

☒ Dve plati uporabe inteligentnih sistemov v medicini

Petra Povalej, Peter Kokol

Laboratorij za načrtovanje sistemov, Fakulteta za elektrotehniko, računalništvo in informatiko, CIMRŠ, Univerza v Mariboru
Smetanova ulica 17, Maribor
petra.povalej, kokol@uni-mb.si

Jernej Završnik

Zdravstveni dom Adolfa Drolca, Vošnjakova 2, Maribor
jernej.zavrsnik@zd-mb.si

Povzetek

Inteligentni sistemi se vse pogosteje uporabljajo na različnih področjih človekovega delovanja (npr. ekonomija, medicina, gospodarstvo ipd.), kjer se pogosto izkažejo kot zelo uporabna orodja. Uporabljamo jih za klasifikacijo, diagnosticiranje, predikcijo, iskanje povezav med različnimi dejavniki, odkrivanje novega znanja ipd. Na učinkovitost uporabe inteligentnih sistemov vpliva množica različnih dejavnikov (npr.: kvaliteta baze podatkov, relacije med vhodnimi in izhodnimi atributi, izbira metodologije inteligentnega sistema, vrsta in oblika podatkov ipd.), na katere imamo lahko v celotnem procesu inteligentne analize večji oz. manjši vpliv. V pričujočem prispevku smo se osredotočili na zanesljivost inteligentnih sistemov na realnih primerih s področja odločanja in diagnosticiranja, ki pokažeta prednosti in slabosti takšne analize. V procesu iskanja najkvalitetnejše rešitve smo preizkusili mnogo različnih pristopov gradnje klasifikacijskega modela (klasični, hibridni in multimetodni). Kljub vloženemu trudu se je iskanje kvalitetnega klasifikacijskega modela za neinvazivno določanje ravni holesterola v krvi izkazalo za neuspešno, saj zbrani parametri pri zbiranju podatkov (vhodni atributi) očitno niso dajali zadostne informacije o ravni holesterola v krvi. Nasprotno pa se je uporaba inteligentnih sistemov za pomoč pri diagnosticiranju prolapsa mitralne valvule (PMV) izkazala kot zelo učinkovita in zanesljiva.

Abstract

Two Sights of Using Intelligent Systems in Medicine

Intelligent systems have been often successfully applied on various fields such as business, economy, medicine, etc. They have been used for classification, diagnosing, prediction, knowledge discovery, etc. The efficiency of intelligent systems analysis very much depends on a type, format and quality of data gathered. However, the existence of appropriate relationships among attributes and outcomes and of course the selection of the right methodology is also crucial. In this paper we focused on the reliability of intelligent systems applied on a real-world problems where the benefits and drawbacks of such an analysis can be seen. With the aim of finding the most reliable solutions, different approaches for classification model construction were used. Nonetheless, the construction of reliable classification model for non-invasive determination of a blood cholesterol level proved to be unsuccessful mostly because of the lack of relationships among the attributes and the decisions. On the other hand, the use of intelligent systems analysis was very successful and reliable for diagnosing mitral valvule prolapse.

1 Uvod

Inteligentni sistemi so se izkazali kot zelo uporabno orodje na različnih področjih človekovega udejstvovanja, še posebno pa se širi njihova uporaba v medicini. Pogosto se uporabljajo za pomoč pri odločanju, klasifikaciji, predikciji, iskanju vzorcev, odkrivanju novih znanj, inteligentni analizi podatkov ter optimizaciji.

Zavedati se moramo, da na učinkovitost uporabe inteligentnih sistemov na realnih bazah podatkov močno vpliva vrsta različnih dejavnikov, kot na primer: kvaliteta baze podatkov, relacije med vhod-

nimi in izhodnimi atributi, oblika in vrsta podatkov ter izbira ustrezne metodologije gradnje inteligentnih sistemov. Kvalitetna priprava podatkov (izbira ustreznih parametrov) ter sam postopek zbiranja podatkov (natančnost pri merjenju, natančnost pri vnašanju podatkov ipd.) sta pogosto že odločilnega pomena za nadaljnji potek in uspešnost inteligentne analize. Prav ti dejavniki so ponavadi "krivi" za uspeh oz. neuspeh uporabe inteligentnih sistemov na realnih bazah podatkov.

V pričujočem prispevku se bomo osredotočili na oblikovanje uporabniku prijaznega klasifikacijskega modela, ki temelji na odločitvenih drevesih. Z izbiro predstavitve izluščenega znanja se bomo le delno omejili, saj bomo za gradnjo odločitvenih dreves uporabili več različnih pristopov (klasični, hibridni, multimetodni).

Predstavili bomo primer uporabe inteligentnih sistemov za pomoč pri določanju ravni holesterola v krvi, ki se je izkazal kot neuspešen predvsem zaradi slabe korelacije med vhodnimi atributi in odločitvenimi razredi ter relativno slabe kvalitete baze podatkov. V drugem primeru bomo prikazali zelo učinkovito uporabo različnih metodologij inteligentnih sistemov za oblikovanje klasifikacijskega modela za pomoč pri diagnosticiranju prolapsa mitralne valvule (PMV).

2 Inteligentni sistemi v medicini

Najpomembnejša lastnost inteligentnih sistemov je njihova zmožnost učenja ter sprejemanja odločitev na podlagi informacij iz preteklosti. Učenje vedno poteka na že zbranih podatkih, ki jih razdelimo na učno in testno množico. Inteligentni sistem mora poiskati pravila s procesom strojnega učenja na učni množici ter nato pravila preveriti s klasifikacijsko natančnostjo na neznanih podatkih iz testne množice. To je bistvenega pomena, saj poskušamo na takšen način doseči splošno uporabnost dobljenega modela na novih, še neznanih realnih primerkih.

Obstaja več metodologij inteligentnih sistemov, od katerih so najbolj znane nevronske mreže, odločitvena drevesa, genetski algoritmi, grobe množice ter hibridni sistemi, ki predstavljajo več (ponavadi dva) pristopov združenih v enega. S hibridnim pristopom poskušamo združiti prednosti, ki so značilne za posamezne pristope ter na takšen način doseči bolj kvalitetne rezultate. V literaturi [1] pogosto zasledimo hibride genetskih algoritmov in odločitvenih dreves [4], nevronskih mrež in odločitvenih dreves, grobih množic in nevronskih mrež [13] ipd.

Vsak posamezen pristop izgradnje ima določene prednosti in slabosti in prav zaradi tega moramo izbrati metodologije prilagoditi problemu ter zahtevam, ki jih moramo zadovoljiti pri reševanju le-tega. Nekatere metodologije, na primer odločitvena drevesa, nam dajejo dobro sliko glede odločitve, medtem ko druge metodologije, na primer nevronske mreže, ne dajejo pojasnil o sprejeti odločitvi.

Pogosto je zelo pomembno, da so odločitve, ki jih predlaga inteligentni sistem, razumljive ljudem, ki stojijo za procesom. Razumljivost inteligentnih sistemov je še posebno pomembna v medicini. Vsak zdravnik namreč odgovarja za svoje odločitve (npr. diagnozo), zato je s tega vidika zelo pomembno, da razume, kateri dejavniki so pripeljali do dane odločitve. Inteligentni sistem v tem primeru zdravniku predstavlja orodje za pomoč pri odločanju.

Razumljiva in enostavna predstavitev klasifikacijskega modela je pomembna tudi za nadzor delovanja modela. V primeru, da začne klasifikacijski model v nekem trenutku producirati napačne odločitve, je možno odkriti in analizirati vzroke za napačno delovanje le, če uporabljamo dobro razumljiv klasifikacijski model. V nasprotnem primeru, če klasifikacijski model deluje kot »črna skrinjica« (npr. nevronske mreže), je zelo težko odkriti, kaj je privedlo do napačnega obnašanja.

V medicini se inteligentni sistemi pogosto uporabljajo tudi za potrjevanje znanih dejstev in odkrivanje novega znanja (oz. novih – še ne znanih povezav med danimi atributi in diagnozo). Glavni problem, ki se pojavlja v tem primeru, je ocena kvalitete novega znanja. Novo znanje se lahko izkaže le kot posledica prekomerne prilagoditve klasifikacijskega modela učnim objektom in tako predstavlja nesmiselne povezave. S tega vidika je zelo pomembno sodelovanje s strokovnjakom z danega področja (npr. kardiolog, pediater ipd.), ki oceni kvaliteto novo odkritega znanja.

