

Primerjava treh tipov modelov za napovedovanje uspešnosti zaključka šolanja

Silvana Gasar¹, Marko Bohanec² in Vladislav Rajkovič²
¹Srednja šola Jesenice, Jesenice
²Institut Jožef Stefan, Ljubljana

Povzetek: Za učinkovito preprečevanje neustreznih izbir srednjih šol in s tem učne neuspešnosti šolski svetovalci potrebujejo orodje za napovedovanje uspešnosti zaključka šolanja posameznega dijaka. S postopki rudarjenja podatkov na bazi dijakov in z ekspertnim modeliranjem smo izdelali več modelov za napovedovanje uspešnosti zaključka šolanja na posamezni srednješolski smeri. Med postopki rudarjenja smo uporabili statistične analize, razvrščanje v skupine in dve metodi strojnega učenja: izgradnjo klasifikacijskih odločitvenih dreves in hierarhičnih odločitvenih modelov. Ekspertni sistem, katerega jedro predstavlja hierarhični večparametrski odločitveni model, smo razvili »ročno«, s pomočjo lupine ekspertnega sistema DEX. Izgrajene modele smo validirali in ovrednotili njihovo uporabnost. Napovedna točnosti modelov DEX in odločitvenih dreves je enaka in zelo zadovoljiva, saj dosega točnost izkušenega svetovalca. Glede na zahtevnost in težave pri izdelavi modelov ter učinkovito prilagajanje na relativno hitre spremembe našega šolskega sistema v nadaljnjem razvoju napovednih modelov predlagamo uporabo odločitvenih dreves.

Ključne besede: učni uspeh, napovedovanje, poklicno usmerjanje, odločanje, ekspertni sistemi, modeli podatkovno rudarjenje

Comparison of tree types of models for the prediction of final academic achievement

Silvana Gasar¹, Marko Bohanec² and Vladislav Rajkovič²
¹Jesenice Secondary School, Jesenice, Slovenia
²Jožef Stefan Institute, Ljubljana, Slovenia

Abstract: For efficient prevention of inappropriate secondary school choices and by that academic failure, school counselors need a tool for the prediction of individual pupil's final academic achievements. Using data mining techniques on pupils' data base and expert modeling, we developed several models for the prediction of final academic achievement in an individual high school educational program. For data mining, we used statistical analyses, clustering and two machine learning methods: developing classification decision trees and hierarchical decision models. Using an expert system shell DEX, an expert system, based on a hierarchical multi-attribute decision model, was developed manually. All the models were validated and evaluated from the viewpoint of their applicability. The predictive accuracy of DEX models and decision trees was equal and very satisfying, as it reached the predictive accuracy

**Naslov / address: mag.Silvana Gasar, univ. dipl. psih., Srednja šola Jesenice, Ruparjeva 2, Jesenice, Slovenija, e-mail: silvana.gasar@telesat.si; domača stran na Internetu: <http://silvana.telesat.si/>*

of an experienced counselor. With respect on the efficiency and difficulties in developing models, and relatively rapid changing of our education system, we propose that decision trees are used in further development of predictive models.

Key words: academic achievement, prediction, occupational guidance, decision making, expert systems, models, data mining

CC=3580

Domnevamo, da problem učne neuspešnosti v srednji šoli (SŠ) izvira predvsem iz neustrezne izbire šole, katere zahteve niso usklajene z lastnostmi učenca. O izbiri šole in poklica govorijo številne teorije (Brančič, 1986; Rodić, 1980), nanjo pa lahko gledamo tudi kot na proces večparametrskega odločanja. Zaradi negotovosti v procesu odločanja vedno obstaja tveganje, da bo izbor šole nekvaliteten, posledice pa negativne, saj kvaliteto odločitve pokaže šele njena realizacija (Rajkovič in Bohanec, 1988).

Za kakovosten izbor SŠ učenci nujno potrebujejo strokovno pomoč (Oman, 1976; Rodić, 1980). Za izobraževalno in poklicno svetovanje pri nas skrbijo predvsem šolske svetovalne službe, toda svetovalci so preobremenjeni z drugimi nalogami (Resman, Bečaj, Bezić, Čačinovič-Vogrinčič in Musek, 1999). Sistematičnih raziskav o izvajanju izobraževalnega svetovanja pri nas in drugje v svetu skorajda ni. Ocena ustreznosti izbire šole je najpogosteje intuitivna. Za primerno ukrepanje, ki bi preprečilo učno neuspešnost, je v času vpisa in prijave na SŠ že prepozno (Lapajne, 1984). Takrat svetovalci lahko dijaka samo opozorijo na verjetnost neustrezne izbire, če tako ocenjujejo na podlagi skromnih in pogosto nepopolnih podatkov, toda le redki dijaki poslušajo opozorila. Svetovalci bi potrebovali orodje, s katerim bi lahko zanesljivo napovedali uspešnost zaključka šolanja in napoved razumljivo razložili dijaku in njegovim staršem.

Napoved učnega uspeha omogoča preprečevanje neuspeha (Lapajne, 1984), posredno pa tudi mnogih nezaželenih družbenih pojavov, kot so: nezaposlenost, delinkventnost, narkomanija, samomorilnost (Resman in dr., 1999). Redke raziskave napovedne veljavnosti podatkov o učencih za slovenske SŠ so pokazale, da šolske ocene lahko služijo kot prediktorji bodočega učnega uspeha, in da velja »načelo najšibkejšega člana verige«, po katerem najvišji splošni učni uspeh omejujejo zlasti predmeti, v katerih so učenci najšibkejši (Lapajne, 1984). Na splošno je učni uspeh odvisen od interakcije fizičnih, fizioloških, socialnih in psiholoških dejavnikov učenja (Musek in Pečjak, 1995), o katerih imajo svetovalci malo podatkov in jih pogosto tudi ne znamo veljavno izmeriti, zato je pri napovedovanju uspešnosti zaključka šolanja nujno potrebno večparametrsko ocenjevanje po kakovostnih kriterijih. Pri tem se šolski svetovalci lahko poslužujejo različnih sodobnih informacijskih tehnologij in metodologij. Baze podatkov o dijakih skrivajo zakonitosti učnega uspeha, ki jih lahko odkrijemo s postopki rudarjenja podatkov. Odkrivanje znanja iz podatkov (Knowledge Discovery

from Data) je netrivialen proces odkrivanja implicitnega, doslej neznanega in potencialno uporabnega znanja iz podatkov, ki poteka po naslednjih fazah: 1. priprava podatkov, 2. rudarjenje podatkov (Data Mining), 3. interpretacija, vrednotenje in predstavitev (Han in Kamber, 2001). Presega omejitve tradicionalnih statističnih metod (Michalski in Kaufman, 1998; Viera, 2000), vendar jih vključuje v svoj okvir (Hand, 1998, Maindonald, 1999). Cilj je pridobivanje novega znanja za boljše odločanje, klasifikacijo ali predikcijo (Michalski in Kaufman, 1998). Najpomembnejše metode rudarjenja so: statistične metode, vizualizacija, metode strojnega učenja, gradnja asociacijskih pravil in razvrščanje v skupine (Han in Kamber, 2001). Razvrščanje v skupine primere razvrsti po »podobnosti«, tako da so objekti znotraj skupin čimbolj podobni in objekti različnih skupin čimbolj različni (Bohanec, 2002; Ferligoj, 1989). Omogoča redukcijo, zgoščevanje podatkov, olajša pregledovanje in določitev tipologije (Ferligoj, 1989; Vieira, 2000). Strojno učenje, tj. učenje, pri katerem »se uči« stroj-računalnik (Kononenko, 1997) - računalniški program (»učni algoritem«) se uči na bazi podatkov preteklih rešenih primerov (Mitchell, 1997) - omogoča avtomatsko izgradnjo (običajno tudi validacijo, najpogosteje z navzkrižnim preverjanjem 10. reda) napovednih modelov, kakršna so npr. odločitvena drevesa in hierarhični modeli HINT.

Osnovna naloga algoritma (računalniški program) za gradnjo odločitvenih dreves je iz tabelarično predstavljenih podatkov (»učnih primerov«) zgraditi odločitveno drevo, ki ponazarja preslikavo med izbranimi vhodnimi polji ali »atributi« (s pomočjo katerih skušamo napovedovati) tabele in izbranim izhodnim poljem ali »razredom« (atribut, ki ga skušamo napovedati - v našem primeru uspešnost zaključka šolanja). Vhodni atributi so lahko diskretni ali zvezni. V odvisnosti od tega, ali je razred diskretna ali zvezna spremenljivka, ločimo med klasifikacijskimi in regresijskimi odločitvenimi drevesi. Algoritem izbere najboljši atribut A (po eni od mer nečistoče ali količini informacije) in glede na vrednosti A razdeli učno množico ter rekurzivno zgradi poddrevesa za vsako podmnožico (Bohanec, 2002; Kononenko, 1997; Mitchell, 1997). Odločitveno drevo je sestavljeno iz: 1) notranjih vozlišč, ki ustrezajo pogojem, 2) vej, ki ustrezajo podmnožicam vrednosti atributov in 3) listov, ki ustrezajo razredom (Kononenko, 1997; Mitchell, 1997). Odločitvena drevesa se pogosto uporabljajo za klasifikacijo in predikcijo – nove primere klasificiramo tako, da od korena potujemo navzdol po ustreznih vejah, sledimo pogojem v notranjih vozliščih in pridemo do lista, ki klasificira primer (Kononenko, 1997, Mitchell, 1997). Lahko se ukvarjajo s številnimi atributi, so razmeroma neobčutljiva na manjkajoče podatke in »šum« v podatkih ter so lahko razumljiva (Mitchell, 1997). Kot mere kvalitete odločitvenih dreves se uporabljajo: klasifikacijska točnost (delež pravilno klasificiranih testnih primerov), razumljivost in velikost drevesa (Bohanec, 2002, Clark in Niblett, 1987). Zaradi prevelike prilagoditve učnim podatkom so spodnji deli dreves manj zanesljivi (nereprezentativni), kar rešujemo z rezanjem, ki praviloma da manjše, bolj razumljivo drevo, z višjo klasifikacijsko točnostjo (Kononenko, 1997; Mitchell, 1997). HINT (Hierarchical INduction Tool) (Zupan, 1997) se lahko uporablja za avtomatizirano konstrukcijo (in validacijo) hierarhičnih odločitvenih modelov tipa DEX (glej sliko 3), ki jih omenjamo v nadaljevanju. Na bazi

»učnih podatkov« (preteklih rešenih primerov) HINT odkrije smiselne hierarhije konceptov – modele z dobro klasifikacijsko točnostjo (Bohanec, Zupan in Rajkovič, 2000). Razvoj modela je lahko avtonomen (popolnoma avtomatiziran, delo v celoti opravi računalniški program oz. algoritem) ali pa poteka v interakciji z odločevalcem, v procesu nadzorovane dekompozicije (v vsaki stopnji razvoja modela človeški strokovnjak izbira najbolj smiselno od možnih kombinacij podkonceptov v koncepte, poimenuje koncepte ipd. Bohanec, Zupan, Bratko in Cestnik, 1997).

Napovedne modele lahko razvijemo tudi v okviru izgradnje ekspertnih sistemov. Ekspertni sistemi so računalniški programi, ki delujejo podobno kot človeški eksperti – posnemajo proces človeškega sklepanja in se pogosto približajo odločitvam ljudi (Mallach 1994). Pripomorejo k večji enotnosti in kvaliteti odločanja. Vsebuje bazo znanja, mehanizem sklepanja in uporabniški vmesnik (Dutta, 1993; Mallach, 1994). Lahko jih razvijemo z uporabo lupin, kot je npr. v nalogi uporabljeni DEX (Bohanec in Rajkovič, 1995, 1999). Lupino ekspertnega sistema sestavljata mehanizem sklepanja (računalniški program) in uporabniški vmesnik (kar na zaslonu vidi uporabnik), bazo znanja pa izdelata uporabnik sam (Dutta, 1993; Mallach, 1994). Bazo znanja za ekspertni sistem lahko sestavimo »ročno«, tako da vanjo vnesemo znanje človeškega eksperta (najpogosteje v obliki produkcijskih oz. odločitvenih pravil »če- potem« - glej sliko 4) ali pa jo izgradimo avtomatsko z uporabo metod strojnega učenja, kakršna je npr. HINT. Določena znanja in zakonitosti, ki jih »ročno« vnesemo v bazo znanja, lahko pridobimo tudi z rudarjenjem podatkov. Pred uporabo mora biti vsak ekspertni sistem validiran, po potrebi izboljššan in revalidiran (Mallach, 1994).

Glavni cilj raziskave je bil s pomočjo rudarjenja podatkov in ekspertnega modeliranja izdelati več različnih modelov za napovedovanje uspešnosti zaključka šolanja po določenem srednješolskem programu in smeri, jih preizkusiti in ovrednotiti njihovo uporabnost. Problem raziskave opredeljujejo naslednja vprašanja:

- Ali je iz baze podatkov o dijakih mogoče ugotoviti splošne zakonitosti učnega uspeha?
- Ali, in kako, je na podlagi teh zakonitosti mogoče zgraditi napovedne modele?
- Kakšna je napovedna točnost in uporabnost takih modelov?
- Ali, in kako, taki modeli prispevajo k boljšemu napovedovanju uspešnosti?
- Ali lahko izboljšajo uspešnost samo?
- Kako jih uvesti v vsakodnevno prakso?

Metoda

Raziskava je zajela baze podatkov o dijakih treh različnih smeri izobraževanja, zajete v šestih zaporednih šolskih letih (1993/94 – 1998/99), v eni od slovenskih SŠ s programom Evidenca 3 (Galle, 1996). Uporabili smo različne računalniške programe za rudarjenje podatkov in izgradnjo modelov. Osnovni postopek je vključeval odkrivanje

zakonitosti učnega uspeha z rudarjenjem podatkov in avtomatsko izgradnjo napovednih modelov s pomočjo strojnega učenja. Sledila je »ročna« izgradnja napovednih modelov po metodi ekspertnega modeliranja, tej pa evalvacija in primerjava različnih modelov.

Rudarjenje podatkov in avtomatska izgradnja modelov

V opisani raziskavi smo uporabili statistične metode, vizualizacijo (grafično prikazovanje), razvrščanje v skupine (clustering) in dve metodi strojnega učenja: gradnjo odločitvenih dreves in hierarhičnih odločitvenih modelov HINT. V ta namen smo uporabljali računalniške programe SPSS (SPSS Inc., 2002), Weka (University of Waikato, 2002) in Orange (Zupan in Demšar, 2002). Rezultate smo interpretirali in ovrednotili z vidika nivoja tveganja, klasifikacijske točnosti, velikosti modelov, smiselnosti, novosti in uporabnosti.

