

Večparametrski hierarhičen model za svetovanje pri izbiri izbirnega predmeta

Hierarchical multiattribute model for giving advice your optional subject

Robert Murko

Osnovna šola Videm, Videm pri Ptuj 47, 2284 Videm pri Ptuj
robert.murko2@quest.arnes.si

Povzetek

V prispevku želim s pomočjo metod umetne inteligence iz baze podatkov o uspehu pri posameznih predmetih ugotoviti pravila za uspešno izbiro izbirnih predmetov v tretji triadi. Postaviti želim večparametrski hierarhičen model za izbor izbirnega predmeta glede na dosežen uspeh pri posameznih predmetih in zaključni uspeh v prejšnjem razredu. Osredotočil se bom na tri predmete, in sicer iz družboslovnega sklopa bo to drugi tuji jezik nemščina in vzgoja za medije, iz naravoslovno-tehničnega sklopa pa izbirni predmet obdelava gradiv. Iz baze podatkov o uspehu v drugi triadi bom poiskal kriterije, ki so glavni prediktorji za uspeh pri izbirnem predmetu. Z metodo odkrivanja znanja iz podatkov bom poiskal te povezave. V prispevku se bom zaenkrat omejil na matično šolo. Bazo znanja si bom zagotovil s pomočjo strojnega učenja in ostalih postopkov rudarjenja na bazi podatkov o učencih in izbirnih predmetih. Prispevek je pripravljen kot začetek dela na tem področju in zaradi večje preglednosti ter lažjega razumevanja osnovne ideje so zajeti samo nekateri podatki in ne vsi, ki so na razpolago in omogočajo še kvalitetnejše rezultate. Cilj je olajšati izbiro izbirnega predmeta učencem in jim tako zagotoviti najvišjo kvaliteto izobraževanja.

Ključne besede: umetne inteligence, izbirni predmeti, devetletka, izbira, odločitveno drevo

Abstract

The aim of the project is to determine the rules for a successful choice of elective subjects in the third triad of primary school by the use of artificial intellect methods and by taking data from achievements at different subjects. Another aim is to set up hierarchical multiattribute model for choosing an elective subject according to achievements in certain subjects and final achievement mark in last period primary school. The focus of the project are three subjects, second foreign language German and Media education from the social-humanistic complex and Material handling from natural-technical complex. The data about achievements in the second triad are the basis for establishing criteria that are the main predictors for being successful at certain elective subject. These relations are found by using a method where knowledge is detected from data. The project is limited to the central school. The base of attainments is going to be assured by using machine learning and other mining procedures on the basis of students' and elective subjects' data. The project is prepared as a starting point in this field therefore to get a clearer view and for the sake of better understanding of the principals not all but only a few available data is included that enable better quality results. The

purpose is to alleviate the choice of choosing an elective subject and assure the students the best possible quality of education.

Key words: artificial intellect, an elective subject, 9-year primary school, a choice, a decision making tree

1. Uvod

Izbirni predmeti v devetletni osnovni šoli so nov način prilagajanja osnovne šole individualnim razlikam in interesom učencev. Devetletna osnovna šola uvaja možnost izbire med več izbirnimi predmeti. Mehanizem učinkovitega omogočanja poglobljanja in širitve znanja v tej obliki poznajo skoraj vsi šolski sistemi razvitih držav. Z izbirnimi predmeti želimo v različnih sistemih mladim omogočiti, da poudarijo tako imenovane močne strani lastnih interesov in sposobnosti. Pri izbranih predmetih se pojavi priložnost, da se učenci in učenke pri njih izkažejo. Izbirni predmeti so vključeni v predmetnik tretjega vzgojno-izobraževalnega obdobja.

Izhodišče prispevka se nanaša na kvaliteten pomoč pri svetovanju in izbiri izbirnega predmeta. Saj so učenci postavljeni pred pomembno in dokaj težko odločitev. Med množico ponujenih izbirnih predmetov morajo izbrati sebi najprimernejšega. Zato je pomemben ustrezen odločitveni model za podporo pri sprejemanju odločitve o izbiri izbirnega predmeta.