2.1 Izbira metodologije in učenje

Izbira metodologije inteligentnih sistemov je v veliki meri odvisna od zastavljenega problema. Razumljivost zgrajenega klasifikacijskega modela je v našem primeru ključnega pomena pri izbiri metodologije. Glavna prednost odločitvenih dreves [7] pred ostalimi metodologijami je njihova enostavnost in razumljivost, kar pojasnjuje pogosto uporabo odločitvenih dreves tudi v medicini. Njihova prednost je tudi, da jih je prav tako možno uporabiti brez računalnika, saj preprosto sledimo hierarhični strukturi znanja.

Zavedati se moramo, da vsestransko uporabna metodologija inteligentnih sistemov, ki daje v vseh primerih najkvalitetnejše rezultate, ne obstaja. Zato je zelo pomembno, da se pri reševanju problema ne omejimo na eno samo metodologijo, temveč z eksperimentiranjem najdemo najboljše. V nadaljevanju

bomo na kratko predstavili osnovne razlike med metodologijami, ki smo jih uporabili v raziskavi.

a) Klasična metodologija gradnje odločitvenih dreves

V procesu klasične gradnje odločitvenih dreves je izbira najprimernejšega atributa v posameznem vozlišču odločitvenega drevesa odvisna od informacijske vsebnosti atributa, ki se oceni na osnovi hevristične funkcije. Najpogosteje uporabljene hevristične funkcije za oceno informacijske vsebnosti atributa, ki smo jih uporabili tudi v naši raziskavi, so naslednje:

- stopnja informacijskega prirastka (*Information gain ratio*) [8], ki izhaja iz entropije (*ID3*),
- χ^2 (*Chi-square*) [1],
- Gini [10],
- J-measure [9].

Poleg zgoraj naštetih osnovnih hevrističnih funkcij smo generirali tudi različne linearne kombinacije [6] le-teh ter njihovo učinkovitost preizkusili na obeh bazah podatkov. Kvaliteto izgrajenih odločitvenih dreves smo poskušali izboljšati z uporabo boostinga – metode za izboljšavo natančnosti poljubnega algoritma za učenje [2].

b) Gradnja odločitvenih dreves z genetskimi algoritmi – hibridni pristop

Gradnja odločitvenih dreves z genetskimi algoritmi je hibridni pristop, ki združuje prednosti genetskih algoritmov (npr.: neobčutljivost na šum, večja verjetnost za pridobitev optimalne rešitve) in prednosti odločitvenih dreves (enostavna in razumljiva predstavitev rezultatov). Gradnja odločitvenega drevesa poteka podobno kot na klasičen način, z razliko, da se v procesu iskanja optimalnega odločitvenega drevesa uporabi evolucijski pristop.

c) Multimetodni pristop

V primerjavi s klasičnimi hibridnimi pristopi multimetodni pristop [3] omogoča uporabo različnih metodologij v enem samem modelu znanja, kjer vsaka izmed metod odraža lastne prednosti in tudi omejitve. Posledično torej pričakujemo, da takšna kombinacija različnih metod v večini primerov pripelje do boljših rezultatov.

2.2 Ocena kvalitete zgrajenega modela

Univerzalen kriterij za oceno kakovosti zgrajenega klasifikacijskega modela na žalost ne obstaja. Pogosto se uporabljajo indikatorji, kot so skupna natančnost,

senzitivnost, specifičnost. V predstavljeni raziskavi smo se odločili za uporabo naslednjih indikatorjev:

- skupna natančnost = $\frac{\text{število pravilno klasificiranih objektov}}{\text{število vseh objektov}}$;
(*average total accuracy*)
- povprečna natančnost odločitvenih razredov
(*average class accuracy*).

Povprečna skupna natančnost odločitvenega drevesa nam pogosto ne daje realne informacije o dejanski kvaliteti odločitvenega drevesa, saj se v večji meri srečujemo s problemi, pri katerih je klasifikacija objektov določenega odločitvenega razreda zelo težka. Tako lahko naučeno odločitveno drevo klasificira testne objekte z visoko skupno natančnostjo, vendar pa je natančnost klasifikacije določenega odločitvenega razreda zelo nizka. Takšno odločitveno drevo je v praksi neuporabno, saj se je specializiralo za klasifikacijo samo določenih odločitev (ne pa vseh). Skupna natančnost odločitvenega drevesa torej ni zadosten indikator kvalitete naučenega drevesa. Mnogo več nam o kvaliteti odločitvenega drevesa pove povprečna natančnost odločitvenih razredov.

