

Univerza v Ljubljani
Fakulteta za računalništvo in informatiko

MAGISTRSKO DELO

**Izboljšave metode skladanja
klasifikatorjev**

Bernard Ženko

Mentor: prof. dr. Ivan Bratko
Somentor: doc. dr. Sašo Džeroski

Ljubljana, 2003

Zahvala

Institutskemu mentorju doc. dr. Sašu Džeroskemu se zahvaljujem, da me je sprejel v svoje delovno okolje ter uvedel v področje strojnega učenja. Njegove ideje, podpora in potrpežljivost so bile ključne za nastanek tega magistrskega dela.

Mentorju prof. dr. Ivanu Bratku se zahvaljujem za koristne komentarje o delu, predstavljenem v tem magisteriju.

Mag. Ljupču Todorovskemu se zahvaljujem za strokovno in tehnično pomoč ter množico koristnih nasvetov.

Doc. dr. Ernestu Ženku se zahvaljujem za jezikovni pregled besedila.

Vsem sodelavkam in sodelavcem Odseka za inteligentne sisteme Instituta Jožef Stefan se zahvaljujem za pomoč ter stimulatивно in prijetno delovno okolje.

Zahvaljujem se Ministrstvu za šolstvo, znanost in šport, da je mi zagotovilo sredstva za podiplomski študij ter Institutu Jožef Stefan, ki mi je zagotovil primerno delovno okolje.

Nenazadnje se zahvaljujem staršema za potrpežljivost in neomajno podporo.

Kazalo

Zahvala	i
Povzetek	iv
Abstract	v
Ključne besede	vi
1 Uvod	1
1.1 Ozadje in motivacija	1
1.2 Znanstveni prispevek	2
1.3 Organizacija magistrskega dela	3
2 Kombiniranje klasifikatorjev	4
2.1 Ansambli klasifikatorjev	4
2.2 Metaučenje	5
2.3 Najbolj znane metode za kombiniranje	6
2.3.1 Metoda bagging	6
2.3.2 Metoda boosting	7
2.3.3 Naključni gozdovi	8
2.3.4 Glasovanje	8
2.3.5 Izberi najboljšega	8
2.3.6 Metoda cascading	9
2.3.7 Skladanje	9

3	Primerjava obstoječih metod za kombiniranje	14
3.1	Metodologija vrednotenja	14
3.1.1	Klasifikacijska napaka	16
3.1.2	Relativno izboljšanje	16
3.1.3	Statistična signifikantnost	16
3.2	Uporabljene metode	18
3.3	Rezultati primerjave	20
4	Izboljšano skladanje klasifikatorjev	24
4.1	Razširitev nabora metaatributov	24
4.2	Skladanje z večličnimi modelskimi drevesi	26
4.3	Primerjava z obstoječimi metodami	26
4.4	Komentar rezultatov	30
4.5	Uporaba metod na praktičnih problemih	31
4.5.1	Množice podatkov	31
4.5.2	Rezultati	35
5	Sklep	37
	Literatura	39
	Izjava	44
	Priloge	45
A.	Rezultati poskusov z obstoