

*UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO*

*Mihec Pezdirc*

## **Prenos znanja med klasifikatorji v strojnem učenju**

*DIPLOMSKO DELO  
NA UNIVERZITETNEM ŠTUDIJU*

*Mentor: doc. dr. Zoran Bosnić*

*Ljubljana, 2011*



Št. naloge: 01701/2010

Datum: 01.09.2010

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **MIHEC PEZDIRC**

Naslov: **PRENOS ZNANJA MED KLASIFIKATORJI V STROJNEM UČENJU**  
**KNOWLEDGE TRANSFER BETWEEN CLASSIFIERS IN MACHINE**  
**LEARNING**

Vrsta naloge: Diplomsko delo univerzitetnega študija

Tematika naloge:

V diplomski nalogi naj kandidat implementira in empirično ovrednoti prenos znanja med klasifikacijskimi modeli (klasifikator-učitelj in klasifikator-učenec) z uporabo pristopa simuliranja funkcij klasifikatorjev. K izbiri primerov za prenos znanja naj pristopi s principom aktivnega učenja, s katerim naj na različne načine poskuša hevristično poiskati ustrezne primere za klasifikatorja-učenca. Dobljene rezultate z uporabo vsaj petih klasifikatorjev in na več testnih domenah naj statistično ovrednoti.

Mentor:

doc. dr. Zoran Bosnić

Dekan:

prof. dr. Nikolaj Zimic



*Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavlanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.*

*Namesto te strani vstavite original izdane teme diplomskega dela s podpisom mentorja in dekana ter žigom fakultete, ki ga diplomant dvigne v študentskem referatu, preden odda izdelek v vezavo.*

# IZJAVA O AVTORSTVU

## *diplomskega dela*

*Spodaj podpisani Mihec Pezdirc,  
z vpisno številko 63050080,  
sem avtor diplomskega dela z naslovom:*

*Prenos znanja med klasifikatorji v strojnem učenju*

*S svojim podpisom zagotavljam, da:*

- *sem diplomsko delo izdelal samostojno pod mentorstvom doc. dr. Zorana Bosnića,*
- *so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela,*
- *soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".*

*V Ljubljani, dne 17. 01. 2011*

*Podpis avtorja:*

# **Zahvala**

*Zahvaljujem se mentorju doc. dr. Zoranu Bosniću, ki mi je pomagal ne samo pri izdelavi, temveč tudi pri izbiri teme diplomskega dela.*

# Kazalo

<i>Povzetek</i>	1
<i>Abstract</i>	2
1. <i>Uvod</i>	3
2. <i>Pregled povezanih področij strojnega učenja</i>	5
2.1. <i>Prenos znanja</i>	5
2.1.1. <i>Uporaba prenosa znanja</i>	6
2.2. <i>Aktivno učenje</i>	7
2.3. <i>Kombiniranje odločitvenih modelov</i>	8
2.4. <i>Oponašanje funkcije in EM algoritem</i>	8
2.4.1. <i>Učenje iz označenih in neoznačenih primerov</i>	9
2.5. <i>Opis uporabljenih modelov za prenos znanja</i>	10
2.5.1. <i>Model k najbližjih sosedov</i>	10
2.5.2. <i>Umetna nevronska mreža</i>	10
2.5.3. <i>Naivni Bayesov klasifikator</i>	12
2.5.4. <i>Odločitveno drevo</i>	12
2.5.5. <i>Metoda podpornih vektorjev</i>	13
3. <i>Predlagane metode prenosa znanja</i>	14
3.1. <i>Prenos znanja s ponovno označitvijo učne množice</i>	14
3.2. <i>Prenos znanja s filtriranjem po verjetnosti napovedi</i>	14
3.3. <i>Prenos znanja s filtriranjem po zanesljivosti napovedi</i>	15
4. <i>Metodologija testiranja metod prenosa znanja</i>	16
4.1. <i>Metodologija testiranja</i>	16
4.1.1. <i>Prečno preverjanje</i>	16
4.1.2. <i>Ocenjevanje statistične značilnosti rezultatov</i>	16
4.2. <i>Okolje in uporabljene knjižnice</i>	17
4.3. <i>Opis domen</i>	18
5. <i>Rezultati</i>	19
5.1. <i>Analiza rezultatov glede na modele strojnega učenja</i>	20
5.2. <i>Analiza rezultatov glede na uporabljene metode prenosa znanja</i>	21
6. <i>Sklepne ugotovitve</i>	22
6.1. <i>Dosežene izboljšave</i>	22
6.2. <i>Odprte teme na področju prenosa znanja</i>	22

<i>A: Rezultati testiranja.....</i>	<i>23</i>
<i>Seznam slik .....</i>	<i>34</i>
<i>Seznam tabel.....</i>	<i>35</i>
<i>Literatura.....</i>	<i>37</i>



## Seznam uporabljenih kratic in simbolov

<i>SVM</i>	<i>metoda podpornih vektorjev (support vector machines)</i>
<i>ANN</i>	<i>umetna nevronska mreža (artificial neural network)</i>
<i>kNN</i>	<i>k najbližjih sosedov (k nearest neighbours)</i>
<i>EM</i>	<i>maksimizacija pričakovanj (expectation maximization)</i>

## Povzetek

*V diplomski študiju smo razvili in testirali tri metode za prenos znanja med modeli strojnega učenja ter poskušali ugotoviti, s kakšno uspešnostjo se znanje lahko prenaša.*

*Prva izmed predlaganih metod izvaja prenos znanja, pri katerem model-učitelj ponovno označi razrede primerov iz učne množice, nad katerimi se je sam učil, nato pa tako označene primere posreduje modelu-učencu. Drugi pristop prenosa znanja izvaja prenos le z uporabo deleža najbolj označenih primerov z najvišjo verjetnostjo razreda, tretji pa izvaja prenos z deležem le najbolj zanesljivo označenih učnih primerov (uporabljena je ocena zanesljivosti iz literature).*

*Prenos znanja testiramo z uporabo modelov: SVM, ANN, kNN, naivni Bayes ter odločitveno drevo. Testiranja so bila opravljena na desetih različnih domenah za vse tri razvite algoritme prenosa znanja in med petimi omenjenimi modeli strojnega učenja.*

*Rezultati kažejo, je z razvitimi metodami znanje možno prenesti iz modela-učitelja v model-učenca. V večini primerov prenosa se pri tem klasifikacijska točnost statistično ne spremeni (v 61 % primerih), v 33% primerov se statistično zmanjša, v 6% primerov pa se poveča. V kontekstu pridobivanja na klasifikacijski točnosti kaže največjo prednost model naivnega Bayesa, kjer je pri prenosu znanja z uporabo vseh treh metod doseglo več modelov doseglo višjo klasifikacijsko točnost kot nižjo.*

### *Ključne besede:*

*prenos znanja, strojno učenje, umetna inteligenca, oponašanje funkcije, aktivno učenje*

## Abstract

*The aim of the thesis is to propose and test three methods for knowledge transfer between various machine learning models and evaluate the success of such transfers in terms of classification accuracy.*

*The first among the proposed methods performs the knowledge transfer by re-labeling the learning set by the teacher-model and passing this newly labeled learning set to the student-model for learning. The second proposed method filters out a chosen ratio of examples that have the lowest probability of the most probable class (with respect to the predicted class probability distribution). The third method uses filtering of examples based on a classification reliability estimate (proposed in the related work in progress).*

*We evaluate knowledge transfer using five different machine learning models: SVM, ANN, kNN, naive Bayes and decision trees. The tests for all three developed knowledge transfer methods and five models were performed on ten different benchmark domains (UCI).*

*The results indicate that the knowledge transfer resulted in a statistically unchanged classification accuracy in 61% of experiments. In 33% of experiments classification accuracy decreased and in 6% of experiments the classification accuracy increased. In terms of gaining classification accuracy, the biggest advantage was shown using the naive Bayes model that after having transferred knowledge from another model, it achieved the highest percentage of classification accuracy increases.*

### **Key words:**

*knowledge transfer, machine learning, artificial intelligence, active learning*

# Poglavje 1

## Uvod

Gradnja klasifikacijskih in regresijskih modelov nadzorovanega strojnega učenja poteka na podlagi atributnega opisa množice primerov poljubne domene. Vsak izmed primerov v domeni je opisan z  $n$  atributi ter enorazredno ali regresijsko spremenljivko. Slika 1.1 prikazuje takšno zgradbo primera z atributnim opisom, ki vsebuje attribute  $A1$ - $A_n$  in ciljno spremenljivko.

$A1$	$A2$	...	$A_n$	Razred/Regresijska spremenljivka
------	------	-----	-------	----------------------------------

**Slika 1.1: Atributna predstavitev primerov v poljubni domeni.**

Induktivno zgrajeni modeli iz takšnih predstavitev vsebujejo znanje, ki ga uporabljamo za napovedovanje novih primerov, isto znanje pa lahko uporabimo tudi pri učenju drugih modelov strojnega učenja. Vprašanje, ki se poraja, je: Ali se lahko znanje uspešno prenaša iz enega modela v drugega?

Motivacij za takšen prenos znanja je lahko več. Osnovni, naučeni model, je lahko prekompleksen, da bi ponudil dobro interpretacijo svoje napovedi. Mogoče označitev novih primerov traja več časa, kot bi želeli, in potrebujemo hitrejši model. Model, ki ga želimo uporabiti, je morda preobčutljiv na šum, ki ga neki drug model pred prenosom znanja lahko zgladi.

Modeli strojnega učenja so med seboj predvsem semantično različni oziroma vsebujejo znanje v različnih oblikah – v obliki odločitvenega drevesa, pravil, nevronov in uteži na povezavah med njimi, ločevalnih hiperravnin. Uporaba principa črne škatle (slika 1.2) za način prenosa znanja med vsemi modeli, ki se gradijo na podlagi atributnih predstavitev, je zaradi toliko različnih predstavitev edini univerzalni način, ki zagotavlja delovanje algoritma za prenos znanja med poljubnimi modeli strojnega učenja.

Vsi modeli, med katerimi se znanje prenaša, se morajo biti sposobni učiti iz atributne predstavitve učnih primerov ter imeti sposobnost napovedovanja in vračanja verjetnostne ocene te napovedi. Edini način prenosa znanja z uporabo principa črne škatle je tako označevanje učnih primerov s strani modela strojnega učenja, poimenovanega učitelj, ki prenaša svoje znanje modelu, poimenovanemu učenec.



**Slika 1.2: Pri prenosu znanja lahko na modele gledamo kot na črne škatle, podrobnosti implementacije niso pomembne. Pomembna sta le izhod in vhod v model.**

Namen tega dela je sprva implementirati način prenosa znanja med modeli strojnega učenja, nato pa opraviti množico testov prenosa znanja na desetih različnih domenah med

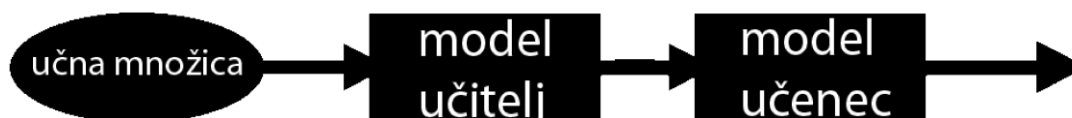
*petimi klasifikatorji (metoda podpornih vektorjev, umetna nevrnska mreža, naivni Bayesov klasifikator, k najbližjih sosedov ter klasifikacijsko drevo) in ugotoviti, ali takšen pristop omogoča doseganje večje točnosti učenja.*

## Poglavje 2

### Pregled povezanih področij strojnega učenja

#### 2.1. Prenos znanja

Osnovni način prenosa znanja temelji na modelu-učitelju, ki prenese svoje znanje modelu-učencu. Preden učitelj lahko prenese znanje, ga seveda mora pridobiti iz učne množice, znanje pa nato lahko prenese na učenca. Pri prenosu znanja se lahko uporabljajo različni pristopi. Najpreprostejši je ponovna označitev učne množice in posredovanje te modelu-učencu. Ta pristop ne zahteva poznavanja zgradbe hipoteze posameznega modela in mu pravimo model črne škatle. Notranjega delovanja modela nam ni treba poznati, vedeti pa moramo, kaj so vhodi in izhodi. Ekvivalenten pojem je pojem ograjevanja, poznan iz objektno usmerjenih programskih jezikov. Ta pravi, da nam ni treba poznati podrobnosti nekih metod, le njihove vmesnike. Koncept takega prenosa znanja prikazuje slika 2.1.