Pri reševanju medicinskih problemov za evalvacijo kvalitete odločitvenih dreves pogosto uporabljamo izraza senzitivnost in specifičnost. Senzitivnost označujemo kot odstotek pravilno klasificiranih zdravih, specifičnost pa kot odstotek pravilno klasificiranih bolnih subjektov:

- senzitivnost = $\frac{tp}{tp + fn}$; (*sensitivity*)
- specifičnost = $\frac{tn}{tn + fp}$; (*specificity*)

Oznake, uporabljene v zgornjih enačbah, imajo naslednji pomen:

- tp = True Positive = število pravilno klasificiranih pozitivnih objektov
- fp = False Positive = število napačno klasificiranih pozitivnih objektov
- fn = False Negative = število napačno klasificiranih negativnih objektov
- tn = True Negative = število pravilno klasificiranih negativnih objektov

3 Dve plati uporabe inteligentnih sistemov v medicini

Uspešnost aplikacije inteligentnih sistemov za reševanje realnih problemov je odvisna od mnogih dejavnikov. V prvi fazi lahko trdimo, da je učinkovitost

inteligentnega sistema v veliki meri odvisna od kvalitete in vrste zbranih podatkov. Pri uporabi realnih baz podatkov se pogosto soočimo z naslednjimi težavami:

- šum v podatkih – pojavlja se zaradi nenatančnosti meritev, napak pri vnašanju podatkov, napačnih podatkov ter manjkajočih podatkov,
- majhne baze podatkov – za slovenske razmere je že baza s podatki o nekaj tisoč pacientih redka, vendar pa včasih tudi nekaj tisoč primerkov v bazi podatkov ni dovolj za izgradnjo učinkovitega modela,
- neenakomerna zastopanost posameznih odločitvenih razredov – posledično lahko pride do prekomernega prilagajanja klasifikacijskega modela bolj zastopani odločitvi,
- varnost podatkov – podatki o bolnikih morajo biti varovani, zato pogosto medicinske ustanove niso pripravljene teh podatkov posredovati.

Poleg kvalitete dane baze podatkov je vsekakor zelo pomembna tudi ustrezna relacija med vhodnimi atributi in odločitvami. V primeru, da te relacije ni ali pa je zelo šibka, seveda ne moremo pričakovati, da bo inteligentni sistem to relacijo našel.

Seveda na učinkovitost inteligentnega sistema ne vpliva samo baza podatkov, temveč je le-ta v veliki meri odvisna tudi od izbire ustrezne metodologije. Že osnovna praksa nam pokaže, da so določene metodologije bolj oz. manj uspešne na različnih bazah podatkov. Prav zaradi tega je potrebno v procesu iskanja najbolj kvalitetne rešitve danega problema preizkusiti čim več različnih metodologij.

V nadaljevanju bomo predstavili primera aplikacije inteligentnih sistemov na dveh realnih medicinskih bazah podatkov.

PRIMER I:

Gradnja modela za pomoč pri določanju ravni holesterola v krvi

V današnjem času se vedno pogosteje srečujemo s problemom previsoke vrednosti maščob v krvi, predvsem pri odraslih. Zadnje študije kažejo, da tudi otroci zaužijejo preveliko količino maščob v vsakodnevni prehrani, kar jim daje slabo predispozicijo za prihodnost. S tega vidika je še posebej pomembno, da že v zgodnjih letih odkrijemo povečane vrednosti maščob v krvi otrok ter tako zmanjšamo verjetnost pojava s tem povezanih težav v odrasli dobi. Ker se otroci bojijo invazivnih raziskav (npr. klasičnega odvzema krvi), smo želeli v raziskavi najti enostavno

izmerljive kazalce nepravilnih vrednosti holesterolov. S pomočjo inteligentnih orodij smo poskušali poiskati korelacije med danimi atributi in vrednostjo maščob v krvi ter možnost neinvazivnega napovedovanja previsokih/prenizkih vrednosti. V ta namen smo skupaj z zdravstvenim domom Adolfa Drolca izvedli raziskavo, ki jo podajamo v pričujočem prispevku.

