

# Odločitvena drevesa za individualno testiranje učencev

Irena Nančovska Šerbec, Alenka Praprotnik, Jože Rugelj

Pedagoška fakulteta, Univerza v Ljubljani, Ljubljana, Slovenija,  
irena.nancovska@guest.arnes.si, alenka.praprotnik@siol.net, joze.rugelj@uni-lj.si.

V članku je predstavljena možnost uporabe odločitvenih dreves za sestavljanje testov, ki se prilagajajo trenutnemu znanju posameznega učenca. V prvi fazi realiziramo spletno testiranje učencev s širšo množico vprašanj in nalog, ki se nanašajo na določen tematski sklop. Rezultate testov shranjujemo v skupno podatkovno bazo. Testiranje je anonimno in uporabniku, oziroma učencu omogoča večkratno vrednotenje testov. S pomočjo mehanizmov strojnega učenja v naslednji fazi zgradimo odločitvena drevesa in druge modele, ki predstavljajo klasifikacijo učencev v razrede uspešnosti. Modeli hkrati predstavljajo hierarhijo absorpcije znanja s strani množice učencev, ki smo jih testirali, oziroma povprečno obvladovanje posameznih ciljev, ki jih določena naloga v hierarhiji drevesa preverja. V zadnji fazi na podlagi odločitvenih dreves realiziramo hitro testiranje učenca, za katero uporabljamo odločitveno drevo kot klasifikator. To pomeni, da se test prilagodi trenutnemu znanju, ki ga učenec kaže v predhodnih fazah testiranja. Možnost uporabe metod strojnega učenja smo potrdili na domeni, ki predstavlja množico ocenjenih testov iz osnovnošolske matematike..

**Ključne besede:** Klasifikacija, odkrivanje znanja iz podatkov, odločitvena drevesa, izbiranje atributov

**DECISION TREES FOR INDIVIDUAL STUDENT TESTING:** In the paper we present a possible use of decision trees for testing, which is adapted to the knowledge of each student. In the first phase we realize web testing of the students with wider set of exercises, regarding to specific topics. Results of testing are saved in a common data base. Testing is anonymous and it enables students multiple test evaluations. In the next phase, by using machine learning recipes, we built decision trees and other models which present students' classification into classes of successfulness. At the same time, models present hierarchy of knowledge absorption by the tested students, or average knowledge of particular goal, captured in each separate exercise. Finally, we realize fast web testing and classification, based on the decision tree. This means that the test is adapted to the knowledge, which is showed through testing. We demonstrate the possibility of use of machine learning on the domain of validated tests of primary school mathematics.

**Key Words:** Classification, Knowledge Discovery from Data, Decision Trees, Attribute Selection

## 1 Uvod

Izobraževani proces je v sedanjem času prepleten in podprt z možnostmi, ki jih ponuja Svetovni splet. Pri tem ni samo obogaten z dosežki izobraževalno-komunikacijske tehnologije, pač pa je tudi pestrejšje organiziran. Ponuja fleksibilnost glede časa, prostora, vsebine in tempa izobraževanja. Učni proces je organiziran v smislu pridobivanja in ustvarjanja novega znanja, v katerem učitelj ni več edini prenašalec znanja in posredovalec določenih dejstev. Učenci znanje pridobivajo iz različnih virov, kar jim omogoča sodobna informacijsko-komunikacijska tehnologija in napredek v razvoju telekomunikacij (Praprotnik 2004).

Učno okolje, obogateno s spletom, omogoča ne samo učenje, ampak tudi ocenjevanje naučenega po meri posameznega učenca. Za učitelja je izbira testov, ki bi bili hkrati reprezentativni in bi zahtevali malo časa za reševanje, pomembna naloga v njegovem strokovnem delu. V pričujočem članku si prizadevamo, da bi z uporabo metod strojnega učenja olajšali delo učiteljem pri ocenjevanju znanja učencev. Radi bi izdelali sistem za samodejno generiranje testov, ki se bo prilagajal trenutnemu znanju učenca oziroma posamezniku.

Pri uspešnih učencih bi testiranje lahko izvedli le na pomembnejših nalogah, saj so osnovne naloge zanje praviloma nezanemljive in dolgočasne. Dobrim učencem, ki se le malo razlikujejo od odličnih, bi predlagali ponavljanje le tistih tem, ki jih pri strojnem ocenjevanju »ločijo« od odličnih. Slabšim učencem, pa bi s pomočjo zgrajenih modelov predlagali le ponovitev pomembnejših tem, saj so zaradi slabega poznavanja osnov večkrat slabše ocenjeni.