Slika 2.1: Osnovni potek prenosa znanja.

Čeprav se znanje modela – črne škatle prenaša posredno z označevanjem učnih primerov, lahko nanj gledamo kot na funkcijo, ki preslika eno hipotezo v drugo hipotezo. Ampak ali ta definicija popolnoma drži? V primeru transduktivnega učenja hipoteze namreč ne zgradimo. Po drugi strani pa lahko rečemo, da je hipoteza kar sama učna množica. Če je prenos znanja torej funkcija, ki preslika eno hipotezo v drugo, je ta funkcija kar model-učitelj, če je model-učenec transduktiven model. To seveda drži, ko govorimo o najpreprostejšem načinu prenosa znanja, npr. metodi prenosa znanja s ponovno označitvijo učne množice, ki je opisana v poglavju Predlagane metode prenosa znanja (stran 14). Ko pa se pri prenosu znanja uporablja dodaten inteligenten način izbiranja primerov za učenje učne množice, pa funkcija prenosa znanja poleg modela vsebuje še ta inteligentni način izbire primerov za učenje.

Če želimo doseči popolno posnemanje učitelja, je najlažje uporabiti model kNN s parametrom  $k=1$ . Večjo ko imamo učno množico, bolj kNN posnema model učitelja. Če pa so v učni množici, ki jo metodi kNN določi model-učitelj, vsi možni primeri iz neke domene, kNN povsem oponaša model-učitelja. Pravzaprav je metoda ekvivalentna temu, da shranimo vse možne klasifikacije metode učitelja v tabelo in jih pri klasifikaciji samo poiščemo.

Glede na to, kako in kateri primeri se označijo, smo ločili več načinov prenosa znanja. Prva delitev na podlagi obstoja učnih primerov, na katerih se je učil model, iz katerega se prenaša znanje, loči dve skupini:

- prenos znanja iz vnaprej podanih učnih primerov,
- prenos znanja s sprotnim določanjem učnih primerov.

Pri prenosu znanja iz vnaprej podanih učnih primerov uporabimo za gradnjo modela-učenca kar iste primere ali podmnožico nabora primerov, ki jih je za učenje uporabil model-učitelj.

Ko pa govorimo o prenosu znanja s sprotnim določanjem učnih primerov, pa se osnovna množica za gradnjo modela-učitelja ne uporablja. Primeri se določijo z vzorčenjem domene. Tak tip prenosa znanja se lahko uporablja npr. med avtonomnimi inteligentnimi agenti, ki se učijo iterativno iz okolja, učnih primerov pa zaradi prostorske omejitve pomnilnika in velike količine podatkov ne morejo shranjevati. V takšni situaciji je možno le graditi in na vsakem koraku učenja popravljati hipotezo, ki pa prostorsko ni tako zahtevna kot učna množica.

Druga delitev pa temelji na že opisanih sorodnih področjih:

- prenos znanja z aktivnim učenjem,
- prenos znanja z direktnim oponašanjem modela,
- prenos znanja s kombiniranjem odločitev več modelov.

Pri prenosu znanja z aktivnim učenjem (stran 7) gre predvsem za inteligentno izbiranje primerov modela-učitelja ali učenca. Tak prenos znanja smo implementirali in je opisan v poglavju *Predlagane metode prenosa znanja* (stran 14).

Prenos znanja z direktnim oponašanjem modela vključuje tako oponašanje funkcije kot tudi uporabo algoritma EM. Sem spada tudi bolj preprosta izvedba prenosa znanja, pri katerem se v učno množico ne dodaja novih učnih primerov, ampak se za gradnjo modela-učenca uporabi le na novo označeno osnovno učno množico.

EM algoritem deluje z iteracijo dveh korakov. Na prenos znanja v osnovi ne gledamo kot na iterativni proces, a si lahko zamislimo prenos znanja, ki deluje v več iteracijah. V prvem koraku model-učitelj prenese znanje na model-učenca, kar združuje oba EM koraka. V drugem koraku model-učenec oceni svoje znanje in modelu-učitelju poda novo množico, iz katere bi se rad učil. Prvi in drugi korak se nato ponavljata do nekega ustavitvenega pogoja. Tak prenos znanja ima dvosmerno komunikacijo med učiteljem in učencem.

Pri kombiniranju odločitev več modelov pa imamo več učiteljev in enega učenca, ki prek neke heuristike izbira odločitve učiteljev, na katerih se bo učil.

### 2.1.1. Uporaba prenosa znanja

Kdaj in s kakšnim namenom se prenos znanja sploh uporablja? Ali je prenos znanja sploh smiseln? Odgovorov na ta vprašanja je morda več, kot bi si sprva predstavljali.

- *Primer 1:*

Recimo, da imamo znanje na voljo le v obliki modela strojnega učenja oziroma hipoteze tega modela. Učne množice nimamo, želimo pa si, da bi neki drugi model naučili istega znanja. Mogoče je izbrani model-učenec časovno hitrejši pri napovedovanju od osnovnega modela. Prenos znanja je v takšnem primeru povsem primerna rešitev.

- *Primer 2:*

Ukvarjamo se z domeno, ki ima prisotnega veliko šuma, model, s katerim bi želeli zajeti znanje, pa je slabo odporen na ta šum. S prenosom znanja si pomagamo, tako da najprej naučimo na šum bolj odporen model, nato pa njegovo znanje prenesemo v zeleni model.

- *Primer 3:*

Včasih poleg nekega kompleksnega modela, npr. SVM, potrebujemo tudi razlago odločitev, ki je ta model preprosto ne more podati v preprosti in razumljivi obliki. Razlaga

odločitev je še posebej pomembna v ekspertnih sistemih, saj uporabnik poleg napovedi pričakuje še razlago napovedi oziroma način sklepanja modela. Znanje, ki ga je zajel SVM, lahko prenesemo v model, ki premore preprosto razlago odločitve. Takšna sta npr. naivni Bayes in odločitveno drevo. Naivni Bayes lahko s pomočjo nomogramov [9] pokaže, kako močno so kateri atributi vplivali na odločitev. Pri odločitvenem drevesu pa je razlaga odločitve kar pot od korena do lista. Po tej poti se prva delitev interpretira kot najpomembnejša, zadnja pa kot najmanj pomembna. Po prenosu znanja model-učenec povsem ne oponaša učitelja. Še vedno se najverjetneje razlikujeta v vsaj nekaj odločitvah. V tem primeru ne moremo podati razlage odločitve z modelom-učencem. Rešitev pri tem je na primer, da imamo s prenosom znanja naučenih več učencev in za razlago uporabimo tistega, čigar napoved se ujema z originalnim modelom.

- **Primer 4:**

Še en razlog za prenos znanja je lahko hitrost učenja. Modeli, kot je ANN, potrebujejo pri učenju veliko iteracij skozi učno množico. Če število potrebnih iteracij zmanjšamo, avtomatično povečamo napako modela. Morda pa lahko vplivamo na število učnih primerov, tako da tiste, ki ne dajo zadosti informacije o domeni, odstranimo iz učne množice. S tem upamo, da ohranimo klasifikacijsko točnost modela, obenem pa zmanjšamo čas gradnje tega modela. Prenos znanja uporabimo, tako da naučimo model-učitelja na osnovni učni množici, nato pa model-učitelj posreduje le delež učnih primerov modelu-učencu. Izbere jih glede na svojo oceno količine informacije, ki jo dobi prek ponovne označitve teh primerov.

- **Primer 5:**

Razlog za prenos znanja je lahko tudi izboljšanje klasifikacijske točnosti ali druge mere uspešnosti modelov strojnega učenja. Če se model-učenec uči iz učne množice, ki jo model-učitelj označi drugače, kot je bila označena originalna učna množica, se bo najverjetneje spremenila tudi uspešnost modela-učenca, ki se uči iz te na novo označene učne množice. Klasifikacijsko točnost modela-učenca primerjamo z modelom istega tipa, ki se uči iz originalne učne množice. Zavedati se moramo, da lahko seveda pride tudi do poslabšanja.

## 2.2. Aktivno učenje

Področje, povezano s prenosom znanja, je aktivno učenje, pri katerem se primeri, ki jih neki model strojnega učenja dobi v učno množico, izberejo na hevristično inteligenten način. V [3] so načini izbiranja primerov opisani kot izbiranje primerov iz domene:

- kjer se model slabo izkaže,
- kjer imamo nizko zaupanje v odločitve,
- kjer pričakujemo, da se bo model spremenil,
- kjer smo prej našli podatke, ki so prispevali k učenju.

Vsi ti načini delujejo v povezavi z modelom, ki ga želimo naučiti. Lahko bi rekli tudi, da nam model sam sporoči, kje potrebuje znanje. Po drugi strani pa se o aktivnem učenju govori tudi takrat, ko učno množico algoritma določimo brez povratne informacije o tem, kaj se dogaja v modelu.

Če je model sposoben iterativnega učenja, lahko po vsaki iteraciji učenja sam pove, kje v domeni še nima dovolj znanja. Eden izmed modelov, ki je sposoben iterativnega učenja, je nevronska mreža. Pristop za aktivno učenje te je razvil Mckay v [4].

Kot že omenjeno, želimo pri prenosu znanja uporabiti princip črne škatle, in zato pristopov, ki so namenjeni posameznim modelom, ne moremo uporabiti.



*Če model ni sposoben iterativnega učenja, se lahko množica učnih primerov inteligentno izbere pred učenjem. Lahko pa se model z vsakim novim primerom zgradi ponovno v celoti.*

*Pri prenosu znanja se lahko tako kot pri aktivnem učenju množica učnih primerov modela-učenca izbere na inteligen ten način. Učenec lahko sam pove, katerih primerov iz domene še ne razume najboljše, ali pa mu učne primere ustrezno izbere model-učitelj. Z aktivnim učenjem, brez informacije o tem, kaj se dogaja pri modelu-učencu, pri dveh razvitih metodah prenosa znanja zbiramo primere za njegovo učno množico. Učitelj sam določi, kje v domeni je največ znanja, in posreduje te podatke učencu.*

### **2.3. Kombiniranje odločitvenih modelov**

*Drugo sorodno področje je kombiniranje odločitvenih modelov. Tu se s pomočjo izbrane hevristike iz več modelov naenkrat ugotovi, katera izmed napovedi je najbolj zanesljiva.*

*Pri prenosu znanja ni rečeno, da se model-učenec ne more učiti od več učiteljev. Ti mu lahko ne posredujejo istih odgovorov in mora sam oceniti, kateri so bolj zanesljivi. Najpreprostejša metoda kombiniranja je glasovanje. Vsak klasifikator prispeva en glas za izbrani razred. Model-učenec se v takšnem primeru uči iz označitve razreda, ki ima največ glasov.*

*Prenosu znanja je s področja kombiniranja odločitvenih modelov podobna metoda naučenega kombiniranja z metaučenjem, opisana v [8]. Na validacijski množici vsak klasifikator vrne napoved za vsak testni primer. Napovedi se uporabijo kot atributi za metaučenje. Naloga učnega algoritma je iz napovedi posameznih klasifikatorjev napovedati pravi razred. Torej se funkcija kombiniranja zgenerira s pomočjo metaklasifikatorja. Na metaklasifikator, ki se uči iz napovedi odločitvenih modelov, lahko gledamo kot na model-učenca, a se ta ne uči iz atributov domene, temveč le iz napovedi posameznih klasifikatorjev.*

*Algoritmi, ki smo jih razvili, uporabljajo izhode enega učitelja.*

### **2.4. Oponašanje funkcije in EM algoritem**

*V [8] je oponašanje funkcije opisano kot metoda, pri kateri najprej rešimo klasifikacijski ali regresijski model z modelom strojnega učenja, ki daje najboljšo točnost. Nato pa s pomočjo tako naučenega modela zgeneriramo nove primere, ki jih dodamo v učno množico modela, ki bo oponašal originalni model. Vrednosti atributov, ki jih je potencialno lahko tudi neskončno, pa generiramo naključno, glede na porazdelitev atributov v originalni učni množici.*

*Tako razširjena učna množica je sedaj primerna za gradnjo drugih modelov. Ti potrebujejo več primerov, kot jih je v originalni učni množici, da dosežejo boljšo točnost napovedi. Z večanjem števila učnih primerov se dobljeni model vse bolj približuje funkciji originalnega modela.*

*Metoda je iteracija bolj splošne metode, znane pod drugim imenom, kot statistična metoda maksimizacija pričakovanj, EM (expectation maximization) algoritem. Ta deluje v dveh ponavljajočih se korakih. V terminologiji strojnega učenja je prvi korak EM algoritma enak označevanju novih primerov, drugi, maksimizacija, pa je učenje novega modela na novi zgrajeni učni množici. Metoda ponavlja oba koraka, dokler rezultati ne konvergirajo.*

*Pri prenosu znanja želimo, da model-učenec oponaša funkcijo modela-učitelja. V tem pogledu je oponašanje funkcije enakovredno prenosu znanja. Znanje se lahko prenaša tudi pod drugačnimi pogoji in na drugačne načine, kot je to opisano v poglavju Prenos znanja (stran 5). Oponašanje funkcije, kot je opisano v [8], je torej le eden izmed načinov prenosa znanja.*

### 2.4.1. Učenje iz označenih in neoznačenih primerov

Še eno izmed področij, ki se križa s področjem prenosa znanja, je učenje iz označenih in neoznačenih primerov. Učne množice velikokrat vsebujejo primere, ki nimajo označene odvisne spremenljivke (oziroma razreda, če gre za klasifikacijo). Področje se ukvarja s tem, kako označiti neoznačene primere in s tem prispevati k boljši točnosti napovedi.