Opis baze podatkov

Podatki, ki so bili uporabljeni za gradnjo klasifikacijskega modela za določanje ravni holesterola v krvi, so podatki o 712 petletnih otrocih, ki so bili zbrani v Zdravstvenem domu Adolfa Drolca v Mariboru. 539 otrok je imelo pri pregledu normalno raven holesterola v krvi, 173 otrok pa nenormalno raven. Atributi, ki so služili kot osnova za izgradnjo klasifikacijskega modela, so dejansko podatki, zbrani z enostavnimi meritvami (npr: obseg glave, obseg nadlahtnice, kožna guba, krvni tlak ipd.), ki jih izvedejo pri standardnem sistematskem pregledu otrok.

Uspešnost gradnje klasifikacijskega modela je precej odvisna tudi od razdelitve osnovne baze podatkov na učno in testno množico. Zaradi primerljivosti rezultatov, pridobljenih z različnimi metodologijami, smo v vseh primerih uporabljali isto učno in testno množico. Pri generiranju učne množice smo bili pozorni na čim bolj enakomerno porazdelitev učnih objektov glede na odločitvena razreda (normalen/nenormalen holesterol).

Analiza rezultatov

Podatki, ki so bili uporabljeni za gradnjo klasifikacijskega modela za določanje ravni holesterola v krvi, so zgodovinski podatki o 712 petletnih otrocih, od katerih je 539 otrok imelo normalno raven holesterola v krvi in 173 nenormalno raven. Baza podatkov je torej razmeroma majhna, poleg tega pa je tudi porazdelitev objektov glede na odločitveni atribut zelo neenakomerna.

Rezultati gradnje odločitvenih dreves z vsemi zgoraj opisanimi metodologijami so prikazani v tabeli 1. Metode so razvrščene glede na povprečno natančnost odločitvenih razredov zgrajenega klasifikacijskega modela, ki se giblje v mejah od 48.57% (Boost Greedy J measure + Chi square) do 64.84% (Boost Greedy Gini * Chi square). Povprečna skupna natančnost pa se giblje med 41.91% (Greedy J measure * Gini, Greedy J measure * ID3) in 94.31% (Multimethod). Opazimo lahko precej veliko razliko v kvaliteti posameznih metodologij v povprečni

natančnosti odločitvenih razredov in še mnogo večjo razliko pri upoštevanju povprečne skupne natančnosti zgrajenih klasifikacijskih modelov.

Opazili smo, da se pogosto pojavlja problem klasifikacije objektov z nenormalno ravniho holesterola (nizka senzitivnost). Takšna drevesa so v praksi neuporabna, saj je natančnost klasificiranja z vsemi metodami precej nizka. Zaradi zelo slabega osnovnega klasifikatorja tudi boosting v večini primerov ni bil uspešen pri izboljšavi natančnosti letega. Kljub uporabi različnih metodologij torej nismo uspeli generirati klasifikacijskega modela, ki bi bil dovolj zanesljiv za uporabo v praksi. Vzroke za neuspešnost lahko delno iščemo v sami bazi podatkov (premajhno število vnosov, premalo otrok z nenormalno ravniho holesterola, šum v podatkih), poleg tega pa je povezava med vhodnimi atributi in ravniho holesterola verjetno prešibka, da bi lahko dokaj zanesljivo skleпали o odločitvi.

Tabela 1:
Določanje ravni holesterola v krvi; skupna natančnost (\square) in povprečna natančnost odločitvenih razredov (Δ) na testni množici ter rang metode glede na povprečno natančnost odločitvenih razredov

Method	\square (%)	Δ (%)	Rank
Greedy ID3	53.29	57.92	17
Boost Greedy ID3	54.79	51.30	34
Greedy Gini	61.97	57.58	19
Boost Greedy Gini	51.79	52.18	33
Greedy Chi square	57.48	55.20	24
Boost Greedy Chi square	64.97	54.22	27
Greedy J measure	48.20	55.22	23
Boost Greedy J measure	54.49	61.03	9
Greedy Gini + ID3	55.98	56.88	21
Boost Greedy Gini + ID3	60.18	64.05	2
Greedy Gini * ID3	53.59	60.55	10
Boost Greedy Gini * ID3	60.18	54.16	28
Greedy Chi square + ID3	61.97	60.05	13
Boost Greedy Chi square + ID3	64.37	53.90	31
Greedy Chi square * ID3	60.18	54.16	28
Boost Greedy Chi square * ID3	61.67	59.89	14
Greedy Gini + Chi square	60.47	56.79	22
Boost Greedy Gini + Chi square	62.57	57.90	18
Greedy Gini * Chi square	60.18	54.16	28
Boost Greedy Gini * Chi square	61.67	64.84	1
Greedy J measure + ID3	51.19	54.34	26