Model bom izdelal na podlagi baze podatkov uspeha doseženega v predhodnem šolskem letu. S pomočjo strojnega učenja bom pripravil odločitveni model za pomoč pri svetovanju o izbiri izbirnega predmeta. Vsaka šola, ponudi izbirne predmete, izbrane med 86 izbirnimi predmeti, ki so potrjeni s strani Strokovnega sveta Republike Slovenije, šola ponudbo oblikuje v skladu z interesi in sposobnostmi učencev, ter v skladu z usposobljenim kadrom. Končna odločitev za izbor je na strani otrok in staršev, ki izberejo med ponujenimi predmeti, v veliko pomoč jim je lahko kvaliteten model za podporo pri odločitvi izbire.

2. Strojno učenje in gradnja odločitvenega modela

Strojno učenje (machine learning) je veja raziskav umetne inteligence, ki je v zadnjih desetih letih s praktičnega vidika močno napredovala, kar se odraža v številnih komercialnih sistemih za strojno učenje in njihovi uporabi v industriji, medicini, ekonomiji itd. Strojno učenje se uporablja predvsem za iskanje znanja v podatkih (data mining) ter za avtomatsko generiranje baz znanja za ekspertne sisteme [4].

Naučljivost temelji na teoriji izračunljivosti, saj je učenec omejen na Turingov stroj. Teorija naučljivosti odgovarja na osnovna vprašanja: ali obstaja program, ki se lahko nauči po končnem branju iz neomejenega zaporedja besed nad neko abecedo pravilo, ki bo znalo razlikovati besede iz danega jezika od besed, ki niso iz tega jezika. Če je za dani jezik odgovor pritrdilen, pravimo, da je jezik naučljiv. Izkaže se, da je naučljivost tesno povezana z izračunljivostjo in vse omejitve v zvezi z izračunljivostjo veljajo tudi za naučljivost.

Orodje, ki sem ga uporabil za učenje, je SEE5. To orodje je primerno za obravnavan primer. See5 omogoča izgradnjo odločitvenega drevesa in postavitev odločitvenih pravil. Vzorec je relativno majhen, zato pričakujem velike napake. Za kvalitetnejše rezultate bi moral uporabiti bazo podatkov za več šol, s čimer bi zmanjšal neobjektivne vplive, kot so vpliv učitelja na predmet, vpliv okolja, vpliv tehničnih možnosti za izvedbo pouka. Saj so ti vplivi po šolah različni. S tem bi bili rezultati relevantnejši, vendar pa bo za začetek in vzpodbudo dovolj.

3. Odločitvena drevesa za izbirne predmete: drugi tuji jezik, obdelava gradiv in vzgoja za medije

3.1 Odločitveno drevo

3.1.1 Odločitveno drevo za izbirni predmet drugi tuji jezik nemščina:

```
izbi = tjn:
:...uspeh in 1,2: og (0)
:   uspeh = 3: og (10)
:   uspeh = 4: og (4)
:   uspeh = 5: vzm (6/2)
```

Učenec, ki je dosegel odličen učni uspeh in izbral izbirni predmet vzgoja za medije, bo izbral tudi izbirni predmet drugi tuji jezik nemščina. Učenci s prav dobrim in dobrim učnim uspehom zraven izbirnega predmeta obdelava gradiv izberejo prav tako drugi tuji jezik nemščino.

3.1.2 Odločitveno drevo za izbirni predmet vzgoja za medije:

```
izbi = vzm:
:...nar in 1,2: og (0)
:   nar = 3: og (14)
:   nar = 4: og (2)
:   nar = 5:
:     :...tit in 1,2: tjn (0)
:       tit = 3: og (2)
:       tit = 4: tjn (2)
:       tit = 5: tjn (10)
```

Učenci, ki dosežejo pri tehniki in tehnologiji odličen uspeh ter pri naravoslovju odličen uspeh, izberejo zraven drugega tujega jezika vzgojo za medije. Vzgojo za medije izberejo tudi učenci, ki so dosegli pri naravoslovju prav dober in dober uspeh, vendar v kombinaciji z izbirnim predmetom obdelava gradiv.