Priprava podatkov

Baze dijakov smo pripravili z računalniškim programom SQL Server 2000 (Vieira, 2000). Končni rezultat je bila tabela podatkov o dijakih, opisanih s 96 atributi. Prvih 19 atributov je bilo poznanih pred vpisom v SŠ, npr. spol, datum rojstva, leto zaključka OŠ, ocene in uspeh OŠ. Sledilo je 77 atributov, pridobljenih v zaporednih letnikih SŠ, kot so: smer izobraževanja (»N«, »L« ali »R«), izostanki, disciplinski ukrepi in različne ocene. Ob koncu 1. letnika poznamo dodatnih 20 atributov, ob koncu 2. letnika 18, ob koncu 3. letnika 17 in ob koncu 4. letnika 22 dodatnih atributov. Glavne skupine atributov so bile:

- Dijakovi osebni in demografski podatki: priimek, ime, spol, datum in kraj rojstva, državljanstvo, EMŠO, prebivališče, ime osnovne šole itd.
- Podatki o učnem uspehu v OŠ: ocene posameznih predmetov (slovenščine, matematike, fizike in tujega jezika) in splošnega učnega uspeha v 7. in 8. razredu OŠ.
- Podatki o učnem uspehu in vedenju v prvem, drugem, tretjem in četrtem letniku SŠ: ocene posameznih predmetov (slovenščine, matematike, tujega jezika...), splošni učni uspeh letnika, izrečeni disciplinski ukrepi, ure opravičenih in neopravičenih izostankov itd.

Ker je bilo v bazi ($N = 1794$; smer »L«: $n = 1058$; smer »R«: $n = 395$; smer »N«: $n = 341$) precej manjkajočih podatkov po posameznih letnikih SŠ, smo izdelali tudi manjšo bazo s popolnimi podatki ($N = 889$, smer »L«: $n = 468$; smer »R«: $n = 250$; smer »N«: $n = 171$), uporabili pa smo tudi dele posamezne baze z ozirom na smer izobraževanja. V nadaljevanju je navedeno, na katerih podatkih so bili dobljeni posamezni rezultati. Po pripravi je vsak dijak v bazi imel uspešnost zaključka šolanja izraženo z eno od naslednjih kategorij (»razredom«):

- 5 = v 4 letih je zaključil šolo s prav dobrim ali odličnim uspehom v 4. letniku
- 4 = v 4 letih je zaključil šolo z dobrim ali zadostnim uspehom v 4. letniku
- 3 = v 5 - 6 letih je zaključil šolo s pozitivnim uspehom v 4. letniku
- 2 = po 1 - 2 letih je prekinil šolanje z negativnim uspehom ali kot neocenjen
- 1 = po 3 ali več letih je prekinil šolanje z negativnim uspehom ali kot neocenjen.

Statistične obdelave

Rudarjenje smo pričeli z ugotavljanjem osnovnih statistik (deskriptivne statistike, frekvenčne porazdelitve), mer povezanosti med spremenljivkami (rang korelacije, hi-kvadrat test, koeficient kontingence) in vizualizacijo (histogrami ipd.) v programu SPSS.

Gradnja odločitvenih dreves

Klasifikacijska odločitvena drevesa smo gradili z računalniškim programom Weka, z uporabo algoritma J4.8 (Witten in Frank, 2000), verzijo znanega Quinlanovega algoritma C4.5 (Quinlan, 1993), ki najboljše atribute izbira na podlagi mere razmerja informacijskega prispevka (gain ratio). Najprej smo zgradili drevo iz vseh atributov, zatem pa še različna odločitvena drevesa na podlagi ekspertne (atribute je izbral človeški ekspert), strojne (atribute je izbiral programski algoritem ReliefF) in pragmatično-strojne (ReliefF je izbiral iz omejene podskupine atributov glede na čas, ko jih poznamo) izbire najboljših atributov. Iskali smo kompromis med čimprejšnjo in čim bolj točno napovedjo (točnost napovedi s časom raste, njena uporabnost pa pada), zato smo v gradnjo dreves vključili attribute, znane vsaj do konca 2. letnika SŠ. Zaradi lažjih primerjav smo drevesa vedno gradili iz 16, 30 in 46 atributov (podlaga za določitev števila atributov je bila prvotna ekspertna izbira), pri tem pa spreminjali parametre rezanja. Klasifikacijsko točnost smo ugotavljali s prečnim preverjanjem 10. reda, pri katerem množico primerov razdelimo na 10 približno enakih podmnožic. Za vsako podmnožico zgradimo drevo tako, da za učenje uporabimo unijo preostalih podmnožic (9/10). Zgrajeno drevo uporabimo za klasificiranje primerov dane podmnožice (1/10). Klasifikacijsko točnost drevesa, ki ga zgradimo iz vseh razpoložljivih primerov, ocenimo kot povprečno točnost vseh 10 dreves.

Razvrščanje v skupine

Z računalniškim programom Weka smo po metodi voditeljev (k -means) dijake razvrstili na 3, 4 in 5 skupin (število skupin in vhodne spremenljivke je določil ekspert). Ustreznost opredelitve kategorij uspešnosti zaključka šolanja smo preverili z razvrščanjem v skupine po dveh spremenljivkah, ki opredeljujeta kategorije uspešnosti: trajanju šolanja in zaključnem uspehu 4. letnika SŠ (frekvenčno porazdelitev dobljenih skupin smo primerjali s frekvenčno porazdelitvijo kategorij uspešnosti). Nadaljevali smo z razvrščanjem v skupine na podlagi 16 ekspertno izbranih atributov, znanih pred vpisom

v SŠ, in na podlagi 30 ekspertno izbranih atributov (ki vključujejo 12 od prvotnih 16 atributov), znanih do konca 1. letnika SŠ, saj so nas zanimale tudi morebitne razlike razvrstitev v nadaljevanju SŠ. Smiselnost dobljenih razvrstitev je na podlagi znanja in v povezavi v ostalih rezultati ocenil človeški ekspert – šolski svetovalec.

HINT: učenje hierarhičnih modelov

Z računalniškim program Orange (Zupan in Demšar, 2002) smo na manjši bazi zgradili več modelov HINT iz 16 in 30 ekspertno izbranih atributov (nenadzorovana dekompozicija, metoda minimalne napake) in jih validirali z navzkrižnim preverjanjem 10. reda (klasifikacijska točnost). Smiselnost zgrajenih hierarhičnih konceptov je ocenil človeški ekspert.

»Ročna« izgradnja ekspertnega sistema

Napovedni model smo zgradili z računalniškim programom DEX – lupino ekspertnega sistema za večparametrsko odločanje (Bohanec in Rajkovič, 1995, 1999). DEX omogoča izgradnjo lahko razumljivih kakovostnih modelov (znanje o odločanju je izraženo neposredno z besedami, pravili in hierarhično urejenimi kriteriji), ukvarjanje z netočnimi, negotovimi in manjkajočimi podatki o variantah (v našem primeru so to podatki o dijakih), kvalitetno verifikacijo in razlago rezultatov. Hierarhični odločitveni model DEX lahko razvijemo »ročno« (znanje eksperta vnesemo v bazo znanja) ali pa avtomatsko z uporabo strojnega učenja, kot je npr. metoda HINT (Bohanec in dr., 2000). Uporabili smo kombinacijo obeh pristopov: v »ročni« razvoj modela smo zajeli ekspertno znanje in znanje, dobljeno s postopki rudarjenja.

Sledili smo metodologiji izgradnje modelov DEX (Bohanec in Rajkovič, 1995). Oblikovali smo spisek osnovnih kriterijev (podatkov o dijaku), jih na podlagi vsebinske povezanosti strukturirali v hierarhično drevo, jim določili merske lestvice (zaloge vrednosti na sliki 3) in funkcije koristnosti (slika 4), ki določajo, kako se podkriteriji (npr. »sposobnosti in predznanje«, »motivacija« in »okolščine«) kombinirajo v sestavljen nadrejeni kriterij (npr. »uspešnost«). Pri zalogi vrednosti »uspešnosti« smo ohranili 5 vrednosti, vendar smo jih spremenili tako, da sta glede na kategorije uspešnosti v postopkih rudarjenja oznaki 1 in 2 zamenjani (DEX oznaka 1 je bila prej kategorija 2, DEX oznaka 2 pa kategorija 1). Za zamenjavo smo se odločili zato, ker ordinalna urejenost zalog vrednosti DEX olajšuje postopek agregacije končne ocene. Dijaki, ki prenehajo z izobraževanjem prej in v nižjih letnikih, so manj uspešni kot dijaki, ki »prilezejo« do višjega letnika, vendar izobraževanja ne zaključijo uspešno. S subjektivnega vidika je DEX ocena 2 najslabša možnost, saj pomeni, da je dijak v izobraževanje vložil precej časa in truda, vendar je bil neuspešen. »Prednost« DEX ocene 1 je v tem, da je dijak tako očitno neuspešen že v samem začetku šolanja, da se bo skoraj zagotovo prešolal – možnosti prešolanja so takrat večje, časa in truda pa tudi ni vložil toliko kot prejšnji.