Boost Greedy J measure + ID3	66.16	62.28	4
Greedy J measure * ID3	41.91	61.78	5
Boost Greedy J measure * ID3	60.18	64.05	2
Greedy J measure + Chi square	59.28	61.10	7
Boost Greedy J measure + Chi square	58.98	48.57	38
Greedy J measure * Chi square	51.49	54.50	25
Boost Greedy J measure * Chi square	59.28	61.10	8
Greedy J measure + Gini	52.39	50.03	35
Boost Greedy J measure + Gini	61.37	59.74	15
Greedy J measure * Gini	41.91	61.78	5
Boost Greedy J measure * Gini	59.28	53.68	32
Greedy linear	47.60	57.38	20
Greedy Boost linear	58.08	60.46	11
Greedy Voting	52.09	49.87	37
Genetic	79.69	59.40	16
Multimethod	94.31	50.00	36
Nested Boost	62.27	60.21	12

PRIMER II:

Gradnja modela za pomoč pri diagnosticiranju prolapsa

Prolaps mitralne valvule (PMV) je v svoji definiciji opisan kot odmik srčne zaklopke iz normalne pozicije. Posledice tega odmika so motnje v pretoku krvi (vrtnici), neobičajne bolečine v prsih, možnost nastajanja krvnih strdkov in celo smrt. Simptome PMV je mogoče najti pri petih do desetih odstotkih prebivalstva. Pri milejših oblikah PMV zdravljenje ni potrebno, medtem ko pri težjih oblikah PMV pride v poštev tudi kirurški poseg. Ker je pri postavljanju diagnoze poleg naprave za pregled z ultrazvokom potrebno še precej znanja in izkušenj, so se pojavile težnje po postavitvi novih kriterijev za odkrivanje PMV. Ideja je najti bolezn, simptome in sindrome, povezane s PMV, ki bi jih lahko razpoznal splošni zdravnik in tako odkril potencialne paciente s PMV. Cilj podane raziskave je bil odkriti sindrome PMV-ja pri osebah med 1. in 18. letom.

Opis baze podatkov

Z uporabo Monte Carlo metode vzorčenja je bilo izbranih 900 otrok in mladostnikov starih od 1 do 18 let. Na poziv k rutinskemu pregledu se jih je odzvalo 631, katerih zdravstveno stanje so ocenili na osnovi posebej pripravljenega protokola za sindrom PMV-ja. Protokol je obravnaval 103 parametre, za katere je verjetno, da označujejo prisotnost PMV-ja. Porazdelitev

glede na odločitvene razrede je bila naslednja: 5% "prolaps", 6% "tihi prolaps" in 89% "ni prolapsa". Baza podatkov ni vsebovala manjkajočih vrednosti.

Analiza rezultatov

Učno množico je sestavljalo 501 oseb, v testni množici pa je bilo 130 subjektov. Vsak subjekt je bil opisan s 102 parametroma, ki so opisovali pacientove bolezni, bolezni njegovih staršev, parametre preiskav z ultrazvokom itd. Baze nismo razdelili sami v razmerju 2/3 za učno množico in 1/3 za testno množico, temveč smo zaradi primerljivosti uporabili že obstoječo učno in testno množico. Relativne frekvence posameznih odločitev so bile približno enake v učni in testni množici.

Rezultati klasifikacije so predstavljeni v tabeli 2. Metode v tabeli so razvrščene glede na povprečno natančnost odločitvenih razredov inducirane klasifikacijskega modela. Opazimo lahko, da se povprečna natančnost odločitvenih razredov giblje od 40.55% (Boost Greedy J measure * Gini, Boost Greedy J measure + Gini) do 84.31% (Multimethod). Povprečna skupna natančnost pa se giblje v mejah od 80.77% (Greedy J measure * Gini) do 93.84% (Multimethod). Oba indikatorja kvalitete klasifikacijskega modela torej kažeta na to, da je multimetodni pristop na tej bazi prekosil vse ostale preizkušene metodologije.