S pomočjo prijemov umetne inteligence, kot so odkrivanje znanja iz podatkov (ang. *Knowledge Discovery from Data* - *KDD*), ki je netrivialen proces odkrivanja implicitnega, doslej neznanega in potencialno uporabnega znanja iz podatkov (Bohanec 2004), skušamo odkriti zakonitosti, ki omogočajo hitro in individualno prilagojeno testiranje. S podatkovnim rudarjenjem (ang. *Data Mining-DM*), ki je faza odkrivanja znanja iz podatkov v kateri dejansko pride do »odkrivanja znanja«, skušamo analizirati podatkovno bazo, ki je sestavljena iz testov učencev, katerim je bil predstavljen določen tematski sklop s strani učitelja. Podatkovna baza skriva zakonitosti učnega uspeha učencev pri absorpciji znanja, ki se nanaša na omenjen učni sklop. Z metodami podatkovnega rudarjenja bi radi zgradili sistem za

samodejno generiranje testov, ki iz širše množice nalog izbere tiste, ki so primerne za validacijo posameznega učenca. Pri tem si lahko pomagamo z modeli *odločitvenih dreves*, ki se pogosto uporabljajo za reševanje problemov klasifikacije in napovedovanja (Kononenko 1997; Michalski in Kaufman 1998; Gasar, Bohanec in Rajkovič 2002). Poleg možnosti razvrščanja v skupine uspešnosti bi *odločitvena drevesa* pokazala, katere teme najbolj vplivajo na uspeh učencev. V omenjenem sistemu bi lahko uporabili tudi dr. metode za razvrščanje v skupine ter tako učencem, ki kažejo podoben vzorec uspešnosti, predlagali teme, ki se zdijo zanimive njihovim sošolcem s podobno uspešnostjo. Navsezadnje lahko že v začetni fazi rudarjenja z neposrednim »merjenjem« in izbiranjem kakovostnih nalog, ki so v kontekstu strojnega učenja v našem primeru predstavljene z atributi, sestavimo hitre in učinkovite teste, ki so morda bolj primerni za ustno ocenjevanje kot za spletno testiranje. Pri klasifikacijskih problemih se uporabljajo različne heuristične ocene, ki merijo kakovost atributov (Robnik-Šikonja 2001; Robnik-Šikonja in Kononenko 2003). Z neposrednim izvajanjem različnih algoritmov lahko že a priori izberemo ožjo množico kakovostnih nalog, s katerimi hitro in učinkovito ocenimo učenca, še preden ga uvrstimo v skupino ali ocenimo z natančnejšim klasifikacijskim algoritmom.

Sistem za generiranje testov, ki so prilagojeni učencu, lahko uporabimo v osnovnih šolah, saj modeli za uvrščanje posnemajo ocenjevanje v kontekstu nivojskega pouka. V članku smo predstavili možnost uporabe sistema na področju osnovnošolske matematike za tematski sklop Izrazi, ki tematsko sodi v učni načrt za deveti razred (Praprotnik 2004). Sistem bi lahko uporabili tako v srednjih kot v višjih in visokih šolah pri različnih predmetih, ki zahtevajo pisno ocenjevanje učencev. Posebno možnost uporabe sistema vidimo pri izvajanju ustnih izpitov, kjer pride do izraza potreba po hitrem, učinkovitem in individualno prilagojenem ocenjevanju.

Testi, ki se prilagajajo učenčevemu znanju, so spletni testi, ki učencu zastavljajo nalogo za nalogo. V ozadju teče algoritem za uvrščanje učenca s pomočjo *odločitvenega drevesa*. Najprej mu zastavi nalogo, ki je v korenu drevesa, potem pa se odvisno od pravilnosti njegovega odgovora, »odloči« za naslednjo nalogo in to ponavlja na vsakem koraku, dokler ne pride do klasifikacije oziroma ocenjevanja učenca.

Kljub prednostim sistemi za individualno ocenjevanje učencev ne morejo izpodriniti ali v celoti nadomestiti strokovnjakov, saj kljub njihovi izredni hitrosti, s katero lahko z matematično natančnostjo obdelujejo podatke, človeškega širokega znanja in spomina še zdaleč ne presegajo. Učni uspeh je tudi večkrat odvisen od interakcije fizičnih, fizioloških socialnih in psiholoških dejavnikov učenja, ki so težko merljivi, pa so vseeno vsaj ocenjeni pri ocenjevanju s strani strokovnjakov. Strokovnjaki, učitelji in pedagogi zaradi visoke stopnje paralelizma v možganih izredno hitro in zanesljivo rešijo probleme validacije, ki jih računalnik sploh ne bi znal rešiti (Kononenko 1997).

## 2 Metode

Za obdelavo podatkov testiranja smo uporabili metode za odkrivanje znanja iz podatkov. Cilj je pridobivanje novega znanja, ki ga uporabimo bodisi za boljše razumevanje problema, ki ga modeliramo, bodisi za boljše odločanje, klasifikacijo ali napovedovanje (Mitchell 1997; Michalski in Kaufman 1998).