V [10] je predstavljen problem klasifikacije petih različnih orientacij obraza. Učna množica vsebuje malo označenih in veliko neoznačenih primerov. Predstavljen je pristop, v katerem se različni modeli najprej učijo iz označenih primerov, nato pa s pomočjo EM algoritma in neoznačenih primerov poskušajo še izboljšati svoje napovedi. Testiranja so pokazala, da se izboljša klasifikacijska točnost vseh modelov razen modela strojnega učenja, ki ni sposoben zaznati odvisnosti med atributi. Vendar to samo takrat, ko je bilo modelu na voljo veliko označenih primerov.

Označevanje neoznačenih primerov lahko poteka tudi s kooperativnim učenjem. Algoritem, ki deluje s kooperativnim pristopom dveh modelov naivnega Bayesovega klasifikatorja, sta razvila A. Blum in T. Mitchell [2]. Pokazala sta, da predlagani algoritem (slika 2.2) dosega boljše rezultate na domeni klasificiranja spletnih strani.

Podano imamo:

- množico  $L$  označenih učnih primerov z atributi  $x = x_1 \cup x_2$
- množico  $U$  neoznačenih učnih primerov

Ustvari množico  $U'$  z naključnim zbiranjem iz  $U$ .

$K$  krat ponovi:

- z  $L$ , nauči klasifikator  $h_1$ , ki upošteva le  $x_1$
- z  $L$ , nauči klasifikator  $h_2$ , ki upošteva le  $x_2$
- s  $h_1$ , označi  $p$  pozitivnih in  $n$  negativnih primerov iz  $U'$
- s  $h_2$ , označi  $p$  pozitivnih in  $n$  negativnih primerov iz  $U'$
- dodaj te samooznačene primere v  $L$
- naključno izberi  $2p + 2n$  primerov iz  $U$ , ki predstavljajo novi  $U'$

**Slika 2.2: Algoritem kooperativnega učenja**

Končna klasifikacija uporablja Bayesovo lastnost pogojne neodvisnosti, ki dovoljuje, da se primeri klasificirajo s kombinacijo  $h_1$  in  $h_2$  po formuli 2.1.

$$P(r|x) = P(r|x_1)P(r|x_2) \quad (2.1)$$

Če algoritem bolje pogledamo, vidimo, da je le-ta priredba EM algoritma. Slaba stran je, da vsebuje nekaj parametrov, ki jih je potrebno optimizirati, če želimo doseči večjo uspešnost pri klasifikaciji.

Izboljšava modelov strojnega učenja z uporabo neoznačenih primerov je s stališča prenosa znanja vzpodbudna v situaciji, kjer model-učenec s pomočjo aktivnega učenja izbira primere, ki jih ne zna dobro klasificirati, nato pa postavi učitelju zahtevo za označitev tega predhodno neoznačenega primera.

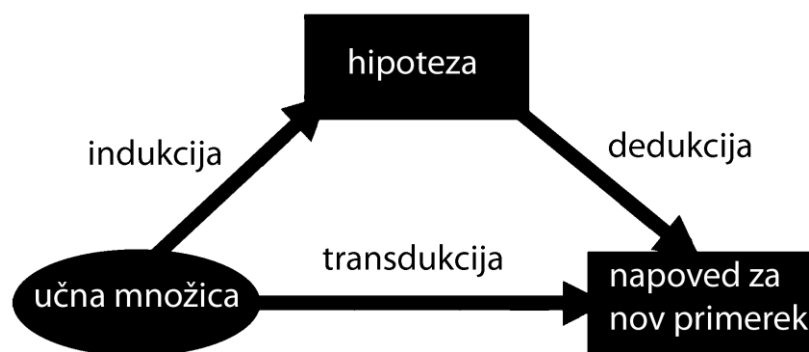
## 2.5. Opis uporabljenih modelov za prenos znanja

Prenos znanja smo testirali s petimi klasifikacijskimi modeli. To so metoda podpornih vektorjev, nevronska mreža z enim skritim nivojem, naivni Bayesov klasifikator,  $k$  najbližjih sosedov ter odločitveno drevo. Sledi kratek opis vseh uporabljenih modelov.

### 2.5.1. Model $k$ najbližjih sosedov

V nasprotju s klasičnimi modeli strojnega učenja, kjer z indukcijo zgradimo hipotezo ter nato s postopkom dedukcije sklepamo na odvisno spremenljivko, gre tu za transduktivno učenje. Hipoteze pri modelu  $k$ NN ne gradimo, ampak na označitev novih primerov sklepamo kar direktno iz učne množice, kot je prikazano na sliki 2.3. Primeru, ki ga klasificiramo, poiščemo  $k$  najbližjih sosedov v učni množici. Iz tako zbranih najbližjih sosedov klasificiramo primer v množico, ki je med sosedi najbolj zastopana v smislu najvišje frekvence vrednosti razreda. Za mero razdalje se lahko uporabita npr. evklidska ali manhattanska razdalja. Podatke je pred klasifikacijo potrebno normalizirati, tako da posamezni atributi z večjim razponom vrednosti nimajo prevelikih vplivov na klasifikacijo. Najpomembnejši parameter v metodi je  $k$ , nastavlja pa se ga glede na šum v podatkih. Če je veliko šuma, se večji  $k$  obnese bolje.

Razlaga odločitve modela  $k$  najbližjih sosedov temelji na podobnosti z najbližjimi  $k$  učnimi primeri.



Slika 2.3: Transdukcija je način sklepanja metode  $k$ NN.

### 2.5.2. Umetna nevronska mreža

Umetna nevronska mreža (ANN) je model, ki poskuša simulirati delovanje nevronov. Lahko se jo uporablja za regresijo ali klasifikacijo.

Znanje v nevronske mreži je shranjeno v obliki uteži na povezavah. Uteži  $w_{ij}$  se uporabljajo pri napovedovanju izhodov posameznih nevronov po enačbi 2.2. Dejansko je to vektorski produkt izhodov nevronov na prejšnji plasti  $x_i$  in uteži  $w_{ij}$  na povezavah med nevronom  $x_j$  in nevroni na prejšnji plasti. Temu se prišteje konstantna aktivacija, včasih poimenovana tudi pristranost (angl. bias) nevrone, označena s  $c_j$ . Nad takšnim izходом se uporabi še sigmoidna izhodna funkcija, kot npr. funkcija 2.3. Izhodne vrednosti nevronov se računajo po vrsti od prvega skritega nivoja do izhodnega nivoja.

$$A(x_j) = \sum_i w_{ij}x_i + c_j \quad (2.2)$$

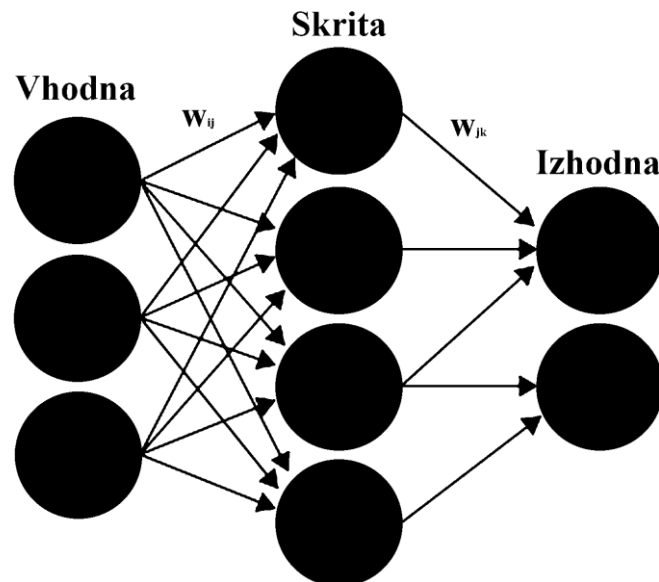
$$f(x) = \frac{1}{1 - e^{-x}} \quad (2.3)$$

Ločimo lahko dva osnovna tipa nevronske mreže:

- nevronske mreže brez skrite plasti,
- nevronske mreže z eno ali več skritih plasti.

Nevronska mreža brez skrite plasti vsebuje vhodne neurone, ki so s sinapsami povezani z vsakim od izhodnih nivojev. Takšen tip nevronske mreže je sposoben zaznati linearne odvisnosti med posameznimi atributi in razredno ali regresijsko spremenljivko.

Če se med vhodni in izhodni nivo vrine še poljubno število med seboj nepovezanih nevronov, dobimo dvonivojsko nevronska mrežo z eno skrito plastjo. Skrita plast je povezana z vhodnimi in izhodnimi nevroni, kot je prikazano na sliki 2.4. Med vhodno in izhodno plast lahko postavimo poljubno število skritih plasti.



Slika 2.4: Umetna nevronska mreža z eno skrito plastjo.

Nevronske mreže s skritimi plastmi so sposobne reševati tudi nelinearne probleme.

Učni algoritem, ki ga uporabljajo nevronske mreže, je posplošeno pravilo delta. To oceni napako na izhodu nevronske mreže po napovedi primera in nato od izhodnih proti vhodnim nevronom popravlja uteži v smeri gradienta, s katerim napaka upada. Velikost koraka v smeri gradienta se določi s tako imenovanim faktorjem učenja. Metoda ne najde nujno globalnega minimuma funkcije napake, temveč lahko obtiči v kakšnem izmed lokalnih minimumov. Eden izmed problemov umetnih nevronske mreže je tudi potencialno dolg čas učenja, saj mora učni algoritem zaradi učenja s posplošenim pravilom delta velikokrat iterirati skozi učno množico primerov.

### 2.5.3. Naivni Bayesov klasifikator

Naivni Bayes je statistična metoda, ki deluje na podlagi iz učne množice ocenjenih pogojnih verjetnosti razredov  $P(r_k|v_i)$  in apriornih verjetnosti razredov  $P(r_k)$ . Tabeli teh dveh verjetnosti predstavljata znanje naivnega Bayesovega klasifikatorja.

Verjetnostno porazdelitev razredov lahko dobimo z enačbo (2.4)

$$P(r_k|V) = P(r_k) \prod_{i=1}^a \frac{P(r_k|v_i)}{P(r_k)} \quad (2.4)$$

Metoda predpostavlja pogojno neodvisnost vrednosti različnih atributov pri danem razredu, zato je manj uspešna pri domenah, kjer pri danih razredih obstajajo odvisnosti med atributi. Dobra stran metode je predvsem to, da že v osnovni ideji vrača verjetnosti razredov v nasprotju z npr. metodo podpornih vektorjev. Metoda ponuja tudi dobro razlago odločitev. Z nomogrami [9] se lahko tudi grafično prikaže vpliv posameznih atributov na razred. Metoda nekoliko slabše deluje pri zveznih spremenljivkah, ki jih je potrebno pred oceno pogojnih verjetnosti diskretizirati.

### 2.5.4. Odločitveno drevo

Odločitvena drevesa spadajo pod simbolične modele strojnega učenja. Osnovni algoritem učenja (slika 2.5) so predstavili Hunt in sod. [6]. Kar se razlikuje pri gradnji dreves, je način izbiranja "najboljšega" atributa.

