

Hierarhični odločitveni modeli in njihova uporaba v zdravstvu

Marko Bohanec⁽¹⁾, Blaž Zupan⁽¹⁾, Vladislav Rajkovič^(2,1)

(1) Institut Jožef Stefan, Jamova 39, 1000 Ljubljana

(2) Univerza v Mariboru, Fakulteta za organizacijske vede, Kranj

Povzetek

Hierarhične odločitvene modele razvijemo z razgradnjo zahtevnih odločitvenih problemov na manjše in manj zahtevne podprobleme. Z njimi lahko klasificiramo oziroma vrednotimo variante ter jih uporabimo za različne analize, simulacije in razlage. Ta prispevek predstavlja osnovne metode za gradnjo in uporabo hierarhičnih modelov ter podaja pregled njihove praktične uporabe v zdravstvu pri nas. Prikazani so dosednji rezultati projektov, ki tečejo na področjih onkologije, radiologije, patronažne nege in nevrofiziologije. Med metodami za gradnjo modelov je posebej izpostavljena nova metoda strojnega učenja HINT, ki temelji na funkcijski dekompoziciji in dopolnjuje klasični ("ročni") način razgradnje problema z možnostjo avtomatizirane sinteze modela iz podatkov. Na primerih so prikazane tudi nekatere druge metode vrednotenja in analize variant, ki so pomembne za uporabo hierarhičnih modelov v zdravstvu.

1. Uvod

Hierarhični večparametrski odločitveni modeli so pripomoček za klasifikacijo oziroma vrednotenje objektov, opisanih v atributnem jeziku [Bohanec, Rajkovič 1995]. Temeljijo na razgradnji problema na manjše in manj zahtevne podprobleme. Podprobleme oziroma vmesne koncepte ponazorimo s spremenljivkami, ki jih med seboj povežemo v strukturo; ta je v splošnem usmerjen aciklični graf, najpogosteje pa drevo. Za vsako notranje vozlišče definiramo predpis, ki opredeljuje odvisnost te spremenljivke od njenih neposrednih naslednikov v strukturi.

Posebej razširjena je uporaba hierarhičnih modelov pri podpori odločitvenih procesov. Gre za probleme, pri katerih je potrebno izmed več variant izbrati tisto, ki najbolj ustreza danim ciljem. Odločitveni model zgradimo in uporabimo predvsem zato, da z njim ovrednotimo variante in na tej osnovi izberemo najboljšo. Poleg tega lahko uporabimo model za najrazličnejše analize in simulacije. Pomembna je tudi vloga modela, ki jo ima kot strukturiran in človeku razumljiv prikaz zakonitosti oziroma ekspertnega znanja o problemskem področju.

Metode za delo s hierarhičnimi odločitvenimi modeli so bile razvite v okviru raziskav psihologije, operacijskih raziskav, odločitvene analize in sorodnih področij [Chankong, Haimes 1983; Vincke 1992; Stewart 1992; Nagel 1993]. Praktično so se uveljavile predvsem pri podpori odločanja na tehničnem in poslovnem področju, na primer pri ocenjevanju investicij in projektov, strateškem planiranju in kadrovskega odločanju. Na teh področjih imamo razmeroma veliko izkušenj tudi pri nas. Razvili smo lupino ekspertnega sistema za podporo večparametrskega odločanja DEX [Bohanec, Rajkovič 1990] in jo uporabili pri reševanju več deset realnih odločitvenih problemov [Urbančič *et al.* 1991].

V zadnjem času se uporaba hierarhičnih modelov postopoma širi tudi na področje medicine. Za to je več razlogov. Nove metode za delo s kvalitativnimi (simboličnimi) modeli so omogočile njihovo uporabo tudi na problemskih področjih, ki jih označujemo kot "mehka" in kamor sodi tudi medicina. Ti problemi so slabše strukturirani, manj formalizirani, njihovo reševanje pa v veliki meri temelji na presoji kvalificiranih strokovnjakov. Pri teh problemih sta posebej pomembna: (1) človeku razumljiva predstavitev znanja o problemskem področju in (2) zmožnost metode pojasnjevati potek in rezultate reševanja konkretnega problema. Bistveno novost na tem področju predstavlja tudi metoda za strojno učenje iz podatkov HINT [Zupan 1997], ki omogoča avtomatizirano gradnjo hierarhičnih modelov na osnovi primerov minulih odločitev, na primer postavljenih diagnoz.

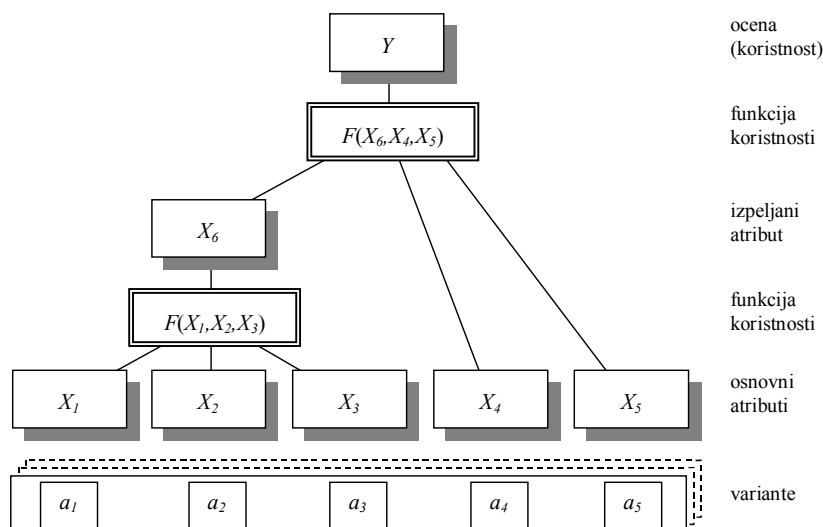
Namen tega preglednega prispevka je predstaviti osnovne metode za delo s hierarhičnimi modeli ter podati pregled njihove praktične uporabe v zdravstvu pri nas. Prikazani so dosednji rezultati projektov, ki tečejo na področjih onkologije, radiologije, patronažne nege in nevrofiziologije. Najbolj podrobno je obdelan projekt s področja onkologije, pri katerem gre za povezavo klasičnih "ročnih" tehnik gradnje modelov z metodami avtomatizirane sinteze znanja iz podatkov.