Funkcije koristnosti v DEX lahko definiramo s pomočjo obtežitev podkriterijev po presoji eksperta področja (v našem primeru šolskega svetovalca z 10 let izkušenj), ali po točkah s pomočjo tabel odločitvenih pravil »če-potem«. Uporabili smo drugi pristop. DEX sam pripravi vse možne kombinacije podkriterijev v tabeli, zato je ekspert določil samo vrednosti sestavljenega kriterija. Spisek kriterijev smo izoblikovali na podlagi ekspertnega mnenja, najboljših odločitvenih dreves in modelov HINT. Merske lestvice kriterijev je določil ekspert na podlagi presoje. Ker smo na podlagi predhodnih rezultatov domnevali, da bo napovedna točnost modela na podlagi podatkov, znanih pred vpisom v SŠ, nizka, smo se odločili za gradnjo dveh modelov: primarnega za napovedovanje uspešnosti ob vpisu in sekundarnega za napovedovanje uspešnosti ob koncu 1. letnika SŠ. Najprej smo zgradili dva splošna modela DEX, ki smo ju na podlagi poskusnih vrednotenj in analiz primerov preoblikovali in prilagodili smeri izobraževanja »L«.

Validacija modelov

Kot je običajno, smo klasifikacijsko točnost avtomatsko zgrajenih modelov ugotavljali z navzkrižnim preverjanjem 10. reda. Za validacijo modelov DEX smo uporabili 10% slučajni stratificirani vzorec ($N = 47$) iz dela manjše baze samo za smer »L«. Vse primere smo ocenili s primarnim in sekundarnim modelom DEX in glede na dejansko uspešnost izračunali klasifikacijsko točnost. Na podlagi enakih podatkov in kategorij uspešnosti, kot jih uporablja DEX, je ocene uspešnosti podal tudi prej omenjen šolski svetovalec. Zaradi obsežnosti dela je bil vključen samo en svetovalec, ki je podatke o dijakih dobil v obliki preglednice »na papirju«, identiteta dijakov pa mu je bila prikrita. Klasifikacijsko točnost izgrajenih modelov smo primerjali z apriorno točnostjo (delež primerov najbolj zastopane kategorije uspešnosti) in s točnostjo svetovalca.

Rezultati in razprava

Rezultati statističnih metod

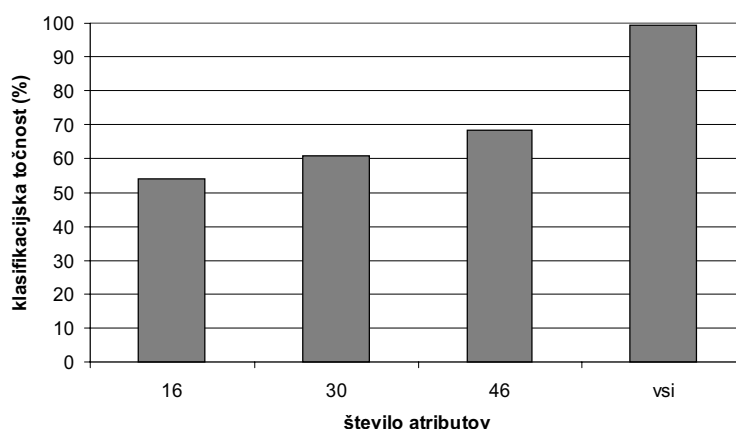
Osnovni vzorec je vključeval 1794 dijakov, večinoma moškega spola iz različnih krajev Slovenije. V 1. letnik SŠ so se vpisovali med leti 1988 in 1997. Ob vpisu so bili v povprečju stari 15 let oz. 180 mesecev. Učna neuspešnost je bila dokaj pogosta – 24,2% dijakov je šolanje v SŠ prekinilo. Največ (51,2%) dijakov je redno zaključilo šolanje z dobrim ali zadostnim uspehom (kategorija uspešnosti 4). Le-ti predstavljajo prevladujoči razred, ki določa apriorno klasifikacijsko točnost (točnost napovedovanja na podlagi najbolj pogoste kategorije), ki bi jo moral ustrezen napovedni model preseči. Statistični rezultati so pokazali, da so numerični atributi baze ustrezni za napovedovanje uspešnosti zaključka in da je dijake različnih smeri izobraževanja potrebno obravnavati

ločeno, saj so se kljub podobnosti nekatere statistike treh smeri pomembno razlikovale. Izostanki in disciplinski ukrepi so bili negativno, različne ocene OŠ in SŠ pa pozitivno povezane z uspešnostjo, kar je povsem razumljivo. Starost ob vpisu v SŠ je negativno korelirala z uspešnostjo, saj so starejši dijaki običajno ponavljalci. Uspešnost zaključka šolanja je najvišje korelirala z zaključnim uspehom posameznih letnikov SŠ, pri čemer so korelacije z višjimi letniki rasle, kar je prav tako razumljivo in smiselno. Manjša baza je vsebovala dijake mlajših generacij, ki so bile nekoliko manj vztrajne in so hitreje opustile šolanje (statistično pomembne razlike) kot starejše. Domnevamo, da je glavni razlog za to v spremembah splošnih socialnih vrednot skozi čas. Življenjska in šolska praksa kaže, da je trdo delo, kot vrednota starejših generacij, danes vse manj cenjeno in nagrajevano. Razen tega med celotno bazo in manjšo bazo ni bilo pomembnih razlik. Razkrile so se tudi nekatere nepričakovane zanimivosti – npr., da ocene predmeta pri določenem učitelju SŠ sploh nimajo nobene veljavnosti za napovedovanje uspešnosti zaključka šolanja, kar zasluži resen razmislek in ustrezno ukrepanje vodstva šole. Gre za enega od osnovnih predmetov, pri katerem se znanje nadgrajuje vsa štiri leta izobraževanja, dokaj pomembno pa je tudi na večini študijskih smeri. Ocene vseh predmetov v vseh letnikih SŠ so pomembno korelirale z uspešnostjo. Tudi ocene tega predmeta v OŠ, 1. in 2. letniku so pomembno korelirale z uspešnostjo (prek 0,30 oz. 0,40, na nivoju tveganja 0,01), v 3. in 4. letniku (kjer je predmet poučeval drug učitelj) pa je bilo opaziti popolno odsotnost korelacij z uspešnostjo ($r = 0$). Ocene omenjenega učitelja prav tako niso korelirale z nobenimi drugimi ocenami, niti istega predmeta v predhodnih letnikih. Rezultati so torej pokazali, da omenjeni učitelj ni uporabljal nobenega sistematičnega in doslednega kriterija ocenjevanja, kar so ob nadaljnjih poizvedbah podprle tudi izjave bivših dijakov o ocenjevanju po sistemu »kakor je kdo učitelju všeč«.