Če je izpolnjen ustavitveni pogoj,  
potem postavi list, ki vključuje vse učne primere;  
sicer

- izberi najboljši atribut  $A_i$ ;
- označi naslednike z vrednostmi atributa  $A_i$ ;
- za vsako vrednost  $V_j$  atributa  $A_i$  ponovi:
  - rekurzivno zgradi poddrevo z ustrezno podmnožico učnih primerov;

**Slika 2.5: Gradnja odločitvenega drevesa [6].**

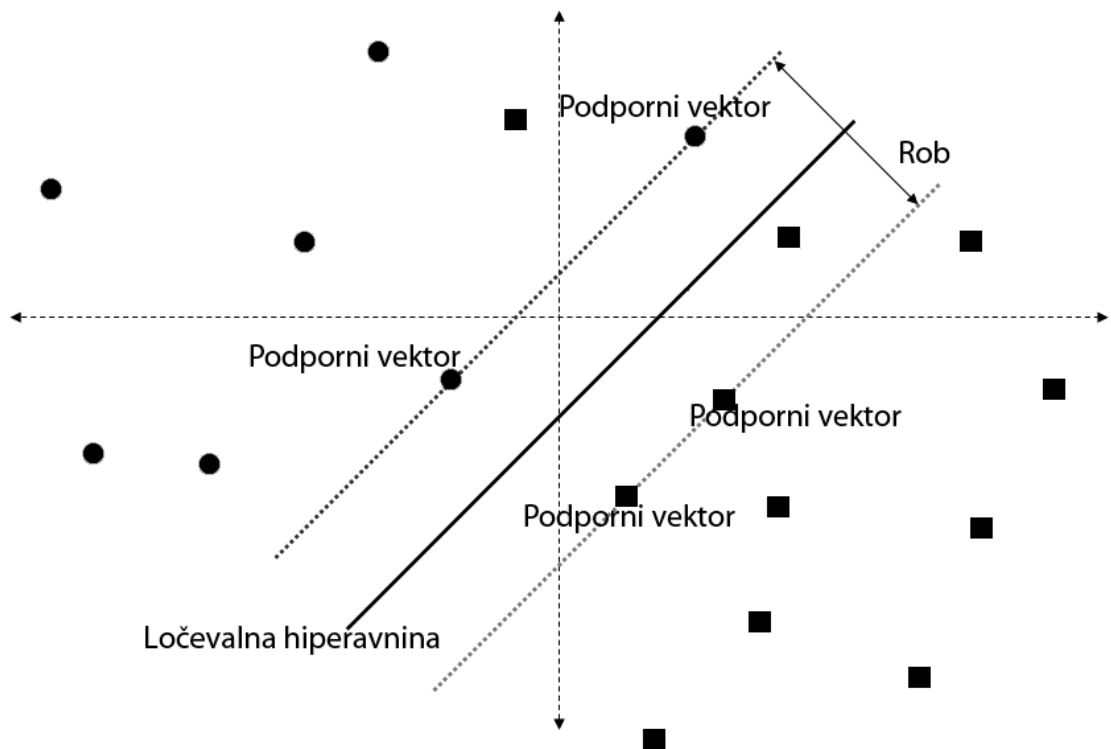
Klasifikacija poteka z uvrščanjem primera, ki ga želimo klasificirati v enega izmed listov drevesa. Začnemo pri korenu drevesa, kjer glede na zgrajeno drevo uvrstimo primer v poddrevo, ki mu pripada primer glede na trenutni razmejevalni atribut. Postopek rekurzivno ponavljamo, dokler primer ne pristane v enem izmed listov, ki vsebuje podmnožico učne množice. Primer klasificiramo v večinski razred te podmnožice.

Odločitvena drevesa ponujajo dobro razlago odločitev. Razlaga je namreč kar pot od korena drevesa do lista, v katerega spada klasificirani primer. Poleg naivnega Bayesovega klasifikatorja je odločitveno drevo edina metoda simboličnega učenja, ki smo jo izbrali za testiranje.

### 2.5.5. Metoda podpornih vektorjev

Pri problemu binarne klasifikacije je osnovni princip metode podpornih vektorjev [5] razmejiti dva razreda s hiperravnino. Če imamo večklasifikacijski problem, ga metoda razbije na več dvorazrednih problemov. Hiperravnino se poišče tako, da se maksimizira rob med podpornimi vektorji. Ti so primeri iz učne množice, ki so na robu ločevalne hiperravnine (slika 2.6). Ko primeri med seboj niso linearno ločljivi, deluje metoda tako, da projicira atributni prostor v večdimenzionalni prostor, v katerem lahko primeri postanejo linearno ločljivi. Rešitev podanega problema je v obliki kvadratičnega optimizacijskega problema, rešljivega z znanimi algoritmi.

Pri klasifikaciji se preprosto pogleda, na kateri strani hiperravnine je primer, ki ga želimo označiti. Čeprav je v metodi kar nekaj parametrov, je bolj stabilna od umetnih nevronske mreže. Metoda trenutno spada med najbolj uspešne metode strojnega učenja.



Slika 2.6: Ločevalna hiperravnina pri metodi podpornih vektorjev.

## Poglavje 3

### Predlagane metode prenosa znanja

#### 3.1. Prenos znanja s ponovno označitvijo učne množice

*Je najbolj osnovna metoda prenosa znanja. Kot tudi preostala dva modela, deluje po principu črne škatle in jo lahko uporabimo skupaj s katerim koli modelom strojnega učenja, ki deluje z atributno predstavitvijo domene. Prenos znanja poteka tako, da se najprej model-učitelj nauči na učni množici, nato ponovno označi vse primere iz učne množice in iz njih naredi novo učno množico. To nato uporabi za učenje model-učenec. Algoritem deluje tako na regresijskih kot tudi na klasifikacijskih modelih. Vendar smo teste uspešnosti opravili le na klasifikacijskih domenah in modelih.*

#### 3.2. Prenos znanja s filtriranjem po verjetnosti napovedi

*Kot pri prenosu znanja s ponovno označitvijo učne množice se uporablja princip črne škatle. Njen vhod so primeri v atributni predstavitvi, izhod pa je porazdelitev verjetnosti po razredih. Zadoščajo že samo verjetnosti najbolj verjetnih razredov.*

*Predpostavka, ki jo postavi ta metoda, je, da je v osnovni učni množici več podatkov, kot jih je potrebnih za gradnjo modela strojnega učenja. Zato lahko učitelj nauči učenca istega koncepta, ki se ga je naučil sam s podmnožico ponovno označenih primerov iz osnovne učne množice.*

*Metoda, ki temelji na aktivnem učenju, deluje tako, da najprej zgradimo model-učitelja, ki nato vse primere iz učne množice ponovno klasificira in vrne verjetnosti vseh napovedanih razredov. Nato sledi faza filtriranja novooznačene učne množice s stratificiranim pristopom. Iz vseh primerov, ki jih je učitelj klasificiral v določen razred, odstranimo delež  $q$  tistih z najnižjo verjetnostjo. Delež filtriranja  $q$  je definiran na intervalu  $[0,1]$ .*

*Obstaja možnost, da je veliko ocen verjetnosti napovedanih razredov enakih po vrednosti. V robnem primeru so lahko vse ocene med seboj enake. Takšnih ocen ne moramo urediti po velikosti in odstraniti  $q$  primerov iz učne množice. Problem rešimo tako, da vsem verjetnostim pred filtriranjem prištejemo naključno vrednost iz intervala:  $[0, 10^{-5}] \cup [-10^{-5}, 0]$ . Lahko bi rekli, da razpršimo verjetnost okoli originalne ocenjene verjetnosti. Tako je metoda veliko bolj robustna in zagotavlja odstranitev natanko deleža  $q$  vseh primerov.*

*Pri vseh testiranjih smo uporabili  $q=1/3$ . Izbiro tega parametra bi se dalo optimizirati z uporabo dodatne podmnožice učne množice, ki bi bila namenjena optimizaciji tega parametra. Lahko bi iskali tudi čim manjšo učno množico, ki prenese še zadosti znanja, da model-učenec dosega zeleno točnost.*

*Model-učenec se s to metodo potencialno uči hitreje od učitelja, saj dobi manj učnih primerov, ki pa so zato potencialno bolj reprezentativni. Regresijskih modelov s to metodo ni mogoče uporabiti, saj ti ne ocenjujejo verjetnosti napovedane regresijske spremenljivke. Za uporabo regresijskih modelov bi bilo potrebno uporabiti mere zaupanja v regresijsko spremenljivko, velikosti intervalov zaupanja ipd.*

### 3.3. Prenos znanja s filtriranjem po zanesljivosti napovedi

Je metoda, ki nadgradije prenos znanja z verjetnostnim filtriranjem učne množice v kombinaciji z aktivnim učenjem. Pri njej model-učitelj na inteligen način izbere učne primere za učenca, kot je opisano v nadaljevanju.

V [7] avtorji predlagajo pristop za ocenjevanje zanesljivosti klasifikacije. Ta temelji na funkciji napake klasifikacije oziroma njeni pričakovani vrednosti. Klasifikacijsko napako poljubnega primera lahko napišemo v obliki enačbe 3.1.

$$e(x) = |y - f(x)| \quad (3.1)$$

kjer je  $y$  pravilni razred, napovedan razred pa je označen z  $f(x)$ . Če označimo napoved modela kot  $\hat{y} = \max_i f_i(x)$ , se napaka poenostavi v  $1 - \hat{y}$ . Pričakovana vrednost napake  $E(e(x))$  se po [7] izračuna z uporabo t.i. referenčne ocene oziroma po formuli 3.2.

$$E[e(x)] = p_y(x) (1 - f_y(x)) + (1 - p_y(x)) f_y(x) \quad (3.2)$$

Resnične verjetnosti pravilno napovedanega razreda  $p_y(x)$  ne poznamo, zato jo aproksimiramo z  $\hat{y}$ . Tako v enačbi 3.3 dobimo referenčno ali privzeto oceno zanesljivosti.

$$O_{ref} = 2(\hat{y} - \hat{y}^2) \quad (3.3)$$

Ta referenčna ocena zanesljivosti ima dve zaželeni lastnosti. Izračunamo jo lahko za vsako klasifikacijo, pri kateri model vrne verjetnost napovedanega razreda. V primeru optimalnega modela pa je ta predikcija celo optimalna, saj je  $O_{ref}$  enak pričakovani vrednosti funkcije napake  $E[e(x)]$ .

Predstavljeno oceno zanesljivosti smo uporabili pri metodi filtriranja po zanesljivosti napovedi. Zaradi lastnosti ocene zanesljivosti lahko ponovno uporabimo princip črne škatle. Kot pri metodi z verjetnostnim filtriranjem pri tej metodi najprej zgradimo model-učitelja, ki nato vse primere iz učne množice ponovno klasificira in vrne verjetnosti vseh napovedanih razredov. Sledi faza filtriranja novooznačene učne množice s stratificiranim pristopom. Iz vseh primerov, ki jih je učitelj klasificiral v določen razred, odstranimo delež  $q$  z največjo vrednostjo ocene  $O_{ref}$ , ki predstavlja oceno napake. Obdržimo označene primere z najmanjšimi napakami. Faktor filtriranja  $q$  je ponovno delež, definiran na intervalu  $[0,1]$ . Vse ocene napake pred filtriranjem naključno razpršimo, kot smo to storili pri metodi prenosa znanja s filtriranjem po verjetnosti napovedi.

Če na oceno  $O_{ref}$  gledamo kot na funkcijo, lahko vidimo, da je definirana na intervalu  $[0,1]$ . Ima en maksimum pri  $\hat{y}=0,5$  od katerega pada z večanjem in manjšanjem  $\hat{y}$ . Ker pri dvorazrednih problemih klasificiramo primer v razred, katerega verjetnost je večja od 0,5, bo vedno veljalo  $\hat{y}>0,5$ . Funkcija na tem intervalu pada. To pomeni, da bomo pri filtriranju odstranili iste primere, kot če bi preprosto uporabili oceno  $\hat{y}$ . Z  $O_{ref}$  lahko filtriramo učno množico modela-učenca drugače le pri večrazrednih problemih, kjer imajo klasificirani primeri poleg  $\hat{y}>0,5$  tudi  $\hat{y}<0,5$ . Bolj zanesljivo klasificirani so torej tudi tisti primeri, ki jim v več kot dvorazrednih primerih označimo razred z verjetnostjo bližje 0.

Pri vseh testiranjih z metodo filtriranja po zanesljivosti napovedi smo uporabili delež  $q = 1/3$ .