3.1.3 Odločitveno drevo za izbirni predmet obdelava gradiv:

```
izbi = og:
:...ang = 1: vzm (0)
:   ang = 2: vzm (6)
:   ang = 3: vzm (4)
:   ang = 5: tjn (10)
:   ang = 4:
:     :...gv in 1,2: vzm (0)
:       gv = 3: tjn (2)
:       gv = 4: vzm (8)
:       gv = 5:
:         :...uspeh in 1,2,3: vzm (0)
```

```
uspeh = 4: vzm (2)
```

```
uspeh = 5: tjn (2)
```

Učenci, ki dosežejo odličen uspeh pri prvem tujem jeziku, izberejo drugi tuji jezik in obdelavo gradiv. Prav tako izberejo učenci z odličnim splošnim učenim uspehom in prav dobrim (4) uspehom pri prvem tujem jeziku ter odličnim uspehom pri glasbeni vzgoji obdelavo gradiv. Učenec, ki ima splošen učni uspeh prav dober, glasbeno vzgojo odlično, prvi tuji jezik prav dober, izbirni predmet vzgojo za medije, izbere obdelavo gradiv.

3.2 Odločitvena pravila za svetovanje pri izbiri izbirnega predmeta:

Vsako odločitveno drevo lahko spremenimo v množico pravil. Vsaka pot od korena do lista ustreza enemu odločitvenemu pravilu. Vsake množice odločitvenih pravil pa ni možno neposredno združiti v drevo. Zato je predstavitev znanja s pravili bolj prožna kot predstavitev z drevesi. [5]

Pravilo 1: (10, lift 2.3)

```
uspeh = 3
```

```
izbi = tjn
```

```
-> class og [0.917]
```

Če je izbirni predmet drugi tuji jezik nemščina in učni uspeh dober (3), potem učenci izberejo predmet obdelava gradiv.

Učenci s splošnim učenim uspehom dober (3) in izbirnim predmetom drugi tuji jezik v kombinaciji izberejo iz sklopa naravoslovno-tehničnih predmetov obdelavo gradiv z vnaprejšnjo verjetnostjo 0,917.

Pravilo 2: (4, lift 2.1)

```
uspeh = 4
```

```
izbi = tjn
```

```
-> class og [0.833]
```

Če je izbirni predmet drugi tuji jezik nemščina in učni uspeh prav dober (4), potem učenci izberejo predmet obdelava gradiv.

Učenci s prav dobrim splošnim učenim uspehom zraven izbirnega predmeta drugi tuji jezik nemščina izberejo obdelavo gradiv z vnaprejšnjo verjetnostjo 0,833.

Pravilo 3: (30/12, lift 1.5)

```
izbi = vzm
```

```
-> class og [0.594]
```

Če je izbirni predmet vzgoja za medije, potem izberejo drug izbirni predmet obdelava gradiv.

Kombinacija izbirnih predmetov iz naravoslovno-tehničnega sklopa obdelava gradiv in iz družboslovnega sklopa vzgoja za medije izberejo z vnaprejšnjo verjetnostjo 0,594.

Pravilo 4: (8, lift 3.1)

```
ang = 4  
gv = 4  
-> class vzm [0.900]
```

Če ima učenec glasbeno vzgojo prav dobro (4) in tuji jezik prav dobro(4), izbere vzgojo za medije.