Najpomembnejša faza odkrivanja znanja iz podatkov je podatkovno rudarjenje. Podatkovno rudarjenje (Gasar, Bohanec, Rajkovič, 2003) pokriva statistične metode, vizualizacijo podatkov, metode strojnega učenja (med katere sodi tudi izgradnja *odločitvenih dreves*, ki jo uporabljamo v članku), povezovalna pravila in razvrščanje v skupine.

Preden se lotimo izgradnje *odločitvenih dreves*, bomo opisali nekaj algoritmov, ki merijo kakovost nalog, ki so v kontekstu strojnega učenja predstavljene z atributi.

### 2.1 Izbiranje atributov

Učni problemi so opisani z množico lastnosti, atributi (v našem primeru so to točke, ki jih dosežejo učenci pri reševanju posamezne naloge v širšem testiranju). Postopki za izbiro podmnožice pomembnih atributov nam v idealnem primeru vrnejo majhno, toda potrebno množico za opis danega koncepta (Robnik-Šikonja 2001) (v našem primeru so to pomembne naloge, ki odražajo razumevanje tematskega sklopa).

Pri klasifikacijskih problemih se uporabljajo različni postopki in mere za ocenjevanje kakovosti atributov so informacijski prispevek, indeks Gini, mera razdalje, j-ocena, Relief algoritmi; MDL, uporabljata se tudi  $\chi^2$  in G statistike (Robnik-Šikonja 2001). Ocenjevanje kakovosti atributov se uporablja na različnih področjih strojnega učenja: pri izbiri podmnožice pomembnih atributov (kot v našem primeru), uteževanju atributov, konstruktivni indukciji, izgradnji odločitvenih in regresijskih dreves.

Naštete mere, razen algoritmov Relief, predpostavljajo, da so atributi pogojno neodvisni od razreda in so neuporabni v primerih, kjer morda obstaja močna pogojna odvisnost med atributi. Pri absorpciji znanja, ki se nanaša na določen tematski sklop, menimo, da je razumevanje kompleksnejših nalog pogojeno z razumevanjem lažjih, zato pričakujemo boljše ocenjevanje kakovosti atributov z algoritmom *ReliefF* (Robnik-Šikonja, Kononenko, 2003), ki predstavlja robustno razširitev algoritma Relief na večrazredne probleme.

Algoritmi *Relief* ocenjujejo kakovost atributov glede na to, kako dobro vrednosti atributov ločijo med primeri, ki so si podobni, oziroma štejejo za »dobro« lastnost atributa, če imajo bližnji sosedji iz istega razreda iste vrednosti atributov, ter za »slabo«, če imajo bližnji sosedji, ki pripadajo različnim razredom, iste vrednosti razreda. Vrnejo oceno kakovosti za vsak atribut posebej v obliki uteži. Nepomembni atributi dobijo uteži, ki so blizu 0.

V članku smo za ocenjevanje atributov preizkusili tudi uteži, ki jih izračuna *metoda podpornih vektorjev*, ki prihaja iz statistične teorije učenja. (Cristianini, Shaw-Taylor 2000; Witten, Frank 2000). Osnovna ideja te metode je nelinearna preslikava primerov v visoko-dimenzionalni prostor, kjer

poiščemo hiper-ravnino z maksimalno margino, ki razmeji preslikane primere v razrede. Učni algoritem, ki temelji na Vapnik-Chervonenkisovi teoriji, s kontroliranjem kapacitete modela onemogoči, da bi se ta prilagodil učni množici, medtem ko optimizacijski algoritmi omogočijo izračun parametrov, ki jih ta teorija zahteva. Uteži podpornih vektorjev, ki definirajo hiper-ravnino, smo vzeli kot mero za kakovost atributov.

## 2.2 Odločitvena drevesa

Odločitvena drevesa so sestavljena iz notranjih vozlišč, ki predstavljajo razvejišča pri klasifikaciji (v našem primeru so to naloge, vprašanja), vej, ki so podmnožice vrednosti vhodnih polj (pri nas je to število doseženih točk, pri reševanju konkretne naloge) ter listov, ki predstavljajo razrede (v našem primeru so to razredi uspešnosti).

Pri izgradnji odločitvenega drevesa je osnovna naloga, da iz predstavljenih podatkov, učnih primerov, zgradimo odločitveno drevo, ki ponazarja preslikavo med izbranimi vhodnimi polji, atributi, in izbranim izhodnim poljem, razredom. Postopek generiranja odločitvenega drevesa iz učne množice imenujemo indukcija drevesa. Pri indukciji začnemo s praznim drevesom in celotno množico učnih objektov. Na vsakem koraku s pomočjo hevristične mere izberemo atribut, ki na poti do trenutnega vozlišča še ni bil uporabljen. Če število

učnih objektov ni dovolj ali če podatki vsebujejo šum, običajno pride do prekomernega prilagajanja učnim objektom. Rezultat tega so velika odločitvena drevesa, ki vsebujejo precej nepomembnih vej. V ta namen obstajajo postopki za rezanje nepotrebnih vej.