## Poglavje 4

# Metodologija testiranja metod prenosa znanja

### 4.1. Metodologija testiranja

*Vse tri predlagane načine prenosa znanja ter vsa testiranja smo implementirali v programskem jeziku in okolju, namenjenemu statistični obdelavi podatkov – okolju R. Algoritme smo testirali in preverili, ali pride do statističnih sprememb v uspešnosti učenja glede na to, ali se model uči iz osnovne učne množice ali s prenosom znanja s katero od implementiranih metod. Uporabili smo torej statistične teste tipa primerjava dveh algoritmov na eni domeni.*

*Teste smo opravili na desetih domenah za vse tri načine prenosa znanja. Prenos znanja smo testirali s petimi modeli strojnega učenja: metoda podpornih vektorjev, dvonivojska nevronska mreža, naivni Bayesov klasifikator, k najbližjih sosedov in odločitveno drevo. Skupaj smo torej opravili  $5 \cdot 5 \cdot 10 \cdot 3 = 750$  testov.*

*Glavni namen testov je bil ugotoviti uspešnost prenosa znanja iz enega modela v drugega, in ne zgraditi najboljših možnih modelov strojnega učenja. Zato iskanje optimalnih množic parametrov za posamezne modele puščamo za nadaljnje delo.*

*Vsi testi kot tudi implementacije razvitih metod so priloženi na zgoščenci.*

#### 4.1.1. Prečno preverjanje

*Metoda, ki se uporablja za ocenjevanje učenja. Množico učnih primerov razdelimo na  $N$  podmnožic enake velikosti, zatem za vsako podmnožico zgradimo hipotezo tako, da za učenje uporabimo unijo preostalih množic. Dobimo  $N$  hipotez, ki jih lahko testiramo na množici, ki je nismo uporabili pri njihovi gradnji. Zavedati se moramo, da je ocena hevristična, saj ne testiramo končne hipoteze, ampak množico hipotez, ki so bile zgrajene na podmnožici končne učne množice. Prečno preverjanje je pravzaprav posplošena metoda izloči-enege, pri kateri se za testiranje hipoteze iz učne množice izloča en po en primer. Zaradi obsežnosti izračuna je lahko ta časovno nesprejemljiva.*

#### 4.1.2. Ocenjevanje statistične značilnosti rezultatov

*Za primerjavo, ali se točnost modela strojnega učenja po prenosu znanja iz enega izmed petih modelov vanj izboljša ali poslabša glede na osnovni algoritem, ki se je učil iz iste učne množice, smo uporabili t-test.*

*T-test smo uporabili skupaj s prečnim preverjanjem tako, da smo najprej naredili 10-kratno prečno preverjanje klasifikacijske točnosti enega izmed izbranih modelov (ANN, SVM, kNN, odločitveno drevo, naivni Bayes). Dobili smo 10 klasifikacijskih točnosti  $CA_{1,i}$ ,  $i = 1, \dots, 10$ , ki smo jih nato povprečili v  $\overline{CA_1}$ . Z 10-kratnim prečnim preverjanjem prenosa znanja v drugi model smo dobili 10 klasifikacijskih točnosti  $CA_{2,i}$ ,  $i = 1, \dots, 10$ , katerih povprečje je  $\overline{CA_2}$ .*

*T-test glede na  $CA_{1,i}$  in  $CA_{2,i}$  ter stopnjo zaupanja  $\beta$  izračuna p-vrednost, ki nam pove, kolikšna verjetnost je, da postavljena ničelna hipoteza drži. Pri t-testu smo postavili naslednje hipoteze:*

- ničelna hipoteza: oba algoritma sta enako uspešna,  $CA_1 = CA_2$ ,

- alternativna hipoteza:  $CA_1 \neq CA_2$ .

Če sprejmemo alternativno hipotezo in če velja  $\overline{CA_1} < \overline{CA_2}$ , je rezultat našega testiranja ugotovitev  $CA_1 < CA_2$  oziroma, če velja  $\overline{CA_1} > \overline{CA_2}$ , je rezultat našega testiranja ugotovitev  $CA_1 > CA_2$ .

Vse t-teste smo opravili s stopnjo zaupanja v alternativno hipotezo  $\alpha = 0,05$  (oziroma zaupanja v ničelno hipotezo 0,95). Če je verjetnost hipoteze manjša ali enaka navedeni vrednosti, lahko ničelno hipotezo zavržemo in sprejmemo postavljeno alternativno hipotezo. Rezultati vseh testov so predstavljeni v dodatku A (stran 23).

## 4.2. Okolje in uporabljene knjižnice

Vsi modeli so implementacije naštetih modelov strojnega učenja iz R zbirke knjižnic [11]:

- **k najbližjih sosedov (R knjižnica - ipred):** Uporabili smo evklidsko razdaljo in parameter  $k=5$ .
- **nevronska mreža z enim skritim nivojem (R knjižnica - nnet):** Uporabili smo naslednje parametre:
  - število nevronov na skriti plasti: (število nevronov na vhodni plasti + število izhodnih nevronov)\*2/3,
  - faktor konvergence uteži: 0.5,
  - največje število iteracij: 800.
- **odločitveno drevo (R knjižnica - rpart)**
- **naivni Bayesov klasifikator (R knjižnica - e1071)**
- **metoda podpornih vektorjev (R knjižnica - e1071):** Pri testiranju smo uporabiliil privzete parametre v R knjižnici e1071, ki so:
  - RBF jedrna funkcija,
  - toleranca ustavitvenega pogoja: 0.001,
  - epsilon = 0.1.

### 4.3. Opis domen

Domene, ki smo jih uporabili pri gradnji modelov ter testiranju prenosa znanja, so zbrane iz javno dostopnega repozitorija strojnega učenja UCI [1]. Izbrane domene so namenjene klasifikaciji in vsebujejo tako diskretne kot zvezne attribute. Povzete so v tabeli 4.1.

	<i>Podatkovna zbirka</i>	<i>Število primerov</i>	<i>Število diskretnih atributov</i>	<i>Število zveznih atributov</i>	<i>Število razredov</i>
1	<i>abalone</i>	4177	1	7	29
2	<i>iris</i>	150	0	4	3
3	<i>glass</i>	214	0	9	7
4	<i>wdbc</i>	569	0	30	2
5	<i>pima-indians diabetes</i>	768	0	8	2
6	<i>zoo</i>	101	16	0	7
7	<i>parkinsons</i>	197	0	22	2
8	<i>ionosphere</i>	351	0	34	2
9	<i>house-votes-84</i>	435	16	0	2
10	<i>optdigits</i>	5620	64	0	10

**Tabela 4.1: Povzetek vseh domen.**

## Poglavje 5

### Rezultati

Podrobni rezultati testiranja treh algoritmov prenosa znanja med petimi modeli strojnega učenja na desetih domenah so predstavljeni v Dodatku A (stran 23). V tem poglavju predstavljamo strnjeni povzetek. Pri testiranju s t-testom smo uporabili stopnjo zaupanja v ničelno hipotezo 0.95.

Povzetki rezultatov v tabelah 5.1, 5.2, 5.3 in 5.4 kažejo število in deleže izboljšav, poslabšanj ter ohranjanja klasifikacijske točnosti. Iz njih ni razvidno, do kako velikih sprememb klasifikacijske točnosti je dejansko prišlo. Te lahko razberemo iz tabel v Dodatku A (stran 23). Poudarimo lahko, da je do največjega izboljšanja prišlo pri prenosu znanja z metodo ponovne označitve učne množice iz modela naivnega Bayesa v model naivnega Bayesa. Klasifikacijska točnost se je povečala za 0,2189. Do največjega poslabšanja klasifikacijske točnosti je prišlo pri prenosu iz naivnega Bayesa v ANN. Klasifikacijska točnost se je zmanjšala za 0,5799.

Večini modelom-učencem klasifikacijska točnost ostane približno enaka (61 %). Zmanjša se nekoliko manjšemu deležu (33 %). Najmanj pa je takih, ki se jim klasifikacijska točnost poveča (6 %).

Prenos v model \ Prenos iz modela	SVM			ANN			naivni Bayes			kNN			odl. drevo		
	-	o	+	-	o	+	-	o	+	-	o	+	-	o	+
SVM	1	9	0	2	8	0	0	8	2	0	8	2	0	9	1
ANN	4	6	0	4	6	0	2	8	0	3	7	0	2	8	0
naivni Bayes	6	4	0	6	4	0	4	5	1	6	3	1	5	5	0
kNN	5	5	0	5	5	0	1	6	3	2	8	0	0	10	0
odl. drevo	5	5	0	4	6	0	2	8	0	3	7	0	0	10	0
vsota	21	29	0	21	29	0	9	35	6	14	33	3	7	42	1

**Tabela 5.1:** Povzetek rezultatov testiranja prenosa znanja s ponovno označitvijo učne množice (metoda 2). (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni.

Prenos v model \ Prenos iz modela	SVM			ANN			naivni Bayes			kNN			odl. Drevo		
	-	o	+	-	o	+	-	o	+	-	o	+	-	o	+
SVM	3	7	0	5	5	0	1	6	3	1	8	1	0	8	2
ANN	3	7	0	6	4	0	2	4	4	4	6	0	3	7	0
naivni Bayes	7	3	0	8	2	0	5	5	0	7	3	0	5	5	0
kNN	4	6	0	7	3	0	0	6	4	4	6	0	1	9	0
odl. Drevo	4	6	0	4	6	0	2	6	2	3	7	0	2	8	0
vsota	21	29	0	30	20	0	10	27	13	19	30	1	11	37	2

**Tabela 5.2:** Povzetek rezultatov testiranja prenosa znanja s filtriranjem po verjetnosti napovedi (metoda 2). (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni.

Prenos v model \ Prenos iz modela	SVM			ANN			naivni Bayes			kNN			odl. Drevo		
	-	o	+	-	o	+	-	o	+	-	o	+	-	o	+
SVM	3	7	0	3	7	0	1	6	3	0	9	1	0	8	2
ANN	4	6	0	5	5	0	2	4	4	4	6	0	3	7	0
naivni Bayes	8	2	0	8	2	0	2	7	1	6	4	0	5	5	0
kNN	4	6	0	6	4	0	0	6	4	4	6	0	1	9	0
odl. drevo	5	5	0	5	5	0	2	6	2	2	8	0	2	8	0
vsota	24	26	0	27	23	0	7	29	14	16	33	1	11	37	2

**Tabela 5.3:** Povzetek rezultatov testiranja prenosa znanja s filtriranjem po zanesljivosti napovedi (metoda 3). (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni.

	Metoda 1			Metoda 2			Metoda 3		
	-	o	+	-	o	+	o	-	+
Vsota	72	168	10	91	143	16	85	148	17
Delež v vseh metodah	0,096	0,224	0,013	0,121	0,191	0,021	0,113	0,197	0,023
Delež v metodi	0,288	0,672	0,040	0,364	0,572	0,064	0,340	0,592	0,068

**Tabela 5.4:** Povzetek rezultatov testiranja treh metod na desetih domenah. (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni. Metoda 1 je prenos znanja s ponovno označitvijo učne množice. Metoda 2 je prenos znanja s filtriranjem po verjetnosti napovedi. Metoda 3 je prenos znanja s filtriranjem po zanesljivosti napovedi. Vrstica "Vsota" vsebuje števila poslabšanj (-), izboljšanj (+) in nespremenjenih (o) klasifikacijskih točnosti po vseh metodah. Vrstica "Delež v vseh metodah" vsebuje deleže poslabšanj (-), izboljšanj (+) in nespremenjenih (o) klasifikacijskih točnosti glede na vseh 750 testov prenosa znanja prek vseh treh metod. Vrstica "Delež v metodi" vsebuje deleže poslabšanj (-), izboljšanj (+) in nespremenjenih (o) klasifikacijskih točnosti glede na posamezno metodo, ki obsega 250 testov.

## 5.1. Analiza rezultatov glede na modele strojnega učenja

Če opazujemo klasifikacijsko točnost osnovnih modelov strojnega učenja, tj. brez prenosa znanja, se prek vseh domen v povprečju najboljše obneseta modela SVM in ANN. Naivni Bayesov klasifikator, kNN ter odločitveno drevo se izkažejo nekoliko slabše.

Ker testiramo prenos znanja iz enega modela v drugega, pričakujemo, da modeli, ki v osnovi zgradijo slabo hipotezo, sicer bodo prenesli svoje znanje, vendar bo klasifikacijska točnost modela-učenca slabša kot klasifikacijska točnost istega tipa modela, brez prenosa znanja vanj. V večini rezultati empirično potrjujejo to trditev. Obstajajo pa tudi primeri, v katerih so slabši klasifikatorji v vlogi modela-učitelja bolj uspešni.