V naslednjem razdelku so opisani osnovni gradniki hierarhičnih odločitvenih modelov ter metode, s pomočjo katerih te gradnike definiramo, jih uredimo v ustrezno strukturo in uporabimo za vrednotenje in analizo variant. Tretji razdelek prikazuje praktično uporabo teh metod v okviru štirih projektov, pri katerih so modeli razviti ročno. Četrty razdelek je osredotočen na avtomatizirano sintezo modelov iz podatkov z metodo HINT, ki jo s stališča klasifikacijske točnosti primerjamo z nekaterimi drugimi metodami strojnega učenja ter predstavimo dosedanje rezultate na področjih nevrofiziologije in onkologije.

2. Hierarhični odločitveni modeli

2.1 Gradniki in struktura modelov

Hierarhični odločitveni model je v splošnem sestavljen iz dveh tipov komponent (slika 1): atributov X_i in funkcij koristnosti F_i . *Atributi* (imenovani tudi parametri ali kriteriji) so spremenljivke, ki ponazarjajo podprobleme odločitvenega problema. Urejeni so v hierarhično strukturo, najpogosteje drevo, ki ponazarja medsebojne odvisnosti med atributi: višjenivojski atributi so odvisni od neposrednih naslednikov. Na sliki 1 je na primer atribut X_6 odvisen od X_1 , X_2 in X_3 . Glede na položaj v strukturi ločimo attribute na *osnovne* (končna vozlišča – listi) in *izpeljane* (notranja vozlišča). Na sliki 1 so osnovni atributi X_1 do X_5 , izpeljana pa sta X_6 in Y . V modelu je za vsak izpeljani atribut določena *funkcija koristnosti* F , ki opredeljuje odvisnost tega atributa od njegovih neposrednih naslednikov v strukturi.



Slika 1: Hierarhični odločitveni model

Variante opišemo z vrednostmi a_i osnovnih atributov. *Vrednotenje* variant poteka “od spodaj navzgor”, to je s postopnim združevanjem vrednosti v skladu s strukturo modela in funkcijami koristnosti. Končno *oceno* (koristnost) variante ponazarja vrednost spremenljivke Y .

Večina klasičnih večparametrskih odločitvenih metod uporablja zvezne attribute in funkcije, določene na osnovi utežene vsote. Takšne modele imenujemo *kvantitativne*. V tem prispevku obravnavamo *kvalitativne* modele, ki jih uporablja sistem DEX. Atributi v teh modelih so diskretni, funkcije koristnosti pa so predstavljene z *odločitvenimi pravili* tipa *če-potem*. Primeri takšnih atributov in funkcij so prikazani v razdelku 3.1.

2.2 Gradnja modelov

2.2.1 Ročna gradnja modelov

Klasični in še vedno prevladujoči pristop h gradnji hierarhičnih odločitvenih modelov lahko označimo z besedo "ročni". Gre za proces, ki navadno poteka v dialogu med ekspertom in odločitvenim analitikom [Mallach 1994] in v okviru katerega se na bolj ali manj sistematičen način postopno oblikujejo in med seboj povezujejo posamezni gradniki modela. Delo tipično poteka po naslednjih fazah:

1. *Identifikacija in strukturiranje atributov:*
 - a. *Seznam atributov:* Oblikujemo nestrukturiran seznam atributov, na osnovi katerih bomo ocenjevali variante. Pomembno je, da pri tem ne spregledamo atributov, ki bistveno vplivajo na oceno (načelo polnosti). Poskusimo izpolniti tudi nekatere druge zahteve, kot so neredundantnost, ortogonalnost in operativnost (merljivost) atributov.
 - b. *Strukturiranje atributov:* Attribute hierarhično uredimo upoštevajoč medsebojne odvisnosti in vsebinske povezave. Nepomembne attribute in tiste, ki so izraženi z ostalimi, zavržemo ter po potrebi oblikujemo nove. Rezultat je hierarhija (oziroma drevo) atributov.
 - c. *Merske lestvice:* Vsem atributom določimo merske lestvice, to je zalogo vrednosti ter morebitne druge lastnosti, na primer urejenost.
2. *Definicija funkcij koristnosti:* V tej fazi postopno za vse izpeljane attribute določimo funkcije koristnosti. Postopek navadno poteka po načelu združevanja: od listov proti korenu definirane hierarhije atributov.

Pri uporabi oznake "ročna gradnja modelov" je potrebno poudariti, da se ta nanaša predvsem na fazi 1a in 1b, ki sta najzahtevnejši in praktično v celoti temeljita na znanju in izkušnjah razvijalcev. Po drugi strani pa sta fazi 1c in 2 v današnjih podpornih računalniških programih razmeroma aktivno podprti z vrsto metod, ki bistveno olajšajo in pohitrijo razvoj. Zajemanje funkcij koristnosti je na primer podprto s podsistemi za zagotavljanje konsistentnosti in polnosti funkcij [Saaty 1993; Bohanec, Rajkovič 1990], možnostjo izbiranja oziroma parametrizacije vnaprej pripravljenih funkcij [Brans, Mareschal 1990], zajemanja v grafični obliki ter z različnimi dialogi, ki jih vodi računalniški program [Humphreys, McFadden 1980].

2.2.2 Strojno učenje modelov: metoda HINT

Na področju medicinske diagnostike in prognostike so se uspešno uveljavile metode strojnega učenja [Kononenko, Kukar 1995], ki iz primerov minulih odločitev (diagnoz, prognoz) bolj ali manj samostojno zgradijo pravilo za klasifikacijo novih primerov in predstavitev zakonitosti, ki jih opisujejo primeri. Pri nas so se posebej uveljavili sistemi za učenje odločitvenih dreves, kot sta Asistent [Cestnik *et al.* 1987] in C4.5 [Quinlan 1993].

HINT (Hierarchy INduction Tool) je nova metoda strojnega učenja, ki na osnovi dane množice učnih primerov poišče definicijo ciljnega koncepta, predstavljeno v obliki hierarhičnega odločitvenega modela [Zupan 1997; Zupan *et al.* 1997a; Bohanec *et al.* 1997a]. To pomeni, da metoda iz primerov, predstavljenih v atributnem jeziku, samostojno ali v dialogu z uporabnikom zgradi hierarhijo atributov, pri čemer definira nove izpeljane attribute in funkcije koristnosti. Funkcije koristnosti predstavi v obliki odločitvenih tabel, tako kot DEX. Metoda HINT na ta način v primerih, ko imamo na voljo učne primere, aktivno podpre tudi fazi 1a in 1b, predstavljeni v razdelku 2.2.1.