Rezultati gradnje odločitvenih dreves

Najvišjo klasifikacijsko točnost so dosegla drevesa, zgrajena iz vseh atributov (nad 99%), ter drevesa iz strojno izbranih atributov (nad 95%). Ob koncu SŠ, ko poznamo vse attribute, je napoved zaključne uspešnosti skoraj popolna, toda praktično neuporabna. Če napoved omejimo na attribute, ki so znani vsaj do konca 1. letnika SŠ (ekspertna in pragmatično-strojna izbira atributov), klasifikacijska točnost odločitvenih dreves znatno upade. Slika 1 kaže, da je klasifikacijska točnost dreves na podlagi 16 podatkov, znanih pred vpisom, le nekaj nad 50%, na podlagi 30 podatkov, znanih do konca 1. letnika, okrog 60% in na podlagi 46 podatkov, znanih do konca 2. letnika, malo pod 70%.

Odločili smo se za uporabo dreves na podlagi ekspertne izbire atributov, njihovo razmeroma nizko klasifikacijsko točnost pa smo poskušali zvišati na različne načine (gradnja dreves po posameznih smereh izobraževanja, drugačne klasifikacije uspešnosti, uporaba »na stroške občutljive klasifikacije« in gradnja dreves na manjši bazi). Poskusi zvišanja klasifikacijske točnosti so pokazali, da je (glede na zaželenost čimprejšnje

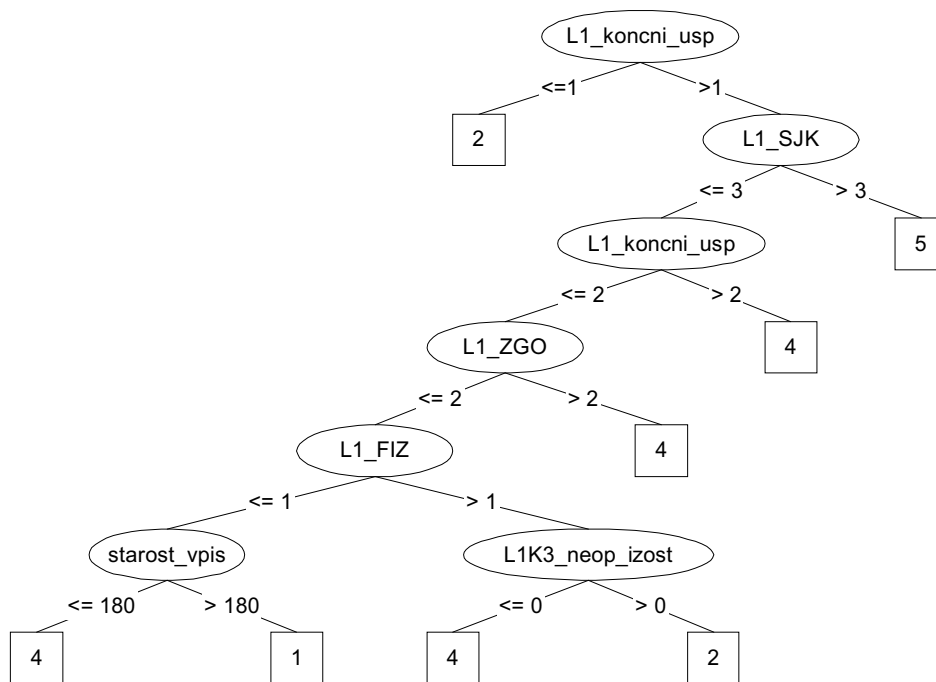


Slika 1: Klasifikacijska točnost dreves iz različnega števila atributov.

napovedi in čim višjo klasifikacijsko točnost) najbolj smiselna gradnja dreves iz ekspertno izbranih atributov, znanih do konca 1. letnika, na manjši bazi glede na smer šolanja. Le-ta je dala manjša, lažje razumljiva drevesa z bistveno višjo klasifikacijsko točnostjo – pogosto za več kot 10% višjo. Klasifikacijska točnost teh »izboljšanih« dreves iz atributov, znanih pred vpisom, je okrog 60%, pri drevesih iz atributov, znanih do konca 1. letnika SŠ, pa okrog 70%. Analiza točnosti napovedovanja posameznih kategorij uspešnosti je pokazala, da drevesa nekatere kategorije (2 in 5) napovedujejo zelo točno, druge (zlasti 3) pa bistveno slabše. Deleži pravilno uvrščenih pozitivnih primerov (true positive rate) in napačno uvrščenih pozitivnih primerov (false positive rate) so pokazali, da napovedi uspešnosti (zlasti kategoriji 3 in 4) še ne izključujejo neuspešnosti, napovedi neuspešnosti (kategoriji 1 in 2) pa dopuščajo le malo možnosti za drugačen izhod.

Kljub razlikam med drevesi treh smeri se je pokazalo, da imata med atributi, znanimi do konca 1. letnika SŠ, največjo napovedno moč zaključni uspeh OŠ oziroma končni uspeh 1. letnika SŠ (vedno sta se pojavljala kot delilca pri korenu drevesa), kar je pričakovano in razumljivo. Najboljše drevo za smer »L« (slika 2) npr. kaže, da bo dijak, ki je imel v 1. letniku pozitiven uspeh in oceno slovenščine višjo od 3, najverjetneje redno zaključil šolanje s prav dobrim ali odličnim uspehom (kategorija uspešnosti 5). Podobno lahko napovemo tudi uspešnost ostalih kandidatov. Po mnenju eksperta atributi v drevesu nakazujejo, da smer zahteva ustrezno splošno inteligentnost, verbalne, spominske in logične sposobnosti, izostanki in višja starost ob vpisu (ponavadi zaradi ponavljanja 1. letnika SŠ na isti ali drugi šoli oz. ponovnega vpisa) pa zmanjšajo možnost uspeha.

Na splošno so odločitvena drevesa razkrila naslednje. Mlajše generacije so nekoliko manj uspešne. Prediktivna moč različnih šolskih ocen se razlikuje glede na OŠ (pri drevesih iz vseh atributov se je bliže korenu drevesa pojavljalo ime OŠ oz.



Slika 2: Najboljše drevo iz 30 atributov za smer »L«: 8 listov (kvadrati) in 7 vozlišč (elipse). Klasifikacijska točnost je 69,66% (Vozlišča predstavljajo naslednje: L1_koncni_usp = končni uspeh 1. letnika SŠ, L1_SJK = ocena slovenščine v 1. letniku, L1_ZGO = ocena zgodovine v 1. letniku, L1_FIZ = ocena fizike v 1. letniku, starost_vpis = starost dijaka ob vpisu v mesecih in L1K3_neop_izost = neopravičeni izostanki v 3. konferenci 1. letnika. Številke v listih predstavljajo kategorije uspešnosti zaključka šolanja).

občina bivanja, ki so mu sledila vozlišča različnih ocen OŠ). Iste ocene ne predstavljajo enakega znanja, ker se OŠ razlikujejo v kriterijih ocenjevanja (Toličič in Zorman, 1977). Pokazalo se je, da imajo isti atributi, predvsem ocene, višjo prediktivno moč za zahtevnejše in bolj selektivne smeri izobraževanja (podobno kot prej, se je atribut smer SŠ pojavljal bližje korenu kot različne ocene, drevesa posameznih smeri pa so imela kljub istim atributom različno klasifikacijsko točnost – za zahtevnejše smeri višjo). Najmanj možnosti za uspešen zaključek izobraževanja na obravnavni šoli imajo dijaki z nezadostnim uspehom 1. letnika SŠ, še posebej če je le-ta posledica šibkega znanja matematike ali/in fizike (notranja vozlišča drevesa: nezadostni uspeh 1. letnika, nezadostni ali neocenjeni matematika in fizika vodijo do razreda uspešnosti 1 ali 2). Razkrile so se tudi spremembe v času, npr. spremenjen kriterij ocenjevanja kot posledica zamenjave učitelja. Vozlišča v drevesu so razkrila, da so do določenega leta ocene določenega predmeta imele relativno visoko napovedno moč za uspešnost, kasneje

pa so jo v veliki meri izgubile. Naknadne poizvedbe v šoli so pokazale, da je v tem času prišlo do zamenjave učitelja predmeta, očitno za manj zahtevnega oz. strogega.

