveh lokacijah, na Poljskem in v Italiji. Na Poljskem smo zbrali podatke 30 bolnikov z diabetesom tipa II in v Italiji 22 bolnikov z diabetesom tipa I. Vsi bolniki so bili opremljeni s prsnim merilnikom, ki zajema signal EKG, dihanja in pospeška, z merilnikom glukoze, z merilnikom krvnega tlaka, s tehtnico ter s pametnim telefonom, preko katerega so se podatki pošiljali v centralno bazo.

V trenutni raziskavi smo uporabili signal EKG, signal dihanja in podatke o izmerjenem nivoju glukoze. Signal EKG in signal dihanja smo najprej obdelali s filtri, ki so odstranili neberljive in šumne dele signalov in hkrati ohranili njihovo morfologijo. Očiščene dele signala smo razdelili na 30-minutne segmente, ki so sovpadali z meritvijo glukoze, in sicer:

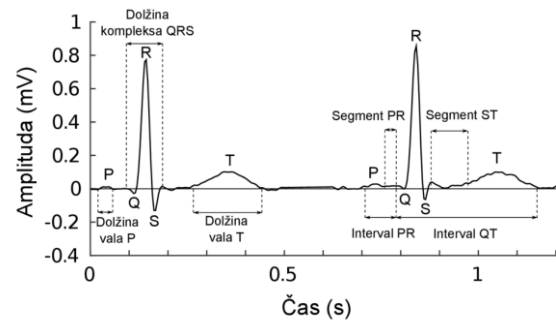
- od 45 do 15 minut pred meritvijo za potrebe napovedovanja nivoja glukoze (hiperglikemija, hipoglikemija, normalna glikemija)
- 15 minut pred in po meritvi za potrebe prepoznavanja nivoja glukoze (hiperglikemija, hipoglikemija, normalna glikemija)

Signal EKG smo obdelali z algoritmom, ki iz signala izloči 13 parametrov, ki opisujejo obliko srčnega utripa. Parametri signala so predstavljeni na sliki 1 in so sledeči:

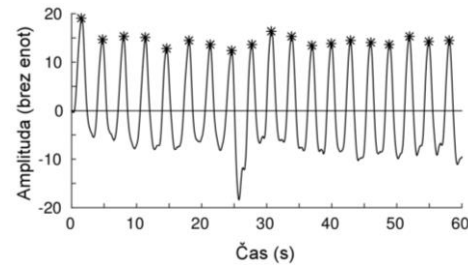
- Segment PR - čas med koncem vala P in začetkom kompleksa QRS
- Interval PR - čas med začetkom vala P in začetkom kompleksa QRS
- Interval QS - čas med začetkom in koncem kompleksa QRS
- Segment ST - čas med koncem kompleksa QRS in začetkom vala T
- Interval QT - čas med začetkom kompleksa QRS in koncem vala T
- Dolžina vala P - čas med začetkom in koncem vala P
- Dolžina vala T - čas med začetkom in koncem vala T
- Vrednost Q - amplituda vala Q
- Vrednost R - amplituda vala R
- Vrednost S - amplituda vala S
- Vrednost P - amplituda vala P
- Vrednost T - amplituda vala T
- Interval RR - čas med dvema zaporednima valoma R

Za vsak parameter smo v 30-minutnih segmentih izračunali srednjo vrednost, standardno deviacijo in trend (naklon linearne aproksimacije).

Signal dihanja smo prav tako razdelili na 30-minutne segmente, ki sovpadajo s časom meritve glukoze. Iz signala smo izluščili število vdihov na minuto ter izračunali standardno deviacijo in trend. Primer signala dihanja z označenimi razpoznanimi vdihmi je predstavljen na sliki 2.



Slika 1. Parametri, pridobljeni z obdelavo signala EKG.



Slika 2. Signal dihanja z označenimi razpoznanimi vdihmi.