Metoda HINT temelji na funkcijski dekompoziciji. Množico učnih primerov obravnava kot delni opis funkcije $y=F(X)$, kjer je X množica osnovnih atributov in y spremenljivka razreda. Funkcijo poskuša razgraditi v funkciji $y=G(A,c)$ in $c=H(B)$, kjer $A \cup B = X$. Tako odkrije nov izpeljani atribut c ter funkciji G in H , delno opisani z množicami primerov, ki sta manjši od prvotne. Dokler je razgradnja mogoča, HINT rekurzivno nadaljuje ta postopek tudi na funkcijah G in H , tako da je končni rezultat razgradnje hierarhija atributov in funkcij.

Izmed več možnih razgradenj v vsakem koraku HINT izbere tisto, ki minimizira bodisi kompleksnost, bodisi klasifikacijsko napako modela. Za uporabo v medicini je posebej primerna druga inačica, ki omogoča obravnavo nekonsistentnih oziroma šumnih učnih primerov.

2.3 Uporaba modelov

Osnovni namen uporabe hierarhičnih odločitvenih modelov je *vrednotenje variant*: vsako varianto, opisano z vrednostmi osnovnih atributov, ovrednotimo z modelom in dobimo njeno oceno. Na osnovi ocen potem variante med seboj primerjamo, jih rangiramo in izberemo najboljšo.

Ta osnovna funkcija še zdaleč ne zadošča za uspešno praktično uporabo modelov. Na končno oceno variant vpliva mnogo dejavnikov in pri vsakem lahko pride do napake. Poleg tega sama končna ocena navadno ne zadostuje za celovito sliko o posamezni varianti. Variante želimo podrobneje analizirati ter razložiti proces in rezultate odločanja. Postavlja se vrsta vprašanj, kot so:

- Kako je bila izračunana končna ocena – na osnovi katerih vrednosti atributov in katerih funkcij?
- Zakaj je končna ocena takšna, kot je? Je v skladu s pričakovanji ali odstopa in zakaj?
- Kateri atributi so najbolj prispevali k takšni oceni?
- V čem se variante bistveno razlikujejo med seboj? Katere so bistvene prednosti in pomanjkljivosti posamezne variante?
- Kakšna je občutljivost ocene glede na spremembe atributov, njihovih vrednosti in funkcij koristnosti? Ali je mogoče in kako variante izboljšati oziroma poslabšati?

Šele z odgovori na ta vprašanja lahko pridemo do celovite slike o variantah in s tem do kvalitetnejše, boljše utemeljene in preverjene odločitve. V ta namen računalniška orodja ponujajo vrsto metod, ki podpirajo tovrstne analize in simulacije. V nadaljevanju bodo na primerih prikazane nekatere metode sistema DEX, ki so posebej pomembne za uporabo modelov v medicini:

- *metode za analizo modelov in predstavitev znanja*: prikaz funkcij koristnosti z izpeljanimi odločitvenimi pravili in metode za analizo uteži atributov [Rajkovič, Bohanec 1991];
- *metode vrednotenja variant* v primeru nepopolnih oziroma šumnih podatkov o variantah [Bohanec *et al.* 1983];
- *metode za razlago vrednotenja in analizo variant*: selektivna razlaga in analiza tipa "kaj-če" [Bohanec, Rajkovič 1995].

3. Ročna gradnja in uporaba hierarhičnih modelov

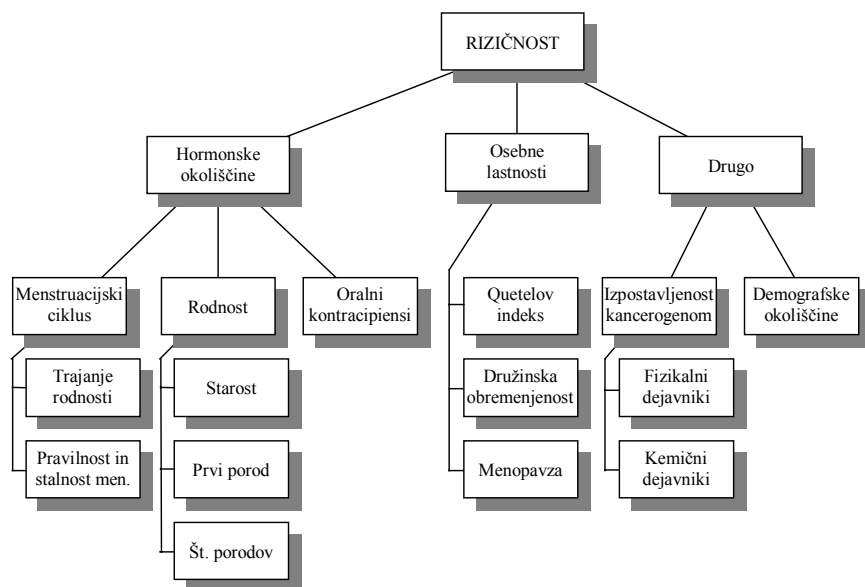
V Sloveniji trenutno potekajo štirje projekti, ki na tak ali drugačen način vključujejo hierarhične odločitvene modele, zgrajene "ročno". V tem razdelku so ti projekti na kratko predstavljeni. Opisana je struktura odločitvenih modelov. Na modelu s področja onkologije so prikazane različne metode vrednotenja, analize variant in razlage znanja, ki jih omogoča sistem DEX.

3.1 Onkologija: Ocena rizičnosti raka na dojki

Rak na dojki je nevarna in nepredvidljiva bolezen, pri kateri je izredno pomembno, da jo odkrijemo čim bolj zgodaj. V svetu se uveljavlja presejalni test (screening), ki zajema posebej ogrožene skupine žensk, kot so na primer ženske v določenih starostnih obdobjih ali ženske, družinsko obremenjene s to boleznijo. Te ženske, ki ne kažejo znakov bolezni in same po sebi ne bi iskale pomoči pri zdravniku, so povabljene na strokovni pregled oziroma mamografijo.

V sodelovanju z eksperti-onkologi je bil razvit prototip modela za ocenjevanje rizičnosti žensk za rak na dojki [Zavašnik 1997]. Strukturo modela prikazuje slika 2. Na osnovi podatkov, kot so starost, menstrualni cikel in rodnost, model oceni stopnjo rizičnosti in jo izrazi z vrednostjo na štiristopenjski lestvici.