Rezultati razvrščanja v skupine

Razvrščanje v skupine po trajanju šolanja in končnem uspehu 4. letnika (s katerima je opredeljena uspešnost zaključka šolanja), je pokazalo, da je dijake najbolj smiselno razvrstiti na 5 skupin in tako podprlo izbrano opredelitev kategorij uspešnosti zaključka šolanja. Po mnenju eksperta (glede na enotnost in vsebinsko razlago skupin) je bila najbolj smiselna razvrstitev dijakov na 5 skupin na podlagi atributov, ki jih poznamo do konca 1. letnika SŠ, saj se nekatere razlike v sposobnostih in motivaciji dijakov (ki jih domnevno odražajo ti atributi) pokažejo šele na zahtevnejšem nivoju izobraževanja. Skupine so se razlikovale predvsem v ocenah predmetov in splošnega učnega uspeha. Slabši učenci v OŠ so se pogosteje vpisovali na manj zahtevno smer, kar ob neomejenem vpisu kaže, da so vsaj delno upoštevali nasvete svetovalcev ali pa so zaradi slabšega uspeha na OŠ že imeli znižane aspiracije. Učenci iz velikih mest imajo višje izobrazbene aspiracije, saj so se ob enako slabemu uspehu OŠ pogosteje vpisali na zahtevnejšo smer, kar je rezultiralo v najslabšem učnem uspehu 1. letnika. Odvisno od smeri, je bil uspeh 1. letnika SŠ za oceno do dve nižji od zaključnega uspeha OŠ. Opravičeni in neopravičeni izostanki so bili močno povezani z uspehom 1. letnika, zaradi odnosa interakcije pa je težko reči, kaj je vzrok in kaj posledica. Najnižji končni uspeh so dosegli dijaki, ki so bili v OŠ najslabši, vendar so izbrali najzahtevnejši program in največ izostajali. Na splošno je bila uspešnost dijakov 1. letnika SŠ nizka, njihove ocene in splošni uspeh so se gibale večinoma od nezadostnih do dobrih, prav dobri in odlični pa so bili zelo redki.

Odločitveni modeli, zgrajeni s HINT

Klasifikacijska točnost modelov HINT je bila nizka, nezadovoljiva (običajno blizu apriorne – 51,2%), kar je kazalo, da bo na podlagi atributov baze težko zgraditi model z visoko napovedno točnostjo. Kljub nekoliko slabši razumljivosti modelov so se pokazale nekatere zakonitosti. Najboljši prediktorji uspešnosti zaključka šolanja so bili: zaključni uspeh OŠ, oceni matematike in fizike v 8. razredu OŠ, tuj jezik v OŠ in splošni učni uspeh 1. letnika SŠ. Po »načelu najšibkejšega člena verige« sta predvsem oceni matematike in fizike omejevali najvišji uspeh 1. letnika (oceni matematike in fizike sta tvorili koncept, od katerega je odvisen uspeh - poimenovali smo ga sposobnost logičnega sklepanja). Izostanki so se povezovali v nov koncept skupaj z ocenama matematike in fizike, kar pomeni, da dijaki niso izostajali naključno ali zaradi zdravstvenih težav, pač pa sistematično, v povezavi s tema predmetoma – verjetno zaradi preverjanj znanj. Obstajalo je več smiselnih struktur podatkov, mnoge interakcije med atributi pa so močno otežile iskanje najboljših. Zaključili smo, da tako razviti modeli niso primerni za neposredno uporabo, vendar so vseeno ponudili zanimive smernice za nadaljnjo »ročno«

izdelavo modelov DEX.

Struktura modelov DEX

Zaradi specifične uporabe (rezultati so neposredno uporabni samo za obravnavano šolo) in obširnosti prikazujemo samo del rezultatov primarnega modela (za napovedovanje ob vpisu) za smer izobraževanja »L«. Strukturo (drevo kriterijev z zalogami vrednosti) modela prikazuje slika 3, tabelo odločitvenih pravil pa slika 4. Na levi strani slike 3 je prikazano drevo kriterijev, iz katerega je razvidno, kateri podkriteriji sestavljajo nadrejeni kriterij (npr. uspešnost vključuje tri podkriterije: sposobnosti in predznanje, motivacijo, okoliščine). Glede na vsebinsko povezanost in medsebojno odvisnost kriterijev strukturo drevesa določi ekspert. Na desni strani slike 3 so zaloge vrednosti, ki jih lahko zavzame posamezen kriterij (npr. uspešnost lahko zavzame 5 numeričnih vrednosti-kategorij, katerih pomen je opisan pod drevesom); le-te določi

Zaloge vrednosti

Kriterij	Zaloga vrednosti
uspešnost	1; 2; 3; 4; 5
sposobnosti in predznanje	neustrezne ; ustrezne; zelo ustrezne
splošna IQ	podpovprečna ; povprečna; nadpovprečna
uspeh OŠ	2; 3; 4; 5
uspeh 7.r	2; 3; 4; 5
verbalne sposobnosti	podpovprečne ; povprečne; nadpovprečne
SJK 8.r	2; 3; 4; 5
tuj jezik 8.r	2; 3; 4; 5
logično rezoniranje	podpovprečno ; povprečno; nadpovprečno
MAT 8.r	2; 3; 4; 5
FIZ 8.r	2; 3; 4; 5
znanje tujih jezikov	neustrezno ; delno ustrezno; ustrezno
tuj jezik OŠ	ni imel ; madžarski; nemški; angleški
tuj jezik	nemški ; angleški; angleški in nemški
znanje slovenščine	pomanjkljivo ; povprečno; zelo dobro
država roj	Nemčija ; Avstrija; Srbija; BiH; Hrvaška; Slovenija
državljanstvo	BIH ; slovensko
motivacija	znižana ; ustrezna
starost vpis	višja ; normalna
redni vpis	ne ; da
okoliščine	negativne ; ustrezne; pozitivne
vozač	da ; ne
status šp	da ; ne
spol	z ; m

uspešnost

uspešnost zaključka šolanja po določenem srednješolskem programu

- 1** prekine po 1-2 letih
- 2** v 3 ali več letih prekine ali ne zaključi s pozitivnim uspehom
- 3** v 5-6 letih s pozitivnim uspehom
- 4** redno z dobrim ali zadostnim uspehom
- 5** redno z odličnim ali pravdobrim uspehom

Slika 3: Struktura in zaloge vrednosti primarnega večparametrskega modela.

sposobnosti in predznanje	motivacija	okolščine	uspešnost
48%	41%	11%	
1 neustrezne	znižana	*	1
2 neustrezne	*	<= ustrezne	1
3 neustrezne	<i>ustrezna</i>	<i>pozitivne</i>	2
4 >= ustrezne	znižana	negativne	2
5 ustrezne	znižana	<i>pozitivne</i>	2
6 >= ustrezne	znižana	ustrezne	3
7 ustrezne	<i>ustrezna</i>	negativne	3
8 ustrezne	ustrezna	>= ustrezne	4
9 zelo ustrezne	znižana	<i>pozitivne</i>	4
10 zelo ustrezne	<i>ustrezna</i>	*	5

Slika 4: Prikaz ene tabele odločitvenih pravil (funkcij koristnosti) primarnega modela (Odstotki v glavi kažejo delež oz. obtežitve podkriterijev v nadrejenem kriteriju uspešnosti).

ekspert. Slika 4 prikazuje eno od tabel odločitvenih pravil tipa »če-potem«, ki definirajo, kako se podrejeni kriteriji sestavljajo v nadrejenega (npr. če so sposobnosti in predznanje neustrezne in motivacija znižana, potem bo uspešnost 1).