Tabela 2: PMV; skupna natančnost (\square) in povprečna natančnost odločitvenih razredov (Δ) na testni množici ter rang metode glede na povprečno natančnost odločitvenih razredov

Method	\square (%)	Δ (%)	Rank
Greedy ID3	88.46	73.39	12
Boost Greedy ID3	91.54	70.94	13
Greedy Gini	88.46	77.87	9
Boost Greedy Gini	87.69	64.15	20
Greedy Chi square	90.00	82.91	5
Boost Greedy Chi square	90.77	57.21	24
Greedy J measure	83.08	50.84	32
Boost Greedy J measure	89.23	53.08	29
Greedy Gini + ID3	92.31	79.27	6
Boost Greedy Gini + ID3	90.77	69.75	16
Greedy Gini * ID3	90.00	78.43	7
Boost Greedy Gini * ID3	92.31	57.77	23
Greedy Chi square + ID3	91.54	83.47	2
Boost Greedy Chi square + ID3	89.23	56.65	25

Greedy Chi square * ID3	90.77	70.66	14
Boost Greedy Chi square * ID3	93.08	70.59	15
Greedy Gini + Chi square	91.54	83.47	2
Boost Greedy Gini + Chi square	89.23	56.65	25
Greedy Gini * Chi square	87.69	69.54	17
Boost Greedy Gini * Chi square	90.00	61.42	22
Greedy J measure + ID3	86.15	51.96	30
Boost Greedy J measure + ID3	89.23	48.60	35
Greedy J measure * ID3	81.54	50.28	33
Boost Greedy J measure * ID3	90.77	53.64	28
Greedy J measure + Chi square	88.46	74.30	11
Boost Greedy J measure + Chi square	91.54	66.46	18
Greedy J measure * Chi square	83.85	55.60	27
Boost Greedy J measure * Chi square	87.69	48.95	34
Greedy J measure + Gini	85.39	51.68	31
Boost Greedy J measure + Gini	89.23	40.55	37
Greedy J measure * Gini	80.77	41.95	36
Boost Greedy J measure * Gini	89.23	40.55	37
Greedy linear	91.54	83.47	2
Greedy Boost linear	91.54	74.51	10
Greedy Voting	90.00	78.43	7
Genetic	92.31	65.83	19
Multimethod	93.84	84.31	1
Nested Boost	93.08	62.54	21

4 SKLEP

Učinkovitost uporabe inteligentnih sistemov pri reševanju realnih problemov je v veliki meri odvisna od same kvalitete baze podatkov ter izbire ustrezne metodologije inteligentnega sistema.

V pričujočem prispevku smo predstavili primer učinkovite in primer neučinkovite aplikacije inteligentnih sistemov na realnih medicinskih bazah podatkov. Uporabili smo več metodologij inteligentnih sistemov, ki temeljijo na različnih pristopih gradnje odločitvenih dreves (klasični, hibridni in multimetodni pristop). Pokazali smo, da je pri iskanju najkvalitetnejše rešitve potrebno preizkusiti več različnih metodologij, saj ena sama univerzalna metodologija ne obstaja. Na danih realnih problemih smo prikazali uspešnost oz. neuspešnost inteligentnih sistemov.

Uporaba inteligentnih sistemov za določanje ravni holesterola v krvi pri petletnih otrocih na osnovi neinvazivnih indikatorjev se je izkazala za neuspešno v

vseh primerih, saj na osnovi dane baze podatkov kljub uporabi različnih metodologij ni bilo mogoče zgraditi kvalitetnega klasifikacijskega modela. Vzroke lahko iščemo predvsem v nezadostni korelaciji med vhodnimi atributi in odločitvami.

V nasprotju z neuspešnostjo prvega primera pa je bila kvaliteta klasifikacijskega modela za pomoč pri diagnosticiranju prolapsa (PMV) zelo visoka. Med vsemi preizkušenimi metodologijami se je kot najuspešnejši izkazal multimetodni pristop, s katerim smo uspeli generirati klasifikacijski model s 84,31% povprečno natančnostjo odločitvenih razredov na testni množici. Visoka natančnost klasifikacije omogoča aplikacijo klasifikacijskega modela tudi v praksi, kot eno izmed metod za pomoč pri diagnosticiranju PMV.