Če poljuben model dosega večjo klasifikacijsko točnost od drugega, ne moremo sklepati, da bo prenos znanja iz njega v slabši model prinesel izboljšanje klasifikacijske točnosti. Mogoče slabši model ni sposoben prejeti znanja in tako pride celo do drastičnega poslabšanja. Tak primer je npr. prenos iz odločitvenega drevesa z osnovno klasifikacijsko točnostjo 0,8324 v SVM, ki ima na neposredni učni množici sicer začetno

klasifikacijsko točnost 0,7331 v domeni "digits", vendar pa pri prenosu znanja iz odločitvenega drevesa doseže le klasifikacijsko točnost 0.2444

Ko pride do izboljšanja, prenos znanja izboljša modele kNN, naivni Bayes in odločitveno drevo. Ko pa sta v vlogi modela-učenca ANN in SVM, se v nobenem izmed testov njuna klasifikacijska točnost ne izboljša.

Klasifikacijska točnost se v veliko primerih ohranja, kar je kazalec, da se prenese iz enega v drug model enaka količina znanja, kot če bi se model-učenec učili iz osnovne učne množice.

Največkrat se izboljša klasifikacijska točnost pri prenosu znanja v naivni Bayesov klasifikator. Izboljšanja, ko je do teh prišlo, smo zabeležili pri prenosu znanja iz vseh modelov v naivni Bayesov klasifikator.

Motivacija za prenos znanja je med drugim tudi hitrost učenja modela-učenca. Če lahko temu zmanjšamo število učnih primerov, se čas učenja metodam potencialno zmanjša. Od vseh izbranih modelov ANN potrebuje največ časa za učenje, a se ta čas s prenosom znanja da zmanjšati. Klasifikacijska točnost se namreč pri prenosih znanja z metodama, ki odstranita 1/3 učne množice, ohranja. Število iteracij, ki jih ANN potrebuje, pa na teh domenah ostane približno enako. Nikoli se ne poveča za 1/3, kar bi pomenilo, da se čas učenja ohranja; ali celo poveča za več kot 1/3, kar bi pomenilo, da se čas učenja poveča za isti faktor.

## 5.2. Analiza rezultatov glede na uporabljene metode prenosa znanja

Iz tabele 5.4 je razvidno, da ima prva metoda (prenos znanja s ponovno označitvijo učne množice) manj izboljšanj klasifikacijske točnosti kot metodi 2 in 3. Preostali dve metodi imata približno isto število izboljšanj. A je zanimivo, da ima prva metoda tudi najmanj poslabšanj – sklepamo, da je glede na število izboljšanj in število poslabšanj najbolj stabilna.

Drugi dve metodi imata več poboljšanj in poslabšanj. Razlog za to lahko iščemo v tem, da filtriramo vnaprej točno določeno 1/3 podatkov. Če bi ta parameter pri tistih modelih, ki so se poboljšali s prvo metodo, zmanjšali na  $1/n$ , kjer je  $n > 3$ , bi prišlo do zmanjšanja števil poslabšanj in povečanja števila izboljšanj. Ti dve metodi bi se z večanjem  $n$ -ja namreč vse bolj obnašali kot prva metoda.

Pri tretji metodi smo zaradi lastnosti ocene zanesljivosti pričakovali potencialno izboljšanje glede na drugo metodo pri večrazrednih domenah. Rezultati kažejo, da sta metodi približno enako uspešni.

Med uspešnostjo metod prenosa znanja in številom razredov, diskretnimi, zveznimi atributi, ni nobene korelacije. Ostaja vprašanje, zakaj imata drugi dve metodi več izboljšanj in tudi več poslabšanj. Predvidevamo, da zato, ker v nekaterih domenah lahko brez večjih posledic odstranimo primere, ki so bili klasificirani z majhno verjetnostjo ali oceno zanesljivosti razreda. Pri drugih je zaradi kompleksnosti domene in majhne učne množice filtriranje veliko bolj vplivalo na modele.

## Poglavje 6

### Sklepne ugotovitve

#### 6.1. Dosežene izboljšave

*Uporabljene metode prenosa znanja lahko uporabimo za izboljšanje klasifikacijske točnosti preprostejših modelov, ki se v osnovi ne obnesejo najboljše. Ker so ti modeli (naivni Bayes, kNN, odločitveno drevo) velikokrat lažje razumljivi oziroma podajajo tudi razlago svojih odločitev, v nasprotju z bolj kompleksnimi modeli (SVM, ANN), je torej prenos znanja uporaben.*

*Prenos znanja z metodama, ki uporabljata filtriranje učne množice, je uporaben, ko želimo pospešiti učenje ANN-ja. Potrebno pa je najprej naučiti model, iz katerega se znanje nato prenese v ANN. Če imamo že zgrajen tak model, ga lahko na nekaterih domenah uporabimo za prenos znanja v ANN s filtriranjem. Ob tem se bo ANN zgradil za velikost faktorja filtriranja hitreje. Paziti moramo, da ta faktor ni prevelik, saj tako izgubimo tudi podatke, ki vsebujejo veliko informacij o domeni, in pri tem klasifikacijska točnost ANN-ja pade. Ne smemo seveda zanemariti tudi časa napovedovanja modela-učitelja.*

*Izmed vseh testiranih metod se največkrat nedvomno izboljšša klasifikacijska točnost naivnega Bayesa. Poleg izboljšanj je pri prenosu znanja prišlo do veliko poslabšanj klasifikacijske točnosti. Dejstvo je, da preden lahko uporabimo modele, zgrajene s prenosom znanja, je le-te potrebno najprej primerjati z osnovnimi modeli.*

#### 6.2. Odprte teme na področju prenosa znanja

*Pri razvitih metodah bi bilo potrebno izboljšati še nastavljanje parametra, ki določa število primerov, ki se jih odstrani iz učne množice. Metoda za to je lahko npr. testiranje pri več vrednostih tega parametra in nato izbira najboljše vrednosti.*

*Prenos znanja, pri katerem učna množica ni več na razpolago, je tema, ki je z razvitimi metodami nismo zajeli. Potrebno bi bilo razviti in testirati metodo, ki deluje na principu črne škatle in v kombinaciji z aktivnim učenjem prenaša znanje. Aktivno učenje bi tu delovalo tako, da bi model-učenec na podlagi tega, kje v domeni še ni zadosti uspešen, zahteval od modela-učitelja označitev novih primerov.*

*Še ena tema, ki bi jo bilo vredno raziskati, je prenos znanja z dvosmerno komunikacijo med modelom-učiteljem in učencem, ki temelji na EM algoritmu. To smo opisali v poglavju 2 (stran 6).*

## Dodatek A

### Rezultati testiranj

*Na naslednjih straneh predstavljeni rezultati testiranj se interpretirajo:*

- *prva tabela pri vsaki domeni označuje klasifikacijske točnosti navedenih modelov strojnega učenja brez prenosa znanja,*
- *v preostalih treh tabelah iste domene je klasifikacijska točnost modela učenca Y, po prenosu znanja iz modela X v model Y na mestu [X,Y] (vrstica, stolpec),*
- *Ob doseženih klasifikacijskih točnostih pri vseh tabelah prenosa znanja je podano ali se klasifikacijska točnost izboljša (+), poslabša (-) ali pa se ta ne spremeni (o).*



## Abalone

	SVM	ANN	naivni Bayes	kNN	odl. drevo
abalone	0,2690	0,2509	0,2359	0,2313	0,2457

**Tabela 7.1: Klasifikacijska točnost modelov-učiteljev na domeni "abalone".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,2715/o	0,0089/-	0,2630/+	0,2679/+	0,2667/+
ANN	0,1471/-	0,1456/-	0,1312/-	0,1476/-	0,1407/-
naivni Bayes	0,2392/-	0,0494/-	0,2335/o	0,2338/o	0,2282/o
kNN	0,2713/o	0,0520/-	0,2530/+	0,2418/o	0,2512/o
odl. drevo	0,2444/-	0,0002/-	0,2336/o	0,2414/o	0,2457/o

**Tabela 7.2: Klasifikacijska točnost modelov-učencev na domeni "abalone" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,2721/o	0,0125/-	0,2684/+	0,2707/+	0,2667/+
ANN	0,1657/-	0,1584/-	0,1682/-	0,1693/-	0,1704/-
naivni Bayes	0,2392/-	0,0489/-	0,2335/o	0,2336/o	0,2282/o
kNN	0,2731/o	0,0626/-	0,2650/+	0,2447/o	0,2548/o
odl. drevo	0,2392/-	0,0489/-	0,2335/o	0,2336/o	0,2282/o

**Tabela 7.3: Klasifikacijska točnost modelov-učencev na domeni "abalone" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,2679/o	0,0103/-	0,2614/+	0,2651/+	0,2636/o
ANN	0,1731/-	0,1675/-	0,1373/-	0,1606/-	0,1657/-
naivni Bayes	0,2392/-	0,0493/-	0,2335/o	0,2334/o	0,2282/o
kNN	0,2696/o	0,0501/-	0,2508/+	0,2354/o	0,2615/o
odl. drevo	0,2392/-	0,0493/-	0,2335/o	0,2334/o	0,2282/o

**Tabela 7.4: Klasifikacijska točnost modelov-učencev na domeni "abalone" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## Iris

	SVM	ANN	naivni Bayes	kNN	odl. drevo
iris	0,9586	0,9799	0,9532	0,9590	0,9283

**Tabela 7.5: Klasifikacijska točnost modelov-učiteljev na domeni "iris".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9586/o	0,9669/o	0,9532/o	0,9722/o	0,9478/o
ANN	0,9598/o	0,9682/o	0,9468/o	0,9667/o	0,9553/o
naivni Bayes	0,9586/o	0,9736/o	0,9473/o	0,9653/o	0,9652/+
kNN	0,9531/o	0,9682/o	0,9473/o	0,9667/o	0,9419/o
odl. drevo	0,9595/o	0,9661/o	0,9532/o	0,9531/o	0,9283/o

**Tabela 7.6: Klasifikacijska točnost modelov-učencev na domeni "iris" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9569/o	0,9799/o	0,9532/o	0,9552/o	0,9560/o
ANN	0,9649/o	0,9741/o	0,9460/o	0,9729/o	0,9207/o
naivni Bayes	0,8910/-	0,9741/o	0,9410/o	0,9413/o	0,9560/o
kNN	0,9653/o	0,9799/o	0,9591/o	0,9487/o	0,9309/o
odl. drevo	0,9657/o	0,9661/o	0,9532/o	0,8889/o	0,9346/o

**Tabela 7.7: Klasifikacijska točnost modelov-učencev na domeni "iris" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9507/o	0,9741/o	0,9473/o	0,9489/o	0,9560/o
ANN	0,9649/o	0,9741/o	0,9460/o	0,9729/o	0,9207/o
naivni Bayes	0,8981/-	0,9674/o	0,9116/o	0,9413/o	0,9560/o
kNN	0,9653/o	0,9799/o	0,9591/o	0,9487/o	0,9059/o
odl. drevo	0,9657/o	0,9661/o	0,9532/o	0,9178/o	0,8846/o

**Tabela 7.8: Klasifikacijska točnost modelov-učencev na domeni "iris" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## Glass

	SVM	ANN	naivni Bayes	kNN	odl. drevo
glass	0,7344	0,6652	0,4090	0,6577	0,6665

**Tabela 7.9: Klasifikacijska točnost modelov-učiteljev na domeni "glass".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,6941/o	0,5250/-	0,4592/+	0,6411/o	0,6254/o
ANN	0,5264/-	0,5319/-	0,3752/o	0,5169/-	0,4880/-
naivni Bayes	0,4039/-	0,4038/-	0,3101/-	0,3835/-	0,3992/-
kNN	0,6298/-	0,6375/o	0,5148/+	0,6334/o	0,6026/o
odl. drevo	0,6825/o	0,5696/-	0,4369/o	0,6346/o	0,6539/o

**Tabela 7.10: Klasifikacijska točnost modelov-učencev na domeni "glass" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,6680/-	0,5240/-	0,5018/+	0,6365/o	0,6351/o
ANN	0,4897/-	0,5107/-	0,4051/o	0,5168/-	0,5019/-
naivni Bayes	0,4390/-	0,4270/-	0,3618/o	0,4025/-	0,4267/-
kNN	0,6284/-	0,5900/-	0,5006/+	0,6262/o	0,5791/-
odl. drevo	0,6853/-	0,5382/-	0,4389/o	0,5631/-	0,6554/o