Validacija modelov

Primerjava odločitev svetovalca, modelov DEX in najboljših odločitvenih dreves je pokazala praktično enako splošno klasifikacijsko točnost in točnosti napovedovanja po posameznih razredih. Točnost primarnih ocen (oz. napovedi na podlagi atributov, znanih do vpisa) vseh treh je okrog 60%, točnost sekundarnih ocen (oz. napovedi na podlagi atributov znanih do konca 1. letnika SŠ) pa okrog 70%. Natančni podatki so navedeni v tabeli 1.

Podrobnejša analiza klasifikacij dijakov po posameznih kategorijah uspešnosti je pokazala, da uspešnim dijakom skoraj nikoli niso napovedali neuspešnosti, dijaki z nizko oceno pa so bili ponavadi res neuspešni. Vsi so se občasno motili tako, da so neuspešnim dijakom napovedali uspešnost (DEX v 12,8% in odločitvena drevesa v 8,5% primerov). Razmeroma dobro so prepoznavali izrazito neuspešne in izrazito

Tabela 1: Splošna klasifikacijska točnost DEX modelov in svetovalca.

	DEX model		Ocena svetovalca	
	Primarni	Sekundarni	Primarna	Sekundarna
Klasifikacijska točnost (%)	59,57	72,34	59,57	63,83

uspešne dijake, slabo pa atipične primere določene kategorije uspešnosti. Napačne napovedi, za več kot eno kategorijo višje ali nižje od dejanske uspešnosti, so bile razmeroma redke (DEX v 12,8% in odločitvena drevesa v 6% primerov). Vsi so najboljše prepoznavali dijake iz kategorije uspešnosti 4, najslabše pa dijake iz kategorije uspešnosti 3. Pogosto so zamenjali primere kategorij 1 in 3, kar je razumljivo, saj obe kategoriji vključujeta dijake, ki so se dolgo izobraževali in ob slabem uspehu postopno napredovali, vendar pa skoraj vse do konca ni bilo mogoče reči, kateri bodo uspešno zaključili izobraževanje (kategorija 3) in kateri ne (kategorija 1).

Zaključki in priporočila za nadaljnje raziskovanje

Zakovitosti, odkrite z rudarjenjem podatkov, večinoma niso nove za izkušene pedagoške delavce, imajo pa velik pomen za prepričevanje dijakov in staršev, ki bolj kot subjektivnemu mnenju svetovalca zaupajo objektivnim rezultatom raziskav. V naši raziskavi se je pokazalo tudi nekaj novih in presenetljivih odkritij, ki vodstvu šole lahko služijo za bolj kvalitetno odločanje v zvezi z zaposlovanjem in vodenjem kadrov.

Članek predlaga tri različne postopke za izgradnjo napovednih modelov, ki lahko služijo kot vodilo nadaljnjim raziskavam. Prvi temelji na metodah rudarjenja podatkov in strojnega učenja, drugi (ki v raziskavi ni bil samostojno uporabljen) pa na »ročni« izgradnji modelov s pomočjo lupine ekspertnih sistemov, kakršna je npr. DEX. Tretji je kombinacija obojega in vključuje »ročno« izgradnjo modelov, v katere je vključeno tako ekspertno znanje kot tudi znanje, pridobljeno s postopki rudarjenja.

Uspeli smo izdelati dve vrsti modelov za napovedovanje uspešnosti zaključka šolanja - odločitvena drevesa in modele DEX – ki svetovalcu lahko služijo kot pripomoček za poskusno uporabo. Pred resno uporabo bi jih morali validirati na novih podatkih dijakov ustrezne šole in smeri. Pri oblikovanju nasvetov naj bi se svetovalec opiral predvsem na lastno znanje in izkušnje, napovedi modelov pa kritično pretehtal in upošteval le ob vseh ostalih podatkih. Pazljivo naj bi preučil zlasti primere razhajanj napovedi. Ob koncu 1. letnika SŠ naj bi (sam in z modeli) dijakom izdelal sekundarno napoved, ki poveča zanesljivost, ter po potrebi korigiral primarno napoved po koncu OŠ. Dober svetovalec naj se ne bi ustavil pri tem, ampak bi na podlagi napovedi skušal spremeniti tudi kritične attribute, če je le mogoče (npr. zvišati dijakovo motivacijo, izboljšati predznanja, izboljšati domače okoliščine). Najbolje bi bilo, če bi dijaka ves čas spremljal isti svetovalec, saj s časom svetovalec dijaka vse bolje pozna in mu lahko tudi bolj kvalitetno svetuje. Svetovalec naj bi vodil tudi evidenco svojih napovedi in napovedi modelov za kasnejšo evalvacijo in validacijo modelov. Za validacijo modelov bi bili seveda primerni le dijaki, pri katerih svetovalec ne bi načrtno spreminjal atributov.

Obe vrsti modelov sta dosegli praktično enako klasifikacijsko točnost. Glede na majhno število in nizko kvaliteto uporabljenih atributov, ki niso »čisti« in natančni pokazatelji dijakovih lastnosti, je bila točnost modelov presenetljivo visoka in zelo

zadovoljiva – dosegli so točnost izkušenega človeškega eksperta ob enakih podatkih. Naš svetovalec je bil izkušen (10 let delovnih izkušenj v izobraževalnem svetovanju), vendar pa točnosti njegovih napovedi ne pove nič o povprečni napovedni točnosti svetovalcev. Čeprav se svetovalci opirajo tudi na druge podatke (pridobljene ob srečanjih z dijakom in starši), ki običajno povečajo točnost napovedi, pa se med seboj precej razlikujejo v izkušnjah, sposobnostih, znanju in intuiciji. Napovedna točnost modelov v praksi je verjetno manjša kot točnost izkušenega svetovalca, toda večja kot točnost svetovalca-novinca. Menimo, da bi predlagana uporaba modelov prispevala k bolj točnemu napovedovanju uspešnosti, k večjemu upoštevanju nasvetov in s tem k izboljšanju splošne učne uspešnosti. Smiselna je zato, ker je zelo izkušenih in sposobnih svetovalcev malo, pomeni pa tudi kakovosten premik v smeri bolj doslednega, objektivnega in sistematičnega ocenjevanja, evidentiranja in evalvacije napovedi, kar je v sedanjosti najbolj šibka točka izobraževalnega svetovanja (Resman in dr., 1999).

Predlagana uporaba modelov ima določene dobre in slabe plati. Prednosti obeh vrst modelov so: (1) zahtevata le tiste podatke, s katerimi vedno razpolagamo, (2) večja doslednost, objektivnost in sistematičnost napovedovanja, (3) lahka uporaba, (4) ponujata »drugo mnenje«, ki zmanjša možnost napak, (5) sekundarna ocena poveča zanesljivost napovedi, (6) razumljivost, objektivnost in argumentiranost napovedi pa pripomore k večjemu zaupanju in upoštevanju nasvetov. Slabost obeh vrst modelov je, da se občasno motijo in niso validirani na novih podatkih, zaradi česar je točnost njunih napovedi v praksi neznana, verjetno pa je vsaj nekoliko nižja od ugotovljene. Praktična uporaba modelov skriva naslednje pasti. Ni mogoče predvideti, kako bo napoved vplivala na motivacijo dijaka, s tem pa tudi na končno uspešnost. Zato je za njeno sporočanje nujno potreben strokovnjak. Poudariti je potrebno, da napoved drži le pod pogoji, v katerih je bila ugotovljena. Ker se motivacija, interesi in ostale osebne lastnosti v puberteti močno spreminjajo, je napoved vedno tvegana. Naš šolski sistem se pogosto spreminja, spreminjajo pa se tudi splošne značilnosti generacij dijakov, kar lahko bistveno poslabša napovedno točnost modelov. Te in kakršne koli druge modele bi torej morali neprestano validirati in prilagajati spremembam. Seveda se kljub opozorilom lahko zgodi, da bi se neizkušeni svetovalci preveč opirali na napovedi modelov in premalo na lastno presojo.