Modeliranja s pomočjo odločitvenih dreves se lotimo tako, da pripravimo množico rešenih primerov. Množico razdelimo na učno, s pomočjo katere gradimo odločitveno drevo, in testno, s pomočjo katere preverjamo zanesljivost zgrajenega drevesa. Sodobni algoritmi za učenje si pomagajo z mehanizmom, ki ga imenujemo n-kratno prečno preverjanje, ki je eden od mehanizmov za preprečevanje prekomernega prilagajanja modela učni množici (Witten, Frank, 2000). To je standardni postopek, ko učno množico razdelimo na deset (n) delov, kjer so vrednosti razredov enako zastopane. Model zgradimo na devetih (n-1) delih, testiramo, ga pa na desetem (n-tem). Postopek ponovimo desetkrat (n-krat). Napaka predstavlja povprečje desetih (n-tih) napak na testni množici. Končni model pa zgradimo na celotni učni množici.

Odločitveno drevo ne opisuje samo učnih podatkov, ampak omogoča tudi klasifikacijo novih primerov. Ko iščemo rešitev za nek nov, še nerešen primer, začnemo v korenu drevesa in se v vsakem atributnem vozlišču odločimo za vejo glede na vrednost pripadajočega atributa v nerešenem primeru, vse dokler ne pridemo do lista in s tem do predlagane rešitve.

**Tabela 1:** Razporeditev točk na testu - vir slike (Praprotnik 2004)

1) naloga (2 primera)	izračun in poenostavitev...3 t poimenovanje...2 t	10 t
2) naloga (3 primeri)	poenostavitev...3 t izračun z vrednostmi spremenljivk...3 t (prvi in drugi primer) izračun z vrednostmi spremenljivk...4 t (tretji primer)	19 t
3) naloga (3 primeri)	skrčitev in izpostavitev skupnega faktorja...3 t	9 t
4) naloga (2 primera)	reševanje besedilnih nalog: rvi primer (pravilen izbor)...3 t drugi primer (pravilen zapis, decimalna št.)...4 t	7 t
5) naloga (3 primeri)	razstavljanje izrazov: prvi in drugi primer (pravilen zapis)...2 t tretji primer (pravilen izbor)...3 t	7 t
6) naloga (1 primer)	izražanje ploščine in obsega lika s spremenljivkami: oščina...6 t obseg...6 t	12 t
7) naloga (2 primera)	enakovrednost izrazov...3 t	6 t
8) naloga (3 primeri)	enakovrednost izrazov...2 t	6 t
9) naloga (1 primer)	enočleniki...5 t	5 t
10) naloga (1 primer)	zapis izraza po besedilu...6 t	6 t
11) naloga (2 primera)	seštevanje enočlenikov: rvi primer (pravilen izbor)...3 t drugi primer (pravilen zapis)...4 t	7 t
12) naloga (1 primer)	zapis enačbe po besedilu in izračun...6 t	6 t

### 3 Opis domene matematičnih testov za tematski sklop Izrazi

Sistem za generiranje testov, ki se prilagodijo učencu, smo preizkusili na domeni rešenih spletnih testov iz matematike, ki se nanašajo na tematski sklop Izrazi. Ta vsebinsko sodi v učni načrt za deveti razred osnovne šole, znotraj teme Aritmetika in algebra. Spletne teste je sestavila Praprotnik Alenka, prof.. Testiranje je realizirala v dveh osnovnih šolah in sicer na Vrhniki in v Horjulu. Testi so bili sestavljeni v skladu z didaktičnimi standardi. Rešeni testi so bili anonimno shranjeni v podatkovno bazo MySQL (Praprotnik 2004).

#### 3.1 Testi iz matematike (vse naloge)

Domena zajema 79 primerov rešenih spletnih testov, ki so jih reševali učenci dveh osnovnih šol. Test je bil sestavljen iz 30 pod-nalog (glej Tabela 1), ki predstavljajo attribute v naš celotni učni množici. Naloge so dveh tipov in sicer naloge tipa »Izberi pravilen odgovor«, kjer učencu ponudimo tri odgovore in se mora odločiti za enega ter naloge tipa »Vpiši odgovor«, kjer mora učenec rešiti nalogo in rešitev zapisati v vnosno polje. Vrednosti atributov so številčne (razen razreda, ki je opisen), oziroma predstavljajo število točk, ki jih je posamezen učenec dosegel pri reševanju posamezne naloge. Na testu je bilo mogoče doseči 100 točk. Razred oziroma atribut Uspeh, ki ga modeliramo, je uvrstitev učenca glede na doseženo število točk in sicer na naslednji način: če je doseženo število točk več kot 65, potem razred dobi vrednost »odlično«, če je število točk med 30 in 65, razred dobi vrednost »dobro« in če je število točk manj kot 30, potem razred dobi vrednost »negativno«.