**Tabela 7.11: Klasifikacijska točnost modelov-učencev na domeni "glass" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,6677/-	0,5196/-	0,4987/+	0,6466/o	0,6475/o
ANN	0,5251/-	0,5388/-	0,4258/o	0,5140/-	0,4865/-
naivni Bayes	0,4390/-	0,4220/-	0,3618/o	0,3951/-	0,4313/-
kNN	0,6147/-	0,5874/-	0,5548/+	0,6196/o	0,5831/-
odl. drevo	0,6853/-	0,5458/-	0,4178/o	0,5978/o	0,6568/o

**Tabela 7.12: Klasifikacijska točnost modelov-učencev na domeni "glass" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

**wdbc**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
wdbc	0,9822	0,9527	0,3769	0,9346	0,9264

**Tabela 7.13: Klasifikacijska točnost modelov-učiteljev na domeni "wdbc".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9751/-	0,9501/o	0,3769/o	0,9241/o	0,9225/o
ANN	0,9737/-	0,9446/o	0,3751/o	0,9387/o	0,9164/o
naivni Bayes	0,3728/-	0,3728/-	0,5958/+	0,3728/-	0,3743/-
kNN	0,9659/-	0,9259/-	0,4859/+	0,9296/o	0,9286/o
odl. drevo	0,9637/o	0,9451/o	0,3763/o	0,9240/o	0,9233/o

**Tabela 7.14: Klasifikacijska točnost modelov-učencev na domeni "wdbc" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9532/-	0,9231/o	0,3765/o	0,9229/o	0,9120/o
ANN	0,9746/o	0,9173/-	0,4618/+	0,9192/o	0,9087/o
naivni Bayes	0,3728/-	0,3714/-	0,3739/o	0,3728/-	0,3758/-
kNN	0,9673/-	0,9228/o	0,5390/+	0,9237/o	0,9101/o
odl. drevo	0,9617/-	0,9242/o	0,3832/o	0,9241/o	0,9034/o

**Tabela 7.15: Klasifikacijska točnost modelov-učencev na domeni "wdbc" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9532/-	0,9247/o	0,3749/o	0,9229/o	0,9135/o
ANN	0,9694/o	0,9202/-	0,4598/+	0,9140/o	0,9059/o
naivni Bayes	0,3728/-	0,3714/-	0,4872/+	0,3728/-	0,3758/-
kNN	0,9673/-	0,9246/-	0,5390/+	0,9237/o	0,9101/o
odl. drevo	0,9617/-	0,9262/-	0,3832/o	0,9241/o	0,9034/o

**Tabela 7.16: Klasifikacijska točnost modelov-učencev na domeni "wdbc" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## pima-indians diabetes

	SVM	ANN	naivni Bayes	kNN	odl. drevo
pima-i. diabetes	0,7572	0,7336	0,7486	0,7189	0,7448

**Tabela 7.17: Klasifikacijska točnost modelov-učiteljev na domeni "pima-indians diabetes".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7587/o	0,7531/o	0,7603/o	0,7515/+	0,7469/o
ANN	0,7594/o	0,7392/o	0,7538/o	0,7442/o	0,7374/o
naivni Bayes	0,7566/o	0,7569/o	0,7351/-	0,7505/+	0,7484/o
kNN	0,7566/o	0,7619/o	0,7538/o	0,7331/o	0,7562/o
odl. drevo	0,7599/o	0,7539/o	0,7526/o	0,7399/o	0,7361/o

**Tabela 7.18: Klasifikacijska točnost modelov-učencev na domeni "pima-indians diabetes" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7586/o	0,7371/o	0,7527/o	0,7469/o	0,7125/o
ANN	0,7656/o	0,7192/o	0,7442/o	0,7336/o	0,7265/o
naivni Bayes	0,7265/o	0,7230/o	0,7131/-	0,7265/o	0,7350/o
kNN	0,7548/o	0,7353/o	0,7513/o	0,7433/o	0,7242/o
odl. drevo	0,7450/o	0,7338/o	0,7397/o	0,7266/o	0,7095/-

**Tabela 7.19: Klasifikacijska točnost modelov-učencev na domeni "pima-indians diabetes" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7592/o	0,7498/o	0,7540/o	0,7483/o	0,7158/o
ANN	0,7553/o	0,7224/o	0,7529/o	0,7360/o	0,7309/o
naivni Bayes	0,7265/o	0,7187/o	0,7131/-	0,7265/o	0,7350/o
kNN	0,7548/o	0,7358/o	0,7513/o	0,7433/o	0,7242/o
odl. drevo	0,7437/o	0,7287/o	0,7344/o	0,7266/o	0,7095/-

**Tabela 7.20: Klasifikacijska točnost modelov-učencev na domeni "pima-indians diabetes" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

**ZOO**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
zoo	0,9486	0,8337	0,9152	0,8423	0,8009

**Tabela 7.21: Klasifikacijska točnost modelov-učiteljev na domeni "zoo".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9486/o	0,8548/o	0,9152/o	0,8600/o	0,8009/o
ANN	0,6689/-	0,6669/-	0,6240/-	0,6290/-	0,7274/o
naivni Bayes	0,9139/o	0,8423/o	0,9048/o	0,8535/o	0,8389/o
kNN	0,8469/-	0,6099/-	0,8212/o	0,8221/o	0,7527/o
odl. drevo	0,7925/-	0,6965/-	0,7852/-	0,7559/-	0,7841/o

**Tabela 7.22: Klasifikacijska točnost modelov-učencev na domeni "zoo" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7527/-	0,8155/-	0,9316/o	0,8130/o	0,7784/o
ANN	0,6502/-	0,6509/-	0,6620/-	0,6022/-	0,6680/o
naivni Bayes	0,8685/o	0,6419/-	0,8456/-	0,7327/-	0,7241/o
kNN	0,7920/-	0,5750/-	0,8730/o	0,7322/-	0,7516/o
odl. drevo	0,6789/-	0,6510/-	0,7168/-	0,7277/-	0,7790/o

**Tabela 7.23: Klasifikacijska točnost modelov-učencev na domeni "zoo" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,8111/-	0,8165/o	0,9320/o	0,8054/o	0,7558/o
ANN	0,6520/-	0,6620/-	0,6763/-	0,6324/-	0,6823/o
naivni Bayes	0,8154/-	0,7654/-	0,8435/o	0,8074/o	0,7729/o
kNN	0,8017/-	0,5994/-	0,8476/o	0,7595/-	0,7375/o
odl. drevo	0,7384/-	0,6788/-	0,7372/-	0,6861/-	0,7201/o

**Tabela 7.24: Klasifikacijska točnost modelov-učencev na domeni "zoo" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## parkinsons

	SVM	ANN	naivni Bayes	kNN	odl. drevo
parkinsons	0,8848	0,8544	0,6980	0,8707	0,8395

**Tabela 7.25: Klasifikacijska točnost modelov-učiteljev na domeni "parkinsons".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,8744/o	0,8393/o	0,7632/o	0,8420/o	0,8675/o
ANN	0,8707/o	0,8497/o	0,7310/o	0,8459/o	0,8560/o
naivni Bayes	0,6814/-	0,6997/-	0,6753/o	0,6537/-	0,7076/-
kNN	0,8598/-	0,8405/o	0,6878/o	0,8744/o	0,8723/o
odl. drevo	0,8602/-	0,8400/o	0,6961/o	0,8171/-	0,7905/o

**Tabela 7.26: Klasifikacijska točnost modelov-učencev na domeni "parkinsons" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,8694/o	0,8317/-	0,7843/+	0,8398/o	0,8677/o
ANN	0,8731/o	0,8166/-	0,7845/+	0,8356/o	0,8423/o
naivni Bayes	0,7024/-	0,7365/-	0,7044/o	0,7163/-	0,6777/-
kNN	0,8598/-	0,8344/-	0,7729/o	0,8149/-	0,8244/o
odl. drevo	0,8074/-	0,8621/o	0,7582/+	0,8230/o	0,8374/o

**Tabela 7.27: Klasifikacijska točnost modelov-učencev na domeni "parkinsons" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,8694/o	0,8417/o	0,7806/+	0,8398/o	0,8677/o
ANN	0,8709/o	0,8253/o	0,7848/+	0,8465/o	0,8531/o
naivni Bayes	0,6987/-	0,7179/-	0,6994/o	0,7204/-	0,6777/-
kNN	0,8598/-	0,8325/o	0,7722/o	0,8149/-	0,8287/o
odl. drevo	0,8368/-	0,8580/o	0,7804/+	0,8267/o	0,8315/o

**Tabela 7.28: Klasifikacijska točnost modelov-učencev na domeni "parkinsons" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## ionosphere

	SVM	ANN	naivni Bayes	kNN	odl. drevo
ionosphere	0,9385	0,8881	0,8882	0,8470	0,8699

**Tabela 7.29: Klasifikacijska točnost modelov-učiteljev na domeni "ionosphere".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9327/o	0,8798/o	0,8860/o	0,8492/o	0,8888/o
ANN	0,9356/o	0,8791/o	0,8752/o	0,8418/o	0,8872/o
naivni Bayes	0,9215/o	0,8670/o	0,7984/-	0,7885/-	0,8774/o
kNN	0,8941/-	0,8487/-	0,8080/-	0,8256/-	0,8452/o
odl. drevo	0,9193/o	0,8803/o	0,8705/o	0,8302/o	0,8771/o

**Tabela 7.30: Klasifikacijska točnost modelov-učencev na domeni "ionosphere" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9378/o	0,8778/o	0,8328/-	0,8579/o	0,8991/+
ANN	0,9346/o	0,8602/o	0,8915/o	0,8703/o	0,8598/o
naivni Bayes	0,7874/-	0,7652/-	0,7453/-	0,7122/-	0,7571/-
kNN	0,9273/o	0,8303/-	0,8936/o	0,8437/o	0,8537/o
odl. drevo	0,9305/o	0,8678/o	0,8655/o	0,8330/o	0,8921/o

**Tabela 7.31: Klasifikacijska točnost modelov-učencev na domeni "ionosphere" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9351/o	0,8681/o	0,8328/-	0,8605/o	0,9050/+
ANN	0,9323/o	0,8609/o	0,8941/o	0,8703/o	0,8473/o
naivni Bayes	0,7958/-	0,7540/-	0,7397/-	0,7072/-	0,7580/-
kNN	0,9273/o	0,8356/-	0,8903/o	0,8437/o	0,8666/o
odl. drevo	0,9253/o	0,8652/o	0,8786/o	0,8120/o	0,8895/o

**Tabela 7.32: Klasifikacijska točnost modelov-učencev na domeni "ionosphere" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**



## house-votes

	SVM	ANN	naivni Bayes	kNN	odl. drevo
house-votes	0,9562	0,9590	0,9031	0,9359	0,9537

**Tabela 7.33: Klasifikacijska točnost modelov-učiteljev na domeni "house-votes".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9562/o	0,9563/o	0,9152/o	0,9228/o	0,9562/o
ANN	0,9534/o	0,9582/o	0,9059/o	0,9339/o	0,9556/o
naivni Bayes	0,9057/-	0,9037/-	0,8849/o	0,9046/-	0,9137/-
kNN	0,9401/o	0,9288/-	0,8895/o	0,9112/-	0,9447/o
odl. drevo	0,9562/o	0,9537/o	0,9145/o	0,9261/o	0,9537/o

**Tabela 7.34: Klasifikacijska točnost modelov-učencev na domeni "house-votes" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9363/o	0,9172/-	0,9110/o	0,9103/-	0,9562/o
ANN	0,9685/o	0,9462/o	0,9174/+	0,9200/o	0,9562/o
naivni Bayes	0,8939/-	0,8782/-	0,8788/o	0,8803/-	0,9562/o
kNN	0,9338/o	0,9257/-	0,9099/o	0,9055/-	0,9562/o
odl. drevo	0,9562/o	0,9562/o	0,9217/+	0,9260/o	0,9562/o

**Tabela 7.35: Klasifikacijska točnost modelov-učencev na domeni "house-votes" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,9363/o	0,9209/-	0,9082/o	0,9103/o	0,9562/o
ANN	0,9685/o	0,9462/o	0,9174/+	0,9200/o	0,9562/o
naivni Bayes	0,8967/-	0,8820/-	0,8820/o	0,8803/-	0,9562/o
kNN	0,9338/o	0,9257/-	0,9099/o	0,9055/-	0,9562/o
odl. drevo	0,9562/o	0,9562/o	0,9217/+	0,9279/o	0,9562/o

**Tabela 7.36: Klasifikacijska točnost modelov-učencev na domeni "house-votes" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## digits