Razlike v kvaliteti, na osnovi različnih metodologij zgrajenih klasifikacijskih modelov, so bile v obeh primerih zelo visoke, kar ponovno kaže v prid sicer časovno potratne uporabe različnih metodologij v procesu iskanja kvalitetnih rešitev.

VIRI IN LITERATURA

- [1] Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J., Classification and Regression Trees, Wadsworth International Group, Belmont, CA, 1984.
- [2] Freund, Y., Schapire, R. E., Experiments with a new boosting algorithm, V: Proceedings Thirteenth International Conference on Machine Learning, Morgan Kaufman, San Francisco, 1996, 148–156.
- [3] Lenič, M., Kokol, P., Combining Classifiers with Multimethod Approach, V: Soft computing systems: design, management and applications, IOS Press, 2002, 374–383.
- [4] Podgorelec, V., Kokol, P., Induction of medical decision trees with genetic algorithms. V: Proceedings of the international ICSC congress on Computational intelligence methods and applications, 1999.
- [5] Podgorelec, V., Kokol, P., Yamamoto, R., Masuda, G., Sakamoto, N., Knowledge discovery with genetically induced decision trees, V: International ICSC congress on Computational intelligence: methods and applications (CIMA'2001), 2001.
- [6] Povalej P., Lenič M., Kokol P., Reliability of decision trees, V: Journal of Medical Systems, in press.
- [7] Quinlan, J. R., C4.5: Programs for machine learning. Morgan Kaufmann publishers, San Mateo, CA, 1993.
- [8] Quinlan, J. R., Induction of decision trees, Machine Learning vol. 1, Kluwer Academic Publishers, 1986.
- [9] Smyth, P., Goodman, R. M. Rule induction using information theory, V: Piatetsky-Schapiro, G., Frawley, W. J. (Eds.), Knowledge Discovery in Databases, AAAI Press, 1991, 159–176.
- [10] White, A. P., Liu W. Z., Bias in information-based measures in decision tree induction, Machine Learning, 1994, 15:321–329.
- [11] Wolpert, D., Macready, W., No free lunch theorems for optimization. IEEE Transactions on Evolutionary Computation, 1997, 1(1):67–82.
- [12] Zorman, M., Hibridni pristopi strojnega učenja. Doktorska disertacija, Maribor, 2001.
- [13] Zorman, M., Kokol, P., Redukcija iskalnega hiperprostora rešitev s hibridom grobih množic in nevronske mreže. V: Proceedings of the tenth ERK Conference Portorož, Slovenia, 2001, 205–208.

Petra Povalej je diplomirala na smeri računalništvo z matematiko na Pedagoški fakulteti Univerze v Mariboru. Je mlada raziskovalka in študentka tretjega letnika doktorskega študija iz računalništva in informatike na tej ustanovi. Njeno glavno raziskovalno področje so inteligentni sistemi. Je dobitnica prve nagrade na tekmovanju študentskih člankov na konferenci Computer Based Medical Systems 2002 in dobitnica nagrade za najboljši poster na konferenci 4th Asia-Pacific Conference on Simulated Evolution and Learning.

Dr. Peter Kokol je diplomiral s področja elektrotehnike in doktoriral s področja računalništva na Univerzi v Mariboru. Njegova raziskovalna področja so inteligentni sistemi, teorija sistemov, teorija kaosa in kvaliteta programske opreme. Je vodja nacionalnih in mednarodnih projektov, med njimi tudi projektov 5QP. Njegova bibliografija obsega več kot 500 enot, od tega več kot 60 izvirnih znanstvenih člankov. Bil je predsednik organizacijskih in programskih odborov več svetovnih konferenc.

Mag. Jernej Završnik, dr. med., je diplomiral in magistriral na Medicinski fakulteti v Zagrebu in specializiral iz pediatrije. Trenutno je predstojnik Oddelka za varstvo žena, otrok in mladine v Zdravstvenem domu Maribor. Njegova raziskovana področja so echocardiografija, pediatrija in medicinska informatika.