#### 3.2 Testi iz matematike (združene naloge)

Drugo domeno smo zgradili z združevanjem pod-nalog. 30 pod-nalog smo združili v 12 smiselni enot oziroma v 12 nalog, kot jo prikazuje Slika 1.

### 4. Poskusi in rezultati

Poskuse na obeh opisanih domenah, ki ju sestavljajo rešeni testi iz osnovnošolske matematike, tematski sklop Izrazi, smo izvedli s programskim paketom WEKA (WEKA, 2004).

#### 4.1 Poskusi z ocenjevanjem kakovosti nalog

Za ocenjevanje kakovosti nalog smo uporabili algoritem *ReliefF* (Robnik-Šikonja in Kononenko 2003) ter metodologijo, s katero metoda s podpornimi vektorji obtežuje attribute (Cristianini in Shaw-Taylor 2000). Namen ocenjevanja kakovosti nalog je bil preveriti domnevo ali so naloge, ki jih učitelj oceni kot »težje« tudi kakovostnejše.

Na domeni, kjer so testi predstavljeni s 30 pod-nalogami, je algoritem *ReliefF* ocenil z najboljšo oceno pod-naloge, ki se nanašajo na enakovrednost izrazov oziroma naloge 7a),

8b), 7b), poenostavitev oziroma naloga 2b) ponovno enakovrednost izrazov ali naloga 8c), besedilna naloga (izraz po besedilu) 10), izračuni: veččlenikov, oziroma nalogi 2aa) in bb), izračun obsega po sliki oziroma naloga 10a).

Na drugi domeni je algoritem *ReliefF* ponovno izbral 7) nalogo, oziroma nalogo o enakovrednosti izrazov, na drugo mesto je uvrstil besedilno nalogo 10) (izraz po besedilu) in na tretje je uvrstil 8) nalogo, ki je enakovrednost izrazov.

Z algoritmom s katerim metoda s podpornimi vektorji obtežuje podporne vektorje, so, v domeni s 30 pod-nalogami, bile izbrane naslednje naloge: 8b) enakovrednost izrazov, 6) izračun obsega na sliki, 10) besedilna naloga (izraz po besedilu), 2b) poenostavitev, 3a) skrajšitev in izpostavljanje 7b) in 7a) enakovrednost izrazov, 3c) skrajšitev in izpostavljanje, 6) izračun ploščine po sliki itd.

Z istim algoritmom smo ocenili attribute še na drugi domeni z 12 atributi. Na prvo mesto je bila uvrščena naloga 8) o enakovrednosti izrazov, na drugo naloga 3) (skrajšitev, skupni faktorji) in na tretje mesto naloga 6) (izračun ploščine oziroma obsega po sliki).

Namen ocenjevanja kakovosti širše množice nalog je izbiranje ožje množice nalog, ki bi jo učitelj lahko uporabil za hitro ocenjevanje učenca. V kontekstu visokega šolstva bi to lahko pomenilo izbiranje množice vprašanj, ki so primerna za ustne izpite.