Ocenjevanje poteka na osnovi odločitvenih pravil, ki so jih definirali eksperti. Tabela 1 prikazuje primer pravil za ocenjevanje rizičnosti s stališča menstruacijskega ciklusa v odvisnosti od trajanja rodnosti ter pravilnosti in stalnosti menstruacije. Pri tem je trajanje rodnosti bodisi kratko (do 30 let menstruacijskega obdobja), srednje (30 do 40 let) ali dolgo (nad 40 let), menstruacija pa je bodisi redna s periodo do 28 dni (R-28), redna s periodo nad 29 dni (R29+) ali neredna (N).



Slika 2: Struktura modela za ocenjevanje rizičnosti raka na dojki

	Trajanje rodnosti	Pravilnost in stalnost menstr.	Menstruacijski cikelus
1	srednje	R-28	zelo riz.
2	dolgo	R-28	zelo riz.
3	dolgo	R29+	zelo riz.
4	dolgo	N	zelo riz.
5	kratko	R-28	riz.
6	srednje	R29+	riz.
7	kratko	R29+	manj riz.
8	kratko	N	manj riz.
9	srednje	N	manj riz.

Tabela 1: Odločitvena pravila za oceno rizičnosti menstruacijskega ciklusa

Predstavljeni model je bil doslej verificiran s strani onkologov in preizkušen na manjšem vzorcu pacientk. V pripravi je obsežnejše ovrednotenje modela, ki bo potekalo na večjem številu pacientk in v primerjavi z rezultati neodvisnih strokovnih pregledov.

3.1.1 Analiza modela in predstavitev znanja

Prikaz metod za delo s hierarhičnimi modeli začnimo z metodami za analizo modelov in predstavitev znanja. Največ jih uporabljamo pri razvoju modelov zato, da bi preverili pravilnost oziroma vsebinsko ustreznost modelov ter odkrili morebitne napake. Metode delujejo tako, da celotni model oziroma

njegove komponente (funkcije koristnosti) pretvorijo v drugačne predstavitve, ki prikazujejo isto odločitveno pravilo na različne načine in na različnih nivojih podrobnosti.

Za generiranje različnih predstavitev posamezne funkcije koristnosti lahko uporabimo metode strojnega učenja. Tabela odločitvenih pravil (kot je tabela 1) interpretiramo kot množico učnih primerov in nad njo poženemo nek učni algoritem. Primer rezultata, dobljenega z algoritmom za učenje *izpeljanih pravil* [Rajkovič, Bohanec 1991], prikazuje tabela 2 (zvezdica v tabeli označuje katerokoli možno vrednost). Nova predstavitev je bolj kompaktna in v splošnem bolj razumljiva od originalne.

	Trajanje rodnosti	Pravilnost in stalnost menstr.	Menstruacijski cikel
1	≥ srednje	R-28	zelo riz.
2	dolgo	*	zelo riz.
3	kratko	R-28	riz.
4	srednje	R29+	riz.
5	kratko	R29+, N	manj riz.
6	≤ srednje	N	manj riz.

Tabela 2: Tabela 1, predstavljena z izpeljanimi pravili

RIZIČNOST	Regresija	Informativ.	Indeks Gini
Hormonske okoliščine	158	202	234
Menstruacijski cikel	125	123	130
Trajanje rodnosti	125	128	138
Prav./stal. menstr.	75	72	62
Rodnost	111	99	130
Starost	97	145	126
Prvi porod	145	128	145
Št. porodov	58	27	29
Oralni kontraceptivi	65	78	41
Osebnostne lastnosti	88	56	39
Quetelov indeks	29	5	11
Druž. obremenjenost	197	183	236
Menopavza	74	112	53
Drugo	55	42	27
Izpost. kancerogenom	100	100	100
Fizikalni dejavniki	160	166	179
Kemični dejavniki	40	34	21
Demogr. okoliščine	100	100	100

Tabela 3: Indeks pomembnosti atributov v onkološkem modelu, ocenjen s tremi metodami: regresijo, mero informativnosti in indeksom Gini. Najpomembnejši atributi so poudarjeni.

Kadar nas zanima manj podrobnosti o neki funkciji koristnosti, jo predstavimo s *povprečnimi relativnimi utežmi* atributov. Uteži izračunamo bodisi z izbrano regresijsko metodo [Rajkovič, Bohanec 1991], bodisi uporabimo katero od mer za ocenjevanje atributov v strojnem učenju, kot sta informativnost ali indeks Gini [Mingers 1989]. Na primer, pravila v tabeli 1 določajo povprečne uteži atributov “Trajanje rodnosti” in “Pravilnost in stalnost menstruacije”, ki ju regresijska metoda oceni na 64% in 36%. To pomeni, da je prvi atribut v *povprečju* nekaj manj kot dvakrat pomembnejši od drugega.

Predstavitev z utežmi je zanimiva predvsem na nivoju celotnega modela. Čeprav je le približna, omogoča hiter pregled vsebinske ustreznosti modela. Tabela 3 prikazuje uteži atributov onkološkega modela, ocenjene s tremi omenjenimi metodami. Zaradi večje preglednosti so uteži predstavljene z indeksom pomembnosti, ki podaja razmerje med dejansko utežjo atributa in utežjo, ki bi jo dobil ta atribut glede na svoje mesto v modelu, če bi bili vsi atributi enakovredni. Indeks 100 torej pomeni, da funkcije koristnosti v modelu niti ne zvišujejo, niti ne znižujejo pomembnosti tega atributa. Višji oziroma nižji indeks pa pomenita večjo oziroma manjšo relativno pomembnost atributa. Iz tabele 3 lahko razberemo, da k oceni rizičnosti najbolj prispeva sklop atributov “Hormonske okoliščine”. Med ostalimi atributi sta posebej pomembna tudi “Družinska obremenjenost” in izpostavljenost kancerogenim “Fizikalnim dejavnikom”.

3.1.2 Vrednotenje in analiza variant

Za prikaz uporabe modela vzemimo primer ženske, stare 42 let, ki ima 2 otroka, redno menstruacijo, povečano telesno težo (visok Quetelov indeks), ne uporablja oralnih kontracepcijskih sredstev in dela v okolju, ki s svojimi fizikalnimi dejavniki in demografskimi značilnostmi povečuje tveganje raka na dojki. Prvega otroka je rodila v starosti 25 let. V družini (mišljeni sta predvsem njena mati in sestra) ni primerov raka na dojki.