	SVM	ANN	naivni Bayes	kNN	odl. drevo
digits	0,7331	0,9836	0,8324	0,9869	0,7750

**Tabela 7.37: Klasifikacijska točnost modelov-učiteljev na domeni "digits".**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7331/o	0,9835/o	0,8324/o	0,9870/o	0,7750/o
ANN	0,7331/o	0,9835/o	0,8324/o	0,9875/o	0,7750/o
naivni Bayes	0,2444/-	0,8463/-	0,8008/-	0,8954/-	0,7195/-
kNN	0,7387/o	0,9810/o	0,8349/o	0,9862/o	0,7626/o
odl. drevo	0,4209/-	0,7927/-	0,7480/-	0,8747/-	0,7734/o

**Tabela 7.38: Klasifikacijska točnost modelov-učencev na domeni "digits" po prenosu znanja z metodo ponovne označitve učne množice.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7331/o	0,9833/o	0,8324/o	0,9877/o	0,7750/o
ANN	0,6189/-	0,9590/-	0,8720/+	0,9723/-	0,7433/-
naivni Bayes	0,3170/-	0,8517/-	0,8168/-	0,9088/-	0,7086/-
kNN	0,7299/o	0,9788/-	0,8455/+	0,9839/-	0,7639/o
odl. drevo	0,3705/-	0,7687/-	0,7786/-	0,8570/-	0,7101/-

**Tabela 7.39: Klasifikacijska točnost modelov-učencev na domeni "digits" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.**

	SVM	ANN	naivni Bayes	kNN	odl. drevo
SVM	0,7331/o	0,9849/o	0,8324/o	0,9871/o	0,7750/o
ANN	0,6165/-	0,9604/-	0,8724/+	0,9725/-	0,7465/-
naivni Bayes	0,3176/-	0,8512/-	0,8169/-	0,9097/-	0,7105/-
kNN	0,7299/o	0,9802/o	0,8455/+	0,9841/-	0,7639/o
odl. drevo	0,3665/-	0,7680/-	0,7785/-	0,8581/-	0,7090/-

**Tabela 7.40: Klasifikacijska točnost modelov-učencev na domeni "digits" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.**

## Seznam slik

<i>Slika 1.1: Atributna predstavitev primerov v poljubni domeni. ....</i>	<i>3</i>
<i>Slika 1.2: Pri prenosu znanja lahko na modele gledamo kot na črne škatle, podrobnosti implementacije niso pomembne. Pomembna sta le izhod in vhod v model. ....</i>	<i>3</i>
<i>Slika 2.1: Osnovni potek prenosa znanja. ....</i>	<i>5</i>
<i>Slika 2.2: Algoritem kooperativnega učenja ....</i>	<i>9</i>
<i>Slika 2.3: Transdukcija je način sklepanja metode kNN. ....</i>	<i>10</i>
<i>Slika 2.4: Umetna nevronska mreža z eno skrito plastjo. ....</i>	<i>11</i>
<i>Slika 2.5: Gradnja odločitvenega drevesa [6]. ....</i>	<i>12</i>
<i>Slika 2.6: Ločevalna hiperravnina pri metodi podpornih vektorjev ....</i>	<i>13</i>

## Seznam tabel

<i>Tabela 4.1: Povzetek vseh domen.....</i>	<i>18</i>
<i>Tabela 5.1: Povzetek rezultatov testiranj prenosa znanja s ponovno označitvijo učne množice. (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni.....</i>	<i>19</i>
<i>Tabela 5.2: Povzetek rezultatov testiranj prenosa znanja s filtriranjem po verjetnosti napovedi. (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni.....</i>	<i>19</i>
<i>Tabela 5.3: Povzetek rezultatov testiranj prenosa znanja s filtriranjem po zanesljivosti napovedi. (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni. ....</i>	<i>20</i>
<i>Tabela 5.4: Povzetek rezultatov testiranj treh metod na desetih domenah. (+) označuje število izboljšanja klasifikacijske točnosti, (-) poslabšanje klasifikacijske točnosti, (o) klasifikacijska točnost se ne spremeni. Metoda 1 je prenos znanja s ponovno označitvijo učne množice. Metoda 2 je prenos znanja s filtriranjem po verjetnosti napovedi. Metoda 3 je prenos znanja s filtriranjem po zanesljivosti napovedi. Vrstica "Vsota" vsebuje števila poslabšanj (-), izboljšanj (+) in nespremenjenih (o) klasifikacijskih točnosti po vseh metodah. Vrstica "Delež v vseh metodah" vsebuje deleže poslabšanj (-), izboljšanj (+) in nespremenjenih (o) klasifikacijskih točnosti glede na vseh 750 testov prenosa znanja prek vseh treh metod. Vrstica "Delež v metodi" vsebuje deleže poslabšanj (-), izboljšanj (+) in nespremenjenih (o) klasifikacijskih točnosti glede na posamezno metodo, ki obsega 250 testov.....</i>	<i>20</i>
<i>Tabela 7.1: Klasifikacijska točnost modelov-učiteljev na domeni "abalone".....</i>	<i>24</i>
<i>Tabela 7.2: Klasifikacijska točnost modelov-učencev na domeni "abalone" po prenosu znanja z metodo ponovne označitve učne množice.....</i>	<i>24</i>
<i>Tabela 7.3: Klasifikacijska točnost modelov-učencev na domeni "abalone" po prenosu znanja z metodo filtriranja po verjetnosti napovedi. ....</i>	<i>24</i>
<i>Tabela 7.4: Klasifikacijska točnost modelov-učencev na domeni "abalone" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.....</i>	<i>24</i>
<i>Tabela 7.5: Klasifikacijska točnost modelov-učiteljev na domeni "iris". ....</i>	<i>25</i>
<i>Tabela 7.6: Klasifikacijska točnost modelov-učencev na domeni "iris" po prenosu znanja z metodo ponovne označitve učne množice. ....</i>	<i>25</i>
<i>Tabela 7.7: Klasifikacijska točnost modelov-učencev na domeni "iris" po prenosu znanja z metodo filtriranja po verjetnosti napovedi. ....</i>	<i>25</i>
<i>Tabela 7.8: Klasifikacijska točnost modelov-učencev na domeni "iris" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi. ....</i>	<i>25</i>
<i>Tabela 7.9: Klasifikacijska točnost modelov-učiteljev na domeni "glass".....</i>	<i>26</i>
<i>Tabela 7.10: Klasifikacijska točnost modelov-učencev na domeni "glass" po prenosu znanja z metodo ponovne označitve učne množice. ....</i>	<i>26</i>
<i>Tabela 7.11: Klasifikacijska točnost modelov-učencev na domeni "glass" po prenosu znanja z metodo filtriranja po verjetnosti napovedi. ....</i>	<i>26</i>
<i>Tabela 7.12: Klasifikacijska točnost modelov-učencev na domeni "glass" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi. ....</i>	<i>26</i>
<i>Tabela 7.13: Klasifikacijska točnost modelov-učiteljev na domeni "wdbc". ....</i>	<i>27</i>
<i>Tabela 7.14: Klasifikacijska točnost modelov-učencev na domeni "wdbc" po prenosu znanja z metodo ponovne označitve učne množice. ....</i>	<i>27</i>
<i>Tabela 7.15: Klasifikacijska točnost modelov-učencev na domeni "wdbc" po prenosu znanja z metodo filtriranja po verjetnosti napovedi. ....</i>	<i>27</i>

<i>Tabela 7.16: Klasifikacijska točnost modelov-učencev na domeni "wdbc" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	27
<i>Tabela 7.17: Klasifikacijska točnost modelov-učiteljev na domeni "pima-indians diabetes".</i>	28
<i>Tabela 7.18: Klasifikacijska točnost modelov-učencev na domeni "pima-indians diabetes" po prenosu znanja z metodo ponovne označitve učne množice.</i>	28
<i>Tabela 7.19: Klasifikacijska točnost modelov-učencev na domeni "pima-indians diabetes" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.</i>	28
<i>Tabela 7.20: Klasifikacijska točnost modelov-učencev na domeni "pima-indians diabetes" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	28
<i>Tabela 7.21: Klasifikacijska točnost modelov-učiteljev na domeni "zoo".</i>	29
<i>Tabela 7.22: Klasifikacijska točnost modelov-učencev na domeni "zoo" po prenosu znanja z metodo ponovne označitve učne množice.</i>	29
<i>Tabela 7.23: Klasifikacijska točnost modelov-učencev na domeni "zoo" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.</i>	29
<i>Tabela 7.24: Klasifikacijska točnost modelov-učencev na domeni "zoo" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	29
<i>Tabela 7.25: Klasifikacijska točnost modelov-učiteljev na domeni "parkinsons".</i>	30
<i>Tabela 7.26: Klasifikacijska točnost modelov-učencev na domeni "parkinsons" po prenosu znanja z metodo ponovne označitve učne množice.</i>	30
<i>Tabela 7.27: Klasifikacijska točnost modelov-učencev na domeni "parkinsons" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.</i>	30
<i>Tabela 7.28: Klasifikacijska točnost modelov-učencev na domeni "parkinsons" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	30
<i>Tabela 7.29: Klasifikacijska točnost modelov-učiteljev na domeni "ionosphere".</i>	31
<i>Tabela 7.30: Klasifikacijska točnost modelov-učencev na domeni "ionosphere" po prenosu znanja z metodo ponovne označitve učne množice.</i>	31
<i>Tabela 7.31: Klasifikacijska točnost modelov-učencev na domeni "ionosphere" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.</i>	31
<i>Tabela 7.32: Klasifikacijska točnost modelov-učencev na domeni "ionosphere" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	31
<i>Tabela 7.33: Klasifikacijska točnost modelov-učiteljev na domeni "house-votes".</i>	32
<i>Tabela 7.34: Klasifikacijska točnost modelov-učencev na domeni "house-votes" po prenosu znanja z metodo ponovne označitve učne množice.</i>	32
<i>Tabela 7.35: Klasifikacijska točnost modelov-učencev na domeni "house-votes" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.</i>	32
<i>Tabela 7.36: Klasifikacijska točnost modelov-učencev na domeni "house-votes" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	32
<i>Tabela 7.37: Klasifikacijska točnost modelov-učiteljev na domeni "digits".</i>	33
<i>Tabela 7.38: Klasifikacijska točnost modelov-učencev na domeni "digits" po prenosu znanja z metodo ponovne označitve učne množice.</i>	33
<i>Tabela 7.39: Klasifikacijska točnost modelov-učencev na domeni "digits" po prenosu znanja z metodo filtriranja po verjetnosti napovedi.</i>	33
<i>Tabela 7.40: Klasifikacijska točnost modelov-učencev na domeni "digits" po prenosu znanja z metodo filtriranja po zanesljivosti napovedi.</i>	33

## Literatura

- [1] A. Asuncion, D.J. Newman, *UCI Machine Learning Repository* (2010). Dostopno na: <http://archive.ics.uci.edu/ml/>
- [2] A. Blum, T. Mitchell, "Combining Labeled and unlabeled Data with Co-Training", v zborniku *Proceedings of the 1998 Conference on Computational Learning Theory*, 1998.
- [3] D.A. Cohn, M.I. Jordan, Z. Ghahramani, "Active Learning with Statistical Models", *Massachusetts Institute of Technology in Center for Biological and Computational Learning*, 1995.
- [4] D. MacKay, "Information-based objective functions", v *Neural Computation* 4, 1992), str. 590-604.
- [5] D. Meyer, *Support Vector Machines*, Technische Universität Wien, Wien, 2010.
- [6] E.B. Hunt, J. Marin, P.J. Stone, "Experiments in Induction", *Academic Press*, New York, 1966.
- [7] E. Štrumbelj, I. Kononenko, D. Pevec, "Evaluating reliability of single classifications" (članek v pripravi).
- [8] Igor Kononenko, *Strojno učenje*, Fakulteta za računalništvo in informatiko, Ljubljana, 2005, str 106-107, 115.
- [9] M. Možina, J. Demšar, M. Kattan, B. Zupan, "Nomograms for Visualization of Naive Bayesian Classifier", v zborniku *Knowledge Discovery in Databases: PKDD 2004*, Springer Berlin/Heidelberg, 2004.
- [10] S. Baluja, "Probabilistic Modeling for Face Orientation Discrimination: Learning from Labeled and Unlabeled Data", *Justsystem Pittsburgh Research Center & School of Computer Science, Carnegie Mellon University*.
- [11] (2010) CRAN Contributed Packages. Dostopno na: <http://cran.r-project.org/web/packages/>