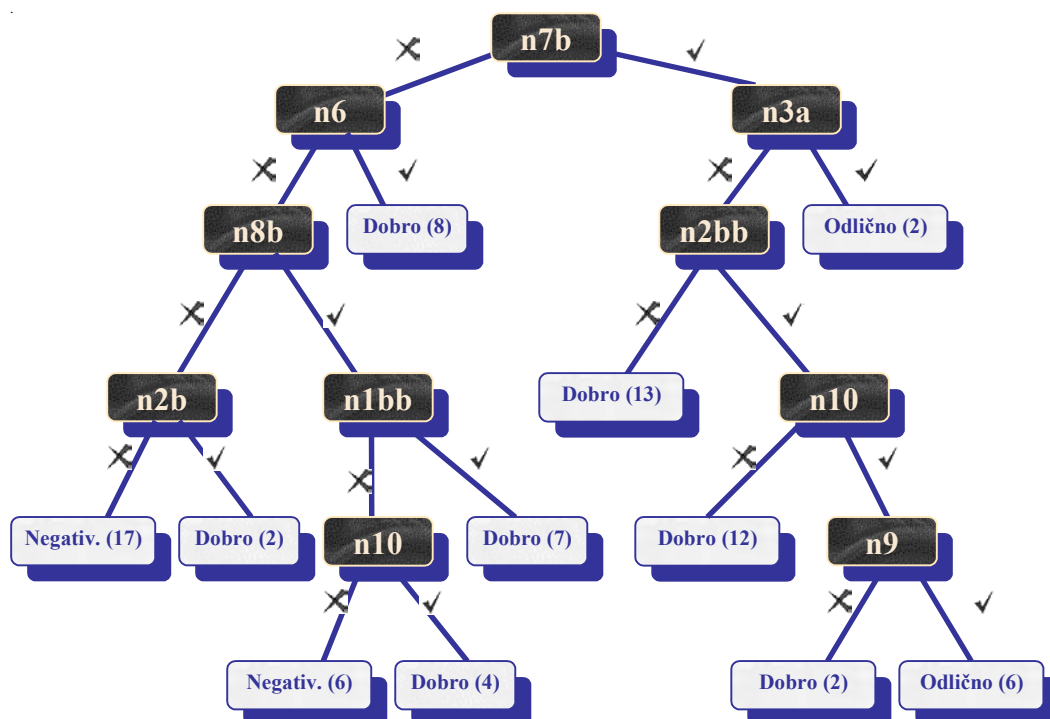
Oba algoritma sta poudarila pomen razumevanja teme *Enakovrednosti izrazov* za celotno razumevanje tematskega sklopa *Izrazi*. Visoko je bila uvrščena tudi naloga, ki od učenca zahteva, da bi zapisal *izraz po besedilu*. Razlika med algoritmoma nastopi pri ocenjevanju naloge, ki od učenca zahteva *izračun obsega* oziroma *ploščine*. *ReliefF* je ocenil za pomembno nalogo za izračun obsega po sliki, ki jo je rešilo 48% učencev medtem, ko je *algoritem*, ki sloni na *metodi s podpornimi vektorji* več poudarka dal tudi nalogi, ki zahteva *izračun ploščine* po sliki, ki jo je rešilo le 18 % učencev. Enako velja za nalogo, ki od učenca zahteva *krčitev izrazov* in *izpostavljanje skupnega faktorja*. To nalogo sta rešila le 2% učencev, *algoritem*, ki sloni na *metodi s podpornimi vektorji*, pa jo je izbral kot pomembno. To pomeni, da omenjeni algoritem izbira »težje« naloge, ki so morda primerne za ocenjevanje učencev za katere domnevamo, da bodo dosegli odličen uspeh. Naloge izbrane z *ReliefF* algoritmom so primerne za ocenjevanje najštevilčnejšega razreda učencev, ki so ocenjeni z dobrim uspehom.

#### 4.2 Odločitveni drevesi na obeh domenah

S podatki obeh domen smo zgradili dve *odločitveni drevesi* (Slika 1, Slika 2). Pri tem smo uporabljali paket J48, v sklopu programskega paketa WEKA (WEKA 2004, Witten in Frank 2000), ki zgradi C4.5 obrezano drevo. Drevesi smo zgradili na celotni učni množici in ju obrezali.

Klasifikacijska točnost, ocenjena z desetkratnim prečnim preverjanjem na domeni s 30 pod-nalogami, je 78%, medtem ko je klasifikacijska točnost na celotni domeni 96%.

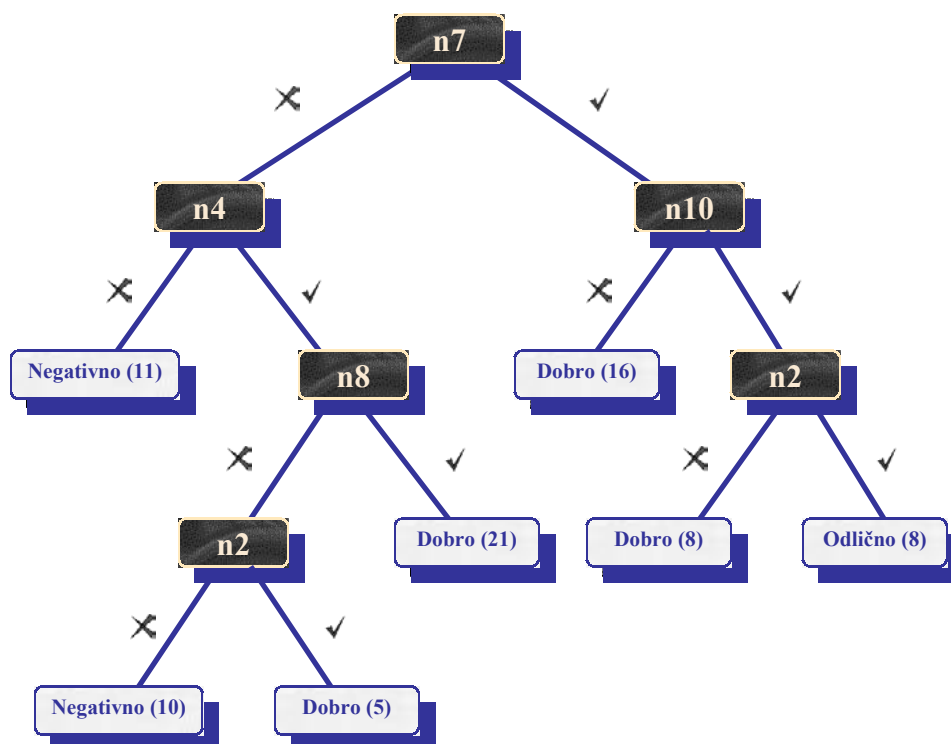
Klasifikacijska točnost, ocenjena z desetkratnim prečnim preverjanjem na domeni z 12 nalogami, je 76%, medtem ko je klasifikacijska točnost na celotni domeni 94%.