Rezultati ocenjevanja rizičnosti so prikazani v prvem stolpcu tabele 4. Ocena rizičnosti je 3, kar na štiristopenjski lestvici pomeni povečano, vendar še ne kritično stopnjo tveganja za rak na dojki. Razloge za takšno oceno lahko poiščemo bodisi s podrobnim pregledom vmesnih rezultatov v tabeli 4, bodisi z metodo *selektivne razlage* [Bohanec, Rajkovič 1995], ki v drevesu atributov poišče poddrevesa s posebej izrazitim – pozitivnim oziroma negativnim – vplivom na končno oceno pri danih podatkih. “Izrazitost” vpliva je opredeljena z mejnimi vrednostmi, ki jih za vsak atribut definira ekspert pri gradnji modela.

Selektivna razlaga našega primera (tabela 5) pokaže, da so najpomembnejši vzroki za povečano rizičnost: starost, telesna teža in predvsem negativni fizikalni in demografski dejavniki okolja.

	Osnovno vrednotenje	Manjkajoči podatki	Analiza "kaj-če"
RIZIČNOST	3	3	2
Hormonske okoliščine	2	<u>3/0,5,2/0,5</u>	2
Menstruacijski cikel	riz.	riz.	riz.
Trajanje rodnosti	srednje	srednje	srednje
Prav./stal. menstr.	R29+	R29+	R29+
Rodnost	riz.	riz.	riz.
Starost	nad 40 let	nad 40 let	nad 40 let
Prvi porod	do 29 let	do 29 let	do 29 let
Št. porodov	do 4	do 4	do 4
Oralni kontraceptivi	ne	*	ne
Osebnostne lastnosti	1	1	1
Quetelov indeks	29+	29+	29+
Druž. obremenjenost	ni	ni	ni
Menopavza	ne	ne	ne
Drugo	zelo.riz.	zelo.riz.	<u>riz.</u>
Izpost.kancerogenom	zelo.riz.	zelo.riz.	<u>riz.</u>
Fizikalni dejavniki	več	več	<u>manj</u>
Kemični dejavniki	ne	*	ne
Demogr. okoliščine	zelo.riz.	zelo.riz.	<u>riz.</u>

Tabela 4: Primeri vrednotenja in analize rizičnosti za rak na dojki

Razlogi ZA večjo rizičnost		Razlogi PROTI večji rizičnosti	
Starost	nad 40 let	Osebnostne lastnosti	1
Quetelov indeks	29+	Druž. obremenjenost	ni
Drugo	zelo.riz.	Menopavza	ne
Izpost.kancerogenom	zelo.riz.	Prvi porod	do 29 let
Fizikalni dejavniki	več	Oralni kontraceptivi	ne
Demogr. okoliščine	zelo.riz.	Kemični dejavniki	ne

Tabela 5: Selektivna razlaga vrednotenja

Tabela 4 prikazuje še dve dodatni možnosti, ki jih imamo pri uporabi modela. Model je možno uporabiti tudi v primeru nepopolnih in nenatančnih podatkov. Denimo, da za obravnavano osebo nimamo podatkov o kemijskih dejavnikih in uporabi oralnih kontraceptivov. Rezultate vrednotenja prikazuje drugi stolpec tabele 4. Končni rezultat je v tem primeru enak, edina razlika s prvim stolpcem pa je v manj natančni oceni hormonskih okoliščin, ki je izražena z verjetnostno porazdelitvijo vrednosti tega atributa.

Druga in za prakso zelo zanimiva možnost uporabe modela je analiza tipa "kaj-če". Zanima nas, kaj se spremeni pri oceni rizičnosti, če spremenimo enega ali več vhodnih podatkov oziroma katero izmed komponent modela. Pri obravnavani osebi nas na primer zanima ocena rizičnosti v primeru, da bi uspeli zmanjšati tveganje, ki izvira iz fizikalnih in demografskih dejavnikov. Tretji stolpec tabele 4 pokaže, da bi v tem primeru dejansko prišlo do zmanjšanja rizičnosti (končna ocena 2). Na podoben način bi lahko odgovorili tudi na druga zanimiva vprašanja, ki se porajajo pri tej osebi, na primer: kakšen vpliv bi imelo zmerno zmanjšanje telesne teže in kaj se bo zgodilo, ko bo vstopila v obdobje menopavze.

3.2 Radiologija: Tehnična analiza napak pri rentgenogramu pljuč

Tehnična kakovost rentgenograma pljuč bistveno vpliva na kvaliteto diagnostičnega postopka. Visoko kakovost je potrebno zagotoviti s čim manjšo dozo sevanja, ki jo prejme pacient. Za ocenjevanje kvalitete posnetkov je bil izdelan model, ki obsega 25 osnovnih kriterijev. Ti sodijo v dve osnovni skupini: tehnična kvaliteta naprav in postopka (prejeta doza, parametri RTG aparata, postopek v temnici) ter kvaliteta dela s pacientom (projekcija, vdih, kontrasti in podobno). Model je namenjen predvsem:

- predstavitvi znanja in zahtev na področju radiologije ter
- kot pripomoček pri izobraževanju študentov (samostojno eksperimentiranje in preverjanje znanja).

3.3 Patronažna nega: Ocena temeljnih življenjskih aktivnosti

Pomemben dejavnik v procesu patronažne zdravstvene nege so takoimenovane temeljne življenjske aktivnosti. Obstaja 14 teh aktivnosti, ki jih je potrebno pri vsakem pacientu spremljati, ocenjevati in vplivati nanje. V načrtu je razvoj odločitvenih modelov za te aktivnosti. Prototipno je izdelan model za odvajanje in izločanje, ki ga sestavlja 10 osnovnih in 5 izpeljanih atributov.