Učenci s prav dobrim uspehom pri prvem tujem jeziku angleščina in prav dobrim uspehom pri glasbeni vzgoji izberejo z vnaprejšnjo verjetnostjo 0,9 izbirni predmet vzgoja za medije. To so verjetno učenci z dokaj dobrim posluhom in jim jeziki ne povzročajo težav.

```
Pravilo 5: (6, lift 3.1)  
  
ang = 2  
izbi = og  
-> class vzm [0.875]
```

Če učenec izbere obdelavo gradiv in je pri angleščini ocenjen zadostno(2), izbere vzgojo za medije.

Učenci s slabšim uspehom pri prvem tujem jeziku (angleščina = 2) izberejo kombinacijo predmetov iz družboslovnega sklopa vzgoja za medije, iz naravoslovno-tehničnega sklopa pa obdelava gradiv. Kar je razumljivo, saj so to verjetno učenci, ki imajo težave z učenjem tujih jezikov.

```
Pravilo 6: (4, lift 2.9)  
  
ang = 3  
izbi = og  
-> class vzm [0.833]
```

Če učenec izbere obdelavo gradiv in doseže pri angleškem jeziku(3) dober uspeh, izbere vzgojo za medije.

Učenci s povprečnim uspehom pri prvem tujem jeziku (angleščina = 3) izberejo kombinacijo predmetov iz družboslovnega sklopa vzgoja za medije iz naravoslovno tehničnega sklopa pa obdelava gradiv. Kar je podobno kot pri prejšnjem pravilu logično, saj so to verjetno učenci, ki imajo manjše težave z učenjem tujih jezikov.

```
Pravilo 7: (2, lift 2.6)  
  
uspeh = 4  
gv = 5  
-> class vzm [0.750]
```

Če učenec doseže prav dober splošni učni uspeh in pri glasbeni vzgoji odličen (5), izbere vzgojo za medije.

Učenci s prav dobrim (4) splošnim učnim uspehom in odličnim (5) uspehom pri glasbeni vzgoji izberejo z vnaprejšnjo verjetnostjo 0,75 izbirni predmet vzgoja za medije.

```
Pravilo 8: (6/2, lift 2.2)  
  
uspeh = 5  
izbi = tjn
```

```
-> class vzm [0.625]
```

Če učenec izbere drugi tuji jezik nemščino in doseže odličen učni uspeh (5) izbere vzgojo za medije.

Učenci z odličnim splošnim učenim uspehom izberejo pogosteje dva izbirna predmeta iz družboslovnega sklopa, pomeni, da zraven vzgoje za medije izberejo drugi tuji jezik nemščino, z vnaprejšnjo verjetnostjo 0625.

```
Pravilo 9: (10, lift 3.0)
```

```
nar = 5
```

```
tit = 5
```

```
izbi = vzm
```

```
-> class tjn [0.917]
```

če ima učenec pri naravoslovju odlično (5), pri tehniki in tehnologiji odlično (5) izbirni predmet vzgoja za medije izbere še tuji jezik nemščino.

Učenci z odličnim uspehom pri naravoslovju in tehniki in tehnologiji izberejo kombinacijo dveh družboslovnih predmetov in sicer vzgojo za medije in drugi tuji jezik nemščino. Ta rezultat je malce presenetljiv, saj bi glede na uspeh pri naravoslovnih predmetih pričakoval, da bodo ti učenci izbrali dva naravoslovno-tehnična predmeta in enega družboslovnega.

```
Pravilo 10: (34/20, lift 1.3)
```

```
izbi = og
```

```
-> class tjn [0.417]
```

Če je učenec izbral obdelavo gradiv izbere tudi tuji jezik nemščino.

Kombinacija izbirnih predmetov iz naravoslovno tehničnega sklopa obdelava gradiv in iz družboslovnega sklopa tuji jezik nemščina je z vnaprejšnjo verjetnost 0,417.

4. Prečno preverjanje (Cross validate):

Preden znanje, dobljeno s strojnim učenjem uporabimo, moramo preveriti in oceniti njegovo kvaliteto. Poleg eksperimentov ocene izgrajenega znanja je za strojno učenje bistvena empirična evalvacija točnosti hipotez s pomočjo statističnih metod, saj omejeni vzorci podatkov lahko narobe predstavljajo splošno porazdelitev podatkov [6].