Slika 1: Odločitveno drevo na celotni učni množici, na domeni s tridesetimi pod-nalogami

Oznaka vozlišča	Naloga	Preverja znanje
n7a	7a) naloga	Enakovrednost izrazov (kvadrat)
n6	6) naloga	Izračun obsega na sliki
n3a	3a) naloga	Skrčitev in izpostavitev (razlika)
n8b	8b) naloga	Enakovrednost izrazov (razlika)
n2bb	2bb) naloga	Izračun veččlenika za vrednosti spremenljivk
n2b	2b) naloga	Poenostavitev izraza
n1bb	1bb) naloga	Razpoznavanje veččlenikov
n10	10) naloga	Zapis izraza po besedilu
n9	9) naloga	Razpoznavanje enočlenikov (zahtevnejše)
Oznaka lista	Številka v oklepaju	Skupno št. uvrščenih primerov/pravilno
Odlično	št. uvrščenih primerov v posameznem listu drevesa	8/8
Dobro		48/46
Negativno		23/22
Oznaka	Pomen	
✗	naloga je bila rešena pravilno	
✓	naloga je bila rešena napačno	

Tabela 2: Tabela z legendami oznak za domeno s tridesetimi pod-nalogami



Slika 2: Odločitveno drevo na celotni učni množici, na domeni z dvanajstimi nalogami

Tabela 3: Tabela z legendami oznak za domeno z dvanajstimi nalogami

Oznaka vozlišča	Naloga	Preverja znanje
n7	7) naloga	enakovrednost izrazov
n4	4) naloga	besedilna naloga
n10	10) naloga	zapis izraza po besedilu
n8	8) naloga	enakovrednost izrazov
n2	2) naloga	poenostavitev in izračun z vrednostmi spremenljivk
Oznaka lista	Številka v oklepaju	Skupno št. uvrščenih primerov/pravilno
Odlično	št. uvrščenih primerov v posamezen list drevesa	8/8
Dobro		50/46
Negativno		21/20
Oznaka	Pomen	
✗	naloga je bila rešena pravilno	
✓	naloga je bila rešena napačno	

Obe odločitveni drevesi sta poudarili pomen razumevanja enakovrednosti izrazov, zapisa izraza po besedilu,

poenostavitve in izračuna z vrednostmi spremenljivk, ki je povsem v skladu s strokovno oceno učiteljev o pomembnosti

tem ter z ocenami kakovosti nalog, ki jih je podal algoritem *ReliefF*.

Zaradi omejenosti učne množice težko govorimo o zanesljivosti modelov, kljub temu pa menimo, da naši poskusi potrjujejo možnost uporabe *odločitvenih dreves* za individualno ocenjevanje učencev, saj so pokazali zadovoljivo klasifikacijsko točnost.

## 5 Zaključki in nadaljnje delo

*Odločitveno drevo* naj bi predstavljalo absorpcijo znanja s strani množice učencev. Učitelju je zato dobra povratna informacija o uspešnosti posredovanja znanja iz določene teme, hkrati pa sta pri absorpciji znanja zajeti tudi narava znanja (določene teme so namreč že a priori težje za razumevanje), socialna struktura učencev (nekaterim je omogočena večja individualna pomoč), psihološki faktor strahu pred novo tehnologijo (določeni učenci so bolj navajeni na delo z novimi informacijskimi tehnologijami, saj imajo dostop do spleta od doma). To pomeni, da *odločitveno drevo* odraža pomembnost konkretne naloge v konkretnem okolju za razvrstitev učencev v skupine.