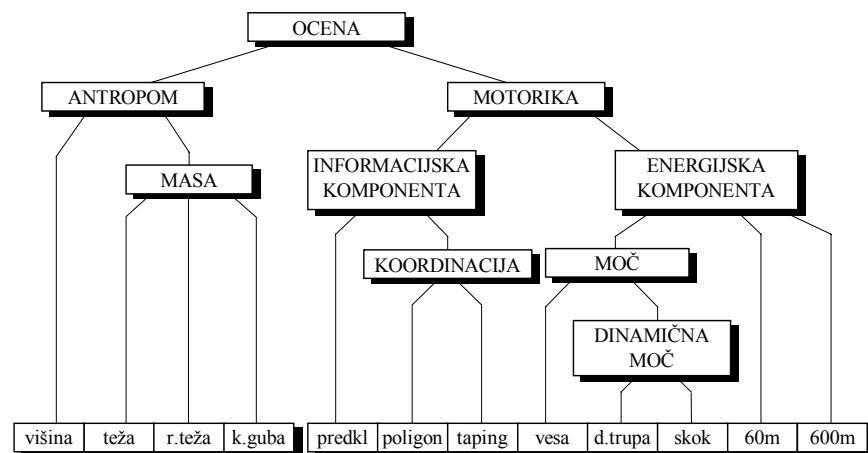
3.4 Usmerjanje otrok v športne panoge

Ta problem – predstavljen je bil na prejšnji delavnici CADAM-95 [Rajkovič *et al.* 1995] – sicer ne sodi neposredno na področje zdravstva, je pa povezan z njim, saj gre za ocenjevanje človekovih sposobnosti na osnovi antropometričnih in motoričnih dejavnikov. Na osnovi meritev, opravljenih v osnovnih in srednjih šolah v okviru t.i. Športnega kartona, odločitveni modeli vrednotijo sposobnosti učencev za doseganje dobrih rezultatov v 22 športnih panogah. Osnovna struktura atributov je prikazana na sliki 3.

Projekt je pomemben za nadaljnji razvoj odločitvenih modelov v zdravstvu predvsem iz naslednjih razlogov:

- gre za doslej najboljše sklop modelov, razvitih pri nas,
- modeli omogočajo vrednotenje in analizo nepopolnih podatkov,
- poseben poudarek je na razumljivosti modelov in razlagi rezultatov,
- uporabljena je kombinacija kvalitativnih (simboličnih) in kvantitativnih (numeričnih) modelov,
- modeli in njihova računalniška podpora (sistem Talent [Bohanec *et al.* 1997b]) so razviti do nivoja praktične uporabnosti ne samo za strokovnjake-kineziologe, ampak tudi za izobraževalno in strokovno delo učiteljev športne vzgoje v šolah.

Sistem Talent trenutno uporablja okrog 100 osnovnih in srednjih šol v Sloveniji.



Slika 3: Struktura modela za usmerjanje otrok v športne panoge

4. Strojno učenje hierarhičnih modelov iz podatkov

V tem razdelku so predstavljeni nekateri rezultati eksperimentov z metodo HINT za strojno učenje hierarhičnih modelov iz medicinskih podatkov. Najprej primerjamo klasifikacijsko točnost metode HINT s točnostjo nekaterih drugih metod s področja strojnega učenja. Drugi eksperiment, opravljen na podatkih s področja nevrofiziologije, pokaže sposobnost metode HINT, da v podatkih odkrije zanimive koncepte oziroma zakonitosti. V tretjem delu so prikazani rezultati rekonstrukcije odločitvenega modela iz nestrukturiranih podatkov, generiranih na osnovi ročno izdelanega onkološkega modela iz razdelka 3.1. Klasifikacijsko točnost rekonstruiranega modela smo primerjali z odločitvenimi drevesi, zgrajenimi z algoritmom C4.5. Predstavljen je tudi okvirni načrt nadaljnjih eksperimentov na področju ocenjevanja rizičnosti raka na dojki.

4.1 Klasifikacijska točnost

Klasifikacijsko točnost metode HINT smo preizkusili na sedmih znanih zbirkah medicinskih podatkov in primerjali dosežene rezultate z rezultati študij, predstavljenih v [Kononenko, Kukar 1995; Kononenko *et al.* 1997]. V teh študijah so bili med drugim preizkušeni učni algoritmi tipa Asistent (dve inačici), LFC (Lookahead Feature Construction), naivni Bayesov klasifikator in k -NN. Prvi trije algoritmi gradijo odločitvena drevesa. Prva inačica algoritma Asistent (I) ocenjuje atribute na osnovi njihovega informacijskega prispevka, druga (R) pa z metodo ReliefF.

Ker so bile zbirke podatkov šumne, je bila uporabljena inačica algoritma HINT, ki temelji na minimizaciji klasifikacijske napake [Zupan 1997]. HINT je bil uporabljen v nenadzorovanem, to je povsem avtomatskem načinu. Potek eksperimenta, vključno z diskretizacijo zveznih atributov, je bil enak, kot v [Kononenko *et al.* 1997]. Klasifikacijska točnost je bila ocenjena kot povprečje 30 eksperimentov, v katerih so bili primeri razdeljeni na učno in testno množico v razmerju 70:30.

	PRIM	BREA	LYMP	RHEU	HEPA	DIAB	HEART
HINT	36.9	79.5	74.1	66.3	79.0	73.7	79.6
Asistent-I	40.8	76.8	77.0	64.8	77.2	71.1	77.6
Asistent-R	38.9	78.5	77.0	63.8	82.3	71.5	77.6
LFC	37.1	76.1	82.4	60.6	79.0	69.2	77.3
naivni Bayes	48.6	78.7	84.7	66.5	86.1	76.3	84.5
k -NN	42.1	79.5	82.6	66.0	82.6	73.9	82.9

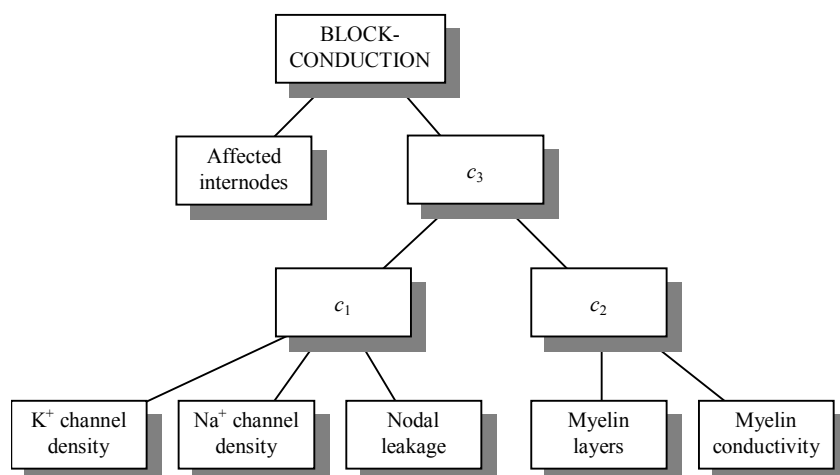
Tabela 6: Klasifikacijska točnost metod v sedmih medicinskih domenah