Meritev glukoze vsebuje poleg same vrednosti nivoja glukoze tudi informacijo o tem, ali je bila izmerjena pred jedjo, po jedi, pred spanjem, ponoči ali ostalo. Glede na vrednost nivoja glukoze smo meritve razdelili v tri skupine:

1. Hipoglikemija – glukoza < 4 mmol/l
2. Hiperglikemija – glukoza > 7 mmol/l
3. Normalno stanje – 4 mmol/l < glukoza < 7 mmol/l

3.2 Pristop za napovedovanje in prepoznavanje glikemij

Za napovedovanje in prepoznavanje glikemij smo uporabili klasičen pristop strojnega učenja, kjer instanco sestavljajo izračunani in izbrani parametri enega ali več signalov. Tako sestavljanje instance obdelamo z algoritmom strojnega učenja za reševanje problema klasifikacije in točnost ovrednotimo z 10-kratnim prečnim preverjanjem.

Pristop smo zastavili na dva načina:

- 1) Začnemo z najbolj relevantnim signalom (EKG) in nato dodamo še dodaten vir informacij kot je dihanje ter čas v dnevu (pred, po jedi, pred spanjem, ponoči in ostalo).
- 2) Čas v dnevu uporabimo za razdelitev prostora napovedovanja na dva dela. Izdelamo dva modela: enega, ki napoveduje glikemije po jedi in enega pred jedjo.

Absolutne vrednosti izračunanih atributov (srednje vrednosti, standardne deviacije in trende trinajstih atributov iz signala EKG ter število vdihov, standardno deviacijo in trend dihanja) smo preračunali tudi v relativne vrednosti glede na posameznega bolnika. Sestavili smo sledeče štiri množice atributov:

- A1. Vsi atributi (absolutne in relativne vrednosti)
- A2. Absolutne vrednosti atributov
- A3. Relativne vrednosti atributov

A4. Najboljših 20 atributov glede na algoritem ReliefF [9]

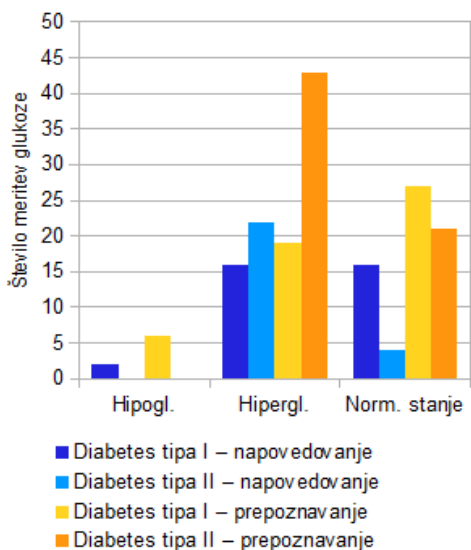
Množica atributov predstavlja instanco označeno s trenutnim stanjem glukoze (hiperglikemija, hipoglikemija, normalno stanje).

4. EKSPERIMENTI IN REZULTATI

Skupno smo izvedli 18 eksperimentov za napovedovanje glikemije in 18 eksperimentov za razpoznavanje glikemije. Za vsako množico atributov iz prejšnje sekcije smo naredili eksperiment z naslednjimi modeli:

- S1. Model, zgrajen z atributi signala EKG,
- S2. Model, zgrajen z atributi signala EKG in signala dihanja,
- S3. Model zgrajen z atributi signala EKG, signala dihanja ter atributom »čas meritve glukoze« (pred jedjo, po jedi, pred spanjem, ponoči ali ostalo)
- S4. Po atributu »čas meritve glukoze« smo razdelili množico na dva dela. Modele smo zgradili z atributi iz točke S2.

Slika 3 predstavlja porazdelitev razredov v zbirki podatkov, kjer lahko vidimo, da imamo največ primerkov hiperglikemije in le nekaj primerkov hipoglikemije v primeru bolnikov z diabetesom tipa I. Bolniki z diabetesom tipa II niso imeli primerov hipoglikemije.



Slika 3. Porazdelitev primerkov glikemij glede na tip diabetesa in glede na tip klasifikacijskega problema.

V eksperimentu pod točko 4 smo instance delili glede na parameter »čas meritve glukoze«, ki imajo vrednost pred jedjo ali po jedi. Porazdelitev po vrednosti je prikazana v tabeli 1 Zaradi premajhnega števila instanc napovedovanja stanja glukoze v krvi nismo testirali v naslednjih skupinah: bolniki z diabetesom tipa I po jedi ter bolniki z diabetesom tipa II pred jedjo in po jedi.