Omenjeni *odločitveni drevesi* sta zgrajeni s premajhno učno množico in zato težko ocenimo njuno zanesljivost. Zanesljivo *odločitveno drevo* bi moralo biti zgrajeno na večji učni množici, zato da bi statistično odražalo povprečni vzorec obnašanja učencev določene starosti pri absorpciji znanja iz teh tem. Spletni testi, ki bi se prilagajali učenčevemu znanju, bi algoritem za uvrščanje učenca s pomočjo *odločitvenega drevesa*. Najprej bi mu zastavili nalogo, ki je v korenu drevesa in bi se po vsaki rešeni nalogi »odločili« za vejo, s tem pa za naslednjo nalogo, glede na pravilnost odgovora tekoče naloge, vse dokler ne pridemo do lista in s tem do ocene. Takšno testiranje bi bilo zelo hitro, saj najdaljša veja *odločitvenega drevesa*, ki ga prikazuje Slika 1 šteje le 5 pod-nalog od 30 možnih. Slabosti modelov vidimo predvsem v tem, da niso 100% zanesljivi. Klasifikacijska točnost omenjenih prototipnih dreves je le okrog 80%. S stališča modela je to dobra klasifikacijska točnost, s stališča učencev, ki jih model uvrsti napačno, pa je to velika napaka. Napačna klasifikacija je še posebej problematična v primeru, ko gre za uspešnega učenca z nizko samozavestjo, ki ga je sistem oziroma model uvrstil v slabši razred. Poleg tega ni mogoče predvideti, kako bo napoved oziroma klasifikacija vplivala na motivacijo učencev, s tem pa posredno tudi na končni rezultat (Praprotnik 2004).

Zanimiva bi bila uporaba *odločitvenih dreves* v vlogi »svetovalca«. Učencu, ki se uvrsti v levi list nekega drevesa, bi lahko sistem ponudil pomoč le za naloge oziroma teme, ki jih je na poti do uvrstitve rešil napačno, s tem pa bi mu omogočil »prehod« v desno pod-drevo omenjenega drevesa. To pomeni, da bi učenec z negativno oceno, »pokazal« kaj je potrebno izboljšati za dober uspeh, dobrim pa kaj za odličnega.

Ob koncu testiranja bi lahko realizirali še spletni vprašalnik za shranjevanje spletnih naslovov, ki učenec pomagajo pri razumevanju tematskega sklopa. Učenec bi

ob koncu testiranja potem pokazali kaj se zdi zanimivo njihovim sošolcem s podobnim vzorcem znanja.

Izdelani modeli so primerni predvsem za poskusno uporabo s strani učencev, ne pa tudi za dokončno ocenjevanje s strani učiteljev. Učitelj se mora pri oblikovanju ocene opirati predvsem na strokovno presojo, ocene samoprilagoditvenih testov pa kritično pretehtati ter jih upoštevati le z določeno zanesljivostjo.

## Reference

- Aied2003, (2003): Artificial Intelligence in Education, [spletni dokumenti] <<http://www.cs.usyd.edu.au/~aied/>>.
- Bohanec, M., (2004): »Strojno učenje, Ljubljana: Institut Jožef Stefan«, [spletni dokument], gradivo za podiplomske študente FOV, Univerza v Mariboru, <<http://www-ai.ijs.si/MarkoBohanec/ai/ai.html>>.
- Cristianini, N., Shawe-Taylor, J., (2000): »An Introduction to Support Vector Machines and Other Kernel-based Learning Methods«, Cambridge University Press.
- Gasar, S., Bohanec, M. in Rajkovič, V. (2002): »Napovedovanje uspešnosti zaključka šolanja«, Organizacija, letnik 2, razprava, številka 8, oktober 2002, str. 508-513.
- Gasar, S., Bohanec, M., Rajkovič, V., (2003): »A combined data mining and decision support approach to educational planning«, V: Mladenović, D. (ur.), Lavrač, N. (ur.), Bohanec, M. (ur.), Moyle, S. (ur.). *Data mining and decision support: integration and collaboration*, The Kluwer international series in engineering and computer science, SECS 745, Boston; Dordrecht; London: Kluwer Academic Publishers, 2003, str. 203-212.
- Kononenko, I., (1997): »Strojno učenje«, Ljubljana: Založba FE in FRI.
- Michalski, R. S. in Kaufman, K. A. (1998): »Data Mining and Knowledge Discovery: a Review of Issues and a Multistrategy Approach«, Chichester, John Wiley & Sons.
- Mitchell, T.M., (1997): »Machine learning«, McGraw-Hill.
- Robnik-Šikonja, M., (2001): »Lastnosti in uporaba hevristične funkcije Relief v strojnem učenju«, doktorska disertacija, Univerza v Ljubljani, Fakulteta za Računalništvo in informatiko.
- Robnik-Šikonja, M. in Kononenko, I., (2003): »Theoretical and Empirical Analysis of Relief and RReliefF«, Machine Learning Journal, 2003, št. 53, str. 23 – 69.
- Praprotnik, A., (2004): »Spletni sistem za pomoč pri učenju in poučevanju«, diplomsko delo, Univerza v Ljubljani, Pedagoška fakulteta.
- WEKA (2004): The University of Waikato, [programska oprema za univerzitetno rabo], <<http://www.cs.waikato.ac.nz/~ml/weka/index.html>>.
- Witten, H. I., in Frank, E., (2000): »Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations«, San Francisco, California, Morgan Kaufmann Publishers.