Rezultate eksperimenta prikazuje tabela 6. HINT je dosegel rezultate, ki so bili povsem primerljivi z ostalimi metodami. V primerjavi z Asistentom in LFC je bil na primer nekoliko slabši v domenah PRIM in LYMP, nekaj boljši v RHEU, DIAB in HEART ter približno enak v BREA in HEPA. Z drugimi besedami, nobena metoda ni dobila prepričljive prednosti. Edina, sicer dobro znana izjema na teh podatkih [Kononenko *et al.* 1997], je naivni Bayesov klasifikator, ki sicer dosega boljše klasifikacijske točnosti, vendar je manj primeren za razlago zakonitosti v podatkih.

V zvezi z razlago zakonitosti v podatkih velja opozoriti na možno težavo pri uporabi algoritma HINT. Kljub temu, da je HINT iz podatkov povsem samostojno zgradil številne hierarhične modele, ki so bili "videti" smiselni, jih nismo interpretirali s stališča vsebinske ustreznosti in razumljivosti. To delo namreč zahteva ekspertno znanje s konkretnega področja; le ekspert je tisti, ki lahko potrdi ali ovrže ustreznost novoodkritega atributa, ga poimenuje ter preveri pravilnost dobljenih funkcij koristnosti. Takšna interpretacija modelov ostaja izziv za nadaljnje delo na tem področju.

4.2 Nevrofiziologija: Učenje konceptov

Ekspertiment z algoritmom HINT, pri katerem je sodeloval tudi ekspert (J.A. Halter), je potekal na področju nevrofiziologije [Zupan et al. 1997b]. Naloga je bila zgraditi hierarhični odločitveni model, ki bi iz šestih osnovnih lastnosti mieliziranega živčnega vlakna sklepal na njegovo prevodnost. Učna množica je obsegala 3000 primerov. HINT, tudi tokrat uporabljen v nenadzorovanem načinu, je zgradil model, katerega strukturo prikazuje slika 4.



Slika 4: Odkrita struktura atributov za oceno prevodnosti živčnega vlakna

Model je interpretiral ekspert. Potrdil je, da so novoodkriti atributi (c_1 , c_2 in c_3) smiselni in da pravilno povzemajo biofizikalne lastnosti živčnega vlakna. Atribut c_2 je interpretiral kot lastnosti mielinske ovojnice in se strinjal s pripadajočo funkcijo koristnosti, ki opisuje monotono naraščajočo odvisnost toka od obeh vhodnih parametrov. Atribut c_1 združuje lastnosti vlakna, ki neposredno vplivajo na kapaciteto izvorov in ponorov ionskih tokov.

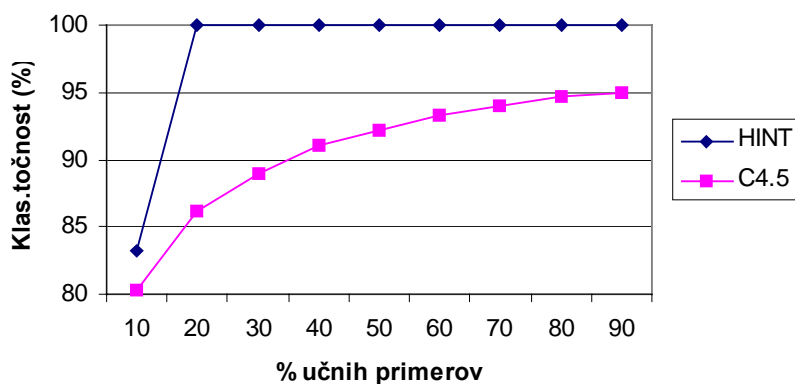
Ekspertiment je potrdil, da interpretacija strojno zgrajenih hierarhičnih modelov – ob primernem poznavanju problemskega področja – ni posebej težavna. HINT je v podatkih odkril relevantne zakonitosti in jih predstavil na način, ki je bil razumljiv strokovnjaku s tega področja.

4.3 Onkologija: Ocena rizičnosti raka na dojki

Na področju ocenjevanja rizičnosti raka na dojki je bil HINT preizkušen z naslednjo nalogo: rekonstruirati ročno izdelani odločitveni model iz nestrukturiranih primerov delovanja tega modela. Z modelom, predstavljenim v razdelku 3.1, je bilo najprej zgeneriranih 20000 naključnih primerov. Vsak

primer je bil opisan le z vrednostmi dvanajstih osnovnih atributov in vrednostjo razreda, ne pa tudi z vrednostmi izpeljanih atributov.

V vsakem koraku eksperimenta je bil iz celotne množice primerov slučajno izbran določen delež primerov in uporabljen kot učna množica za nenadzorovano inačico algoritma HINT z minimizacijo kompleksnosti. Klasifikacijska točnost dobljenega modela je bila ocenjena na preostalih primerih celotne množice. Za primerjavo je bil uporabljen algoritem za gradnjo odločitvenih dreves C4.5 [Quinlan 1993].



Slika 5: Učne krivulje C4.5 in HINT na onkoloških podatkih

Rezultati eksperimentov so prikazani z učnimi krivuljami na sliki 5. Pri rekonstrukciji je bil HINT izredno uspešen, saj je povsem točno rekonstruiral model že iz razmeroma majhnega števila učnih primerov. Pri tem je bistveno prekašal algoritem C4.5.

V okviru nadaljnjega dela na področju ocenjevanja rizičnosti raka na dojki je predvidenih še več eksperimentov z algoritmom HINT. Na Onkološkem inštitutu v Ljubljani so namreč izdelali študijo [Primic-Žakelj *et al.* 1995], v okviru katere so z anketo zbrali podatke o 1248 ženskah, od katerih jih je polovica zbolela za rakom na dojki, polovica pa je bilo zdravih. Te podatke bomo uporabili za:

- preverjanje ročno razvitega odločitvenega modela,
- avtomatizirano gradnjo novih klasifikatorjev z metodo HINT in drugimi metodami strojnega učenja ter
- medsebojno primerjavo ročno in strojno razvitih modelov.