Vsako množico atributov smo testirali z desetimi klasifikacijskimi algoritmi strojnega učenja, ki so na voljo v paketu Weka [9]: naivni Bayes, logistična regresija, SVM, IB3, AdaBoostM1 z RepTree, Bagging z RepTree, JRip, J48, Random Forest in ZeroR kot izhodiščni algoritem, ki vedno vrne večinski razred. Vsak eksperiment je bil ocenjen in preverjen z uporabo 10-kratnega prečnega preverjanja.

Tabela 1. Število instanc glede na stanje glukoze v krvi in glede na čas meritve glukoze (pred jedjo ali po jedi).

Stanje	Napovedovanje				Prepoznavanje			
	Diab. I		Diab. II		Diab. I		Diab. II	
	Pr.	Po	Pr.	Po	Pr.	Po	Pr.	Po
Hipogl.	1	-	-	-	4	2	-	-
Hipergl.	10	-	-	-	8	7	9	32
Normalno	8	-	-	-	12	8	6	15

Rezultati eksperimentov so prikazani v tabeli 2 in tabeli 3, pri čemer tabela 2 vsebuje rezultate napovedovanja glikemij, tabela 3 pa rezultate prepoznavanja glikemij. Prva vrstica označuje tip eksperimenta (S1, S2, S3 ali S4), levi del tabele predstavlja rezultate pri bolnikih z diabetesom tipa I in desni del tabele rezultate pri bolnikih z diabetesom tipa II. Prvi stolpec obeh razdelkov vsebuje oznako množice uporabljenih atributov (Sekcija 3.2), ki je vrnila najboljši rezultat klasifikacije.

Pri napovedovanju glikemij se je izkazala množica A4 (izbranih 20 atributov glede na algoritem ReliefF) za najboljšo v primeru diabetesa tipa I in množica A2 (absolutne vrednosti vseh atributov) za primer diabetesa tipa II. Rezultat z najvišjo točnostjo 84,21% pri bolnikih z diabetesom tipa I dobimo z logistično regresijo in z delitvijo množice glede na čas meritve glukoze. V primeru bolnikov z diabetesom tipa II je najboljši rezultat pridobljen z algoritmom IB3 in sicer 88,46%. Domnevamo, da so se bolniki tipa II merili predvsem ob slabem počutju, saj meritve niso izvajali po jedi ali pred jedjo ampak različno skozi cel dan, zato rezultatov za eksperiment S4 z dodanim atributom čas meritve glukoze ni bilo mogoče izvesti.

Tabela 2. Rezultati napovedovanja hipoglikemij in hiperglikemij v primerjavi z izhodiščnim rezultatom.

	Diabetes tipa I			Diabetes tipa II		
	A	ZeroR (%)	Točnost (%)	A	ZeroR (%)	Točnost (%)
S1	A4	41.18	61.76	A2	84.62	88.46
S2	A4	41.18	64.71	A2	84.62	88.46
S3	A4	41.18	79.41	A2	84.62	88.46
S4	A4	52.63	84.21	-	-	-

Tabela 3. Rezultati prepoznavanja hipoglikemij in hiperglikemij v primerjavi z izhodiščnim rezultatom.

	Diabetes tipa I			Diabetes tipa II		
	A	ZeroR (%)	Točnost (%)	A	ZeroR (%)	Točnost (%)
S1	A3	51.92	63.46	A2	67.19	70.31
S2	A4	51.92	69.23	A2	67.19	71.88
S3	A4	51.92	74.15	A2	67.19	73.44
S4	A4	48.78	78.05	*	66.13	75.81

* A4 za model pred jedjo in A1 za model po jedi

Pri prepoznavanju glikemij (

Tabela 3), se je delitev po času meritve glukoze izkazal za najboljši pristop pri obeh tipih diabetesa. Za diabetes tipa I je

uporaba atributov iz množice A4 in algoritem SVM za model pred jedjo in logistična regresija za model po jedi vrnila najboljši rezultat in sicer 78,05%. Za diabetes tipa II je najboljši rezultat vrnila uporaba atributov iz množice A4 in algoritem SVM za model pred jedjo in atributov iz množice A2 in algoritem Bagging za model po jedi, in sicer 75,81%.