V nadaljevanju sem pripravil rezultate prečnega preverjanja za obdelani primer, dobljeni s programom SEE5. Rezultati prečnega preverjanja so v razredu napake 13 %, z oceno variabilnosti 3 %. Glede na majhno bazo so rezultati zadovoljivi. Za boljše rezultate bo potrebno razširiti bazo podatkov, tako da bo zajetih čim več podatkov o učencih različnih šol v tretji triadi. Potrebno bi bilo s primernimi metodami izbrati samo tiste parametre, ki so pomembno prediktivni pri izbiri izbirnega predmeta v tretji triadi. Za uporabnejši model bi morali zajeti vse slovenske osnovne šole. Tako bi lahko učencem svetovali izbrati najprimernejši izbirni predmet.

5. Zaključek

Z metodo strojnega učenja sem zgradil odločitvena drevesa in odločitvena pravila za izbirne predmete: drugi tuji jezik nemščina, vzgoja za medije iz družboslovno humanističnega sklopa, ter obdelavo gradiv iz sklopa naravoslovno-tehničnih predmetov. Na podlagi atributov baze in odkritih zakonitosti sem pripravil odločitvena drevesa za napovedovanje izbire izbirnega predmeta, ki pa glede na

nezadostno validacijo lahko služi le kot primer za poskusno uporabo. Njegova napovedna točnost je razmeroma majhna in bi jo bilo zaželeno izboljšati. Dobljena drevesa in pravila so za širšo uporabo dokaj neuporabna, saj so izdelana samo za tri izbirne predmete iz velike množice izbirnih predmetov. Vendar se lahko razširijo na širšo populacijo, takrat pa bi dobili zelo uporabne rezultate in učencem olajšali izbiro. Namen tega prispevka je ravno ta, da vzpodbudim razmišljanje v tej smeri, saj imamo obilico podatkov, ki pa jih ne uporabimo. Smiselno bi bilo k vsemu omenjenemu dodati tudi rezultate, ki jih učenci dosegajo na tekmovanjih, oziroma določiti, kateri rang tekmovanj se upošteva in izdelati obsežne modele za pomoč pri svetovanju pri izbiri izbirnega predmeta.

Literatura:

- Bohanec, M. (2001). Metode umetne inteligence. Učno gradivo za podiplomske študente FOV
- Bohanec, M. in Bratko, I. (1994). Trading Accuracy for Simplicity in Decision Trees. Machine Learning, 15, 223-250 str.
- Bohanec, M. in Rajkovič, V. (1999). Multi-Attribute Decision Modeling: Industrial Applications of DEX. Informatica, 23 (4), 487-491 str.
- Kononenko I. (1997). Strojno učenje, Fakulteta za računalništvo in informatiko, Ljubljana
- Kononenko I. (2000). Nekatere vidike strojnega učenja, umetne inteligence in zavesti, Fakulteta za računalništvo in informatiko, Ljubljana.
- Mitchell Tom M. (1997). Machine Learning, McGraw-Hill.
- Ministrstvo za šolstvo in šport: <http://www.mss.gov.si>
- Sample applications using See5: <http://www.rulequest.com/see5-examples.html#NEXT>

Strokovni življenjepis

Robert Murko je študent podiplomskega študija na Fakulteti za organizacijske vede Univerze v Mariboru smer management informacijskih sistemov in je od leta 2000 zaposlen kot računalnikar – organizator informacijske dejavnosti na Osnovni šoli Videm pri Ptuj. Diplomiral je na Pedagoški fakulteti v Mariboru smer Fizika – proizvodno tehnična vzgoja.

Robert Murko is a postgraduate student of the Faculty of Organisational Science of the University of Maribor in the course of Management of informational systems and has been working as a computer science teacher since the year 1999 at the primary school Videm Pri Ptuj. He graduated at the Faculty of Education in Maribor in the course of Physics and Technics.