5. Zaključek

Hierarhični odločitveni modeli se vse bolj uveljavljajo tudi na področju zdravstva. S stališča praktičnih aplikacij so posebej pomembne naslednje lastnosti modelov in zmožnosti podpornih računalniških orodij:

- strukturirana predstavitev znanja o problemski domeni,
- možnost uporabe kvalitativnih modelov, definiranih s simboličnimi spremenljivkami in odločitvenimi pravili,
- analiza modela in vpliva posameznih dejavnikov na končno oceno,
- vrednotenje oziroma klasifikacija novih primerov, tudi ob nepopolnih in nenatančnih podatkih,
- razlaga rezultatov vrednotenja oziroma klasifikacije.

Doslej prevladujoči ročni način izgradnje hierarhičnih modelov dopolnjuje metoda strojnega učenja HINT. Klasifikacijska točnost strojno zgrajenih modelov je primerljiva, v primerih strukturiranih problemskih domen pa celo bistveno boljša od metod, ki gradijo odločitvena drevesa. HINT je sposoben odkrivati nove, relevantne koncepte. Metoda je posebej uporabna za odkrivanje zakonitosti v podatkih, kjer postopek odkrivanja hierarhične strukture usmerja ekspert, medtem ko pri povsem avtomatskem učenju iz podatkov lahko nastopi problem interpretacije dobljenih modelov.

Literatura

- Bohanec, M., Bratko, I., Rajkovič, V. (1983): An expert system for decision making. *Processes and tools for decision support* (ed. Sol, H.G.), North-Holland, 235–248.
- Bohanec, M., Rajkovič, V. (1990): DEX: An expert system shell for decision support. *Sistemica* 1, 145–157.
- Bohanec, M., Rajkovič, V. (1995): Večparametrski odločitveni modeli. *Organizacija* 28, 427–438.
- Bohanec, M., Zupan, B., Bratko, I., Cestnik, B. (1997a): A function-decomposition method for development of hierarchical multi-attribute decision models. *Proceedings of the Fourth Conference of the International Society for Decision Support Systems ISDSS-97*, Lausanne, 503–514.
- Bohanec, M., Kapus, V., Leskošek, B., Rajkovič, V. (eds.) (1997b): *Talent: Uporabniški priročnik*. Ministrstvo za šolstvo in šport in Zavod Republike Slovenije za šolstvo.
- Brans, J.P., Mareschal, B. (1990): The PROMETHEE methods for MCDM. *Readings in multiple-criteria decision aid* (ed. Bana e Costa, C.A.), Springer, 216–252.
- Cestnik, B., Kononenko, I., Bratko, I. (1987): ASSISTANT 86: A knowledge elicitation tool for sophisticated users. *Progress in machine learning* (eds. Bratko, I., Lavrač, N.), 31–45, Sigma Press.
- Chankong, V., Haimes, Y.Y. (1983): *Multiobjective decision making: Theory and methodology*. North-Holland.

- Humphreys, P., McFadden, W. (1980): Experiences with MAUD: Aiding decision structuring versus bootstrapping the decision maker. *Acta Psychologica* 45, 51–69.
- Kononenko, I., Kukar, M. (1995): Machine learning for medical diagnosis. *Računalniška analiza medicinskih podatkov CADAM-95* (ed. Lavrač, N.), IJS Scientific Publishing IJS-SP-95-1, 9–30.
- Kononenko, I., Šimec, E., Robnik Šikonja, M. (1997): Overcoming the myopia of inductive learning algorithms with ReliefF. *Applied Intelligence Journal* 7(1), 39–56.
- Mallach, E.G. (1994): *Understanding decision support systems and expert systems*. Irwin.
- Mingers, J. (1989): An empirical comparison of selection measures for decision-tree induction. *Machine Learning* 3, 319–342.
- Nagel, S. (ed.) (1993): *Computer-aided decision analysis: Theory and application*. Quorum Books.
- Primic-Žakelj, M., Evstifeeva, T., Ravnihar, B., Boyle, P. (1995): Breast-cancer risk and oral contraceptive use in Slovenian women aged 25 to 54. *Int. J. Cancer* 62, 414–420.
- Quinlan, J.R. (1993): *C4.5: Programs for machine learning*. Morgan Kaufmann Publishers.
- Rajkovič, V., Bohanec, M. (1991): Decision support by knowledge explanation. *Environments for supporting decision processes* (eds. Sol, H.G., Vecsenyi, J.), North-Holland, 47–57.
- Rajkovič, V., Bohanec, M., Leskošek, B., Kapus, V. (1995): Zasnova ekspertnega sistema za usmerjanje šolske mladine v športne panoge. *Računalniška analiza medicinskih podatkov CADAM-95* (ed. Lavrač, N.), IJS Scientific Publishing IJS-SP-95-1, 234–244.
- Saaty, T.L. (1993): *Multicriteria decision making: The analytic hierarchy process*. RWS Publications.
- Stewart, T.J. (1992): A critical survey on the status of multiple criteria decision making theory and practice. *Omega, Intl. J. of Mgmt. Sci.*, 20(5/6), 569–586.
- Urbančič, T., Kononenko, I., Križman, V. (1991): *Review of applications by Ljubljana Artificial Intelligence Laboratories*. Institut Jožef Stefan, delovno poročilo DP-6218.
- Vincke, P. (1992): *Multicriteria decision-aid*. John Wiley & Sons.
- Zavašnik, A., Karlo, S., Vagaja, N. (1997): *Uporaba ekspertnega sistema pri ugotavljanju rizičnosti raka dojk*. Seminarska naloga. Univerza v Mariboru, Fakulteta za organizacijske vede Kranj.
- Zupan, B. (1997): *Machine learning based on function decomposition*. Ph.D. Thesis, University of Ljubljana, Faculty of Computer Science and Informatics.
- Zupan, B., Bohanec, M., Bratko, I., Demšar, J. (1997a): Machine learning by function decomposition. *Proceedings of ICML-97* (ed. Fisher, D.H.), Morgan Kaufman Publishers, 421–429.
- Zupan, B., Halter, J.A., Bohanec, M. (1997b): Concept discovery by decision table decomposition and its application in neurophysiology. *Intelligent data analysis in medicine and pharmacology* (eds. Lavrač, N., Keravnou, E., Zupan, B.), Kluwer Academic Publishers, 261–278.