5. ZAKLJUČEK

V prispevku smo predstavili pristop za napovedovanje in razpoznavanje anomalij (hiperglikemija, hipoglikemija) pri bolnikih z diabetesom tipa I in tipa II. Uporabili smo splošni pristop strojnega učenja, kjer smo iz atributov, pridobljenih iz signala EKG in iz signala dihanja, zgradili modele za reševanje obeh klasifikacijskih problemov.

Eksperimente smo izvedli na podatkih 30 bolnikov z diabetesom tipa I in 22 bolnikov diabetesa tipa II. Ugotovili smo, da je najboljši pristop tako pri napovedovanju kot pri razpoznavanju gradnja dveh modelov, enega za napovedovanje ali razpoznavanje glikemij pred jedjo in drugega za napovedovanje ali razpoznavanje po jedi.

Z omenjenim pristopom smo dosegli 84,21% točnost pri napovedovanju glikemij pri bolnikih z diabetesom tipa I. Enakega postopka nismo mogli uporabiti pri bolnikih tipa II, saj je bilo premalo podatkov, označenih s časom merjenja glukoze. Z uporabo tako pomanjkljivih podatkov smo dosegli točnost 88,46%. Z enakim pristopom smo pri razpoznavanju dosegli točnosti 78,05% pri bolnikih s diabetesom tipa I in 75,81% pri bolnikih tipa II. Glede na to, da se razpoznavanje zdi lažja naloga od napovedovanja, je ta rezultat nekoliko presenetljiv in si ga bomo v prihodnje prizadevali pojasniti.

V nadaljnjem delu bomo tudi ocenili, ali je razvita metoda primerna za praktično prepoznavanje in napovedovanje glikemij, ki bi ga lahko uporabili, da bi bolniku svetovali, naj si izmeri raven glukoze v krvi z običajnim invazivnim merilnikom. Poleg tega bomo ugotovili signifikantnosti posameznih parametrov in se ukvarjali z nepravilnostmi delovanja srca.

6. ZAHVALA

Raziskava je bila delno financirana iz evropskega projekta COMMODITY12 (www.commodity12.eu).

7. REFERENCE

- [1] WHO, <http://www.who.int/mediacentre/factsheets/fs312/en/>
- [2] Zephyr BioHarness, <http://www.zephyranywhere.com/>
- [3] Hanefeld, M., Duetting, E., Bramlage, P. 2013. Cardiac implications of hypoglycaemia in patients with diabetes – a systematic review. *Cardiovascular Diabetology*, 135, 12 (Sep. 2013). DOI=10.1186/1475-2840-12-135.

- [4] Frier, M. B., Scherthaner, G., Simon R. Heller, R. S. 2011. Hypoglycemia and Cardiovascular Risks. *Diabetes Care*, 34 (May 2011). 132-137. DOI=10.2337/dc11-s220.
- [5] Snell-Bergeon, J. K., Wadwa, R. P. 2012. Hypoglycemia, Diabetes, and Cardiovascular Disease. *Diabetes Technology & Therapeutics*, 14, 51–58. DOI=<http://doi.org/10.1089/dia.2012.0031>
- [6] Singh, P.J., Larson, G. M., O'Donnell, J. C., Wilson F. P., Tsuji, H., Lloyd-Jones, M. D., Levy, D. 2000. Association of hyperglycemia with reduced heart rate variability (The Framingham Heart Study). *The American Journal of Cardiology*, 86, 3, (Aug 2000), 309-312, DOI=[http://dx.doi.org/10.1016/S0002-9149\(00\)00920-6](http://dx.doi.org/10.1016/S0002-9149(00)00920-6).
- [7] Nguyen, L. L., Su, S., Nguyen, H.T. 2012. Identification of Hypoglycemia and Hyperglycemia in Type 1 Diabetic patients using ECG parameters, in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Sand Diego, USA, August 28 – September 1, 2012), EMBC, IEEE, 2716-2719. DOI = 10.1109/EMBC.2012.6346525
- [8] Plis, K., Bunescu, R., Marling, C., Shubrook, J., Schwartz, F. 2014. A Machine Learning Approach to Predicting Blood Glucose Levels for Diabetes Management. *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*, 2014.
- [9] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, H. I. 2009. The WEKA Data Mining Software: An Update. *SIGKDD Explorations*, 11, 1, 2009.