Izboljšanje napovedne točnosti modelov DEX bi lahko dosegli s spremembami modelov in vključevanjem novih kriterijev, predvsem čimbolj »čistih« in točnih mer različnih sposobnosti, interesov, motivacije in osebne lastnosti dijaka. Ker s temi podatki svetovalci običajno ne razpolagajo, bi to zahtevalo obsežna psihološka testiranja posameznih dijakov. Idealno bi bilo razviti sistem ekspertnih sistemov (modelov), ki bi učencu ob zaključku OŠ predlagal nekaj, zanj najbolj ustreznih, smeri izobraževanja. Razvoj takega sistema bi zahteval natančen »profil« lastnosti dijaka, ki ustreza posamezni smeri izobraževanja, njegova validacija pa bi zahtevala obsežna psihološka testiranja in longitudinalno spremljanje celih generacij dijakov, kar ob sedanjih razmerah šolskega svetovanja ni realno pričakovati.

Zaradi zahtevnosti, dolgotrajnosti in obsežnosti dela pri »ročnem« razvoju in validaciji modelov DEX je prednost na strani odločitvenih dreves, ki so morda tudi bolj primerna glede na naravo uporabljenih podatkov (številne interakcije med atributi). Za njihovo strojno izgradnjo in validacijo potrebujemo samo urejeno bazo podatkov in poznavanje ustreznih računalniških programov. Že izgrajena drevesa pa se lahko uporabijo za predikcijo novih primerov tudi »na papirju«, brez računalnika (Gams in Lavrač, 1987). Po drugi strani je treba priznati, da sta priprava velikih baz podatkov in samo rudarjenje podatkov precej zahtevna. Za povprečne uporabnike osebnih računalnikov je lahko »ročna« izgradnja modelov DEX veliko lažja. Poudarjamo, da je primerjava točnosti človeškega eksperta in modelov služila izključno validaciji, sicer pa ne gre za izbiro »ali-ali«, temveč za izkoriščanje največje prednosti, ki jo ponuja uporaba modelov v rokah izkušenega šolskega svetovalca.

Literatura

- Bohanec, M. (2002). Metode umetne inteligence [Artificial intelligence methods]. WWW site. <http://www-ai.ijs.si/MarkoBohanec/ai/ai.html> [Zadnji dostop 17. maja, 2002].
- Bohanec, M. in Rajkovič, V. (1995). Večparametrski odločitveni modeli [Multi-attribute decision models]. *Organizacija*, 28 (7), 427-438.
- Bohanec, M. in Rajkovič, V. (1999). Multi-attribute decision modeling: industrial applications of DEX. *Informatica*, 23 (4), 487-491.
- Bohanec, M., Zupan, B. in Rajkovič, V. (2000). Applications of qualitative multi-attribute decision models in health care. *International Journal of Medical Informatics*, 58/59, 191-205.
- Bohanec, M., Zupan, B., Bratko, I. in Cestnik, B. (1997). 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, ISDSS97* (503-514). Luusanne: Universite de Lausanne.
- Brančič, B. (1986). *Psihološke teorije izbora zanimanja [Psychological theories of profession choice]*. Beograd: Naučna knjiga.
- Clark, P. in Niblett, T. (1987). Induction in Noisy Domains. V I. Bratko in N. Lavrač (ur.), *Progress in machine learning: proceedings of 2nd European worknig session on learning, EWSL 87*. Wilmslow: Sigma Press.
- Dutta, S. (1993). *Knowledge processing and applied artificial intelligence*. Oxford: Butterworth-Heinemann.
- Ferligoj, A. (1989). *Razvrščanje v skupine: teorija in uporaba v družboslovju [Clustering: theory and application in social sciences]*. Ljubljana: Raziskovalni inštitut FSPN.
- Galle, R. (1996). *Priročnik za uporabo programa: vodenje evidence učencev (Evidenca 3) [Manual for using program: pupils' evidence]*. Ljubljana: SŠER.
- Gams, M. in Lavrač, N. (1987). Review of five empirical learning systems within a proposed schemata. V I. Bratko in N. Lavrač (ur.), *Progress in machine learning: proceedings of 2nd European worknig session on learning, EWSL 87* (201-217). Wilmslow: Sigma Press.

- Han, J. in Kamber, M. (2001). *Data mining: concepts and techniques*. San Francisco: Morgan Kaufman Publishers.
- Hand, D.J. (1998). Data mining: statistics and more. *American Statistician*, 52 (2), 112-118.
- Kononenko, I. (1997). *Strojno učenje [Machine learning]*. Ljubljana: založba FRI.
- Lapajne, Z. (1984). *Proces izbire poklica [Process of profession choice]*. Ljubljana: Univerza Edvarda Kardelja.
- Maindonald, J. (1999). Data mining from a statistical perspective. WWW site. <http://www.maths.anu.edu.au/~johnm/dm/dmpaper.html> [Zadnji dostop 17. maja, 2002].
- Mallach, E.G. (1994). *Understanding decision support systems and expert systems*. Burr Ridge: Irwin.
- Michalski, R.S. in Kaufman, K.A. (1998). Data mining and knowledge discovery: a review of issues and a multistrategy approach. V R.S. Michalski, I. Bratko in M. Kubat (ur.), *Machine learning and data mining: methods and applications*. Chichester: John Wiley & Sons.
- Mitchell, T.M. (1997). *Machine learning*. New York: The McGraw-Hill.
- Musek, J. in Pečjak, V. (1995). *Psihologija [Psychology]*. Ljubljana: Educy.
- Oman, I. (1976). *Teorija in praksa poklicnega usmerjanja [Theory and practice of professional guidance]*. Kranj: Moderna organizacija.
- Quinlan, R.J. (1993). *C4.5: programs for machine learning*. San Francisco: Morgan Kaufman.
- Rajkovič, V. in Bohanec, M. (1988). Sistemi za pomoč pri odločanju [Decision support systems]. *Organizacija in kadri*, 21 (1/2), 127-140.
- Resman, M., Bečaj, J., Bezić, T., Ččinovič-Vogrničič, G. in Musek, J. (1999). *Svetovalno delo vrtcih, osnovnih in srednjih šolah [Counseling in kindergartens, primary and secondary schools]*. Ljubljana: ZRSŠ.
- Rodić, R. (1980). *Izbor zanimanja [The choice of profession]*. Novi sad: Misao.
- SPSS Inc. (2002). SPSS. WWW site. <http://www.spss.com/spssbi/spss/> [Zadnji dostop 17. maja, 2002].
- Toličič, I. in Zorman, L. (1977). *Okolje in uspešnost učencev [Environment and pupils' academic achievements]*. Ljubljana: DZS.
- University of Waikato (2002). Weka 3 – data mining with open source machine learning software in Java. WWW site. <http://www.cs.waikato.ac.nz/ml/weka/> [zadnji dostop 17. maja, 2002].
- Vieira, R. (2000). *Professional SQL Server 2000 programming*. Birmingham: Wrox Press.
- Witten, I.H. in Frank, E. (2000). *Data mining: practical machine learning tools and techniques with Java implementations*. San Francisco: Morgan Kaufmann.
- Zupan, B. (1997). Machine learning based on function decomposition. Doktorska disertacija, Univerza v Ljubljani. WWW site. <http://magix.fri.uni-lj.si/hint/papers.asp> [Zadnji dostop 17. maja, 2002].
- Zupan, B. in Demšar, J. (2002). Orange. WWW site. <http://magix.fri.uni-lj.si/orange/> [Zadnji dostop 17. maja, 2002].

Prispelo/Received: 26.06.2002

Sprejeto/Accepted: 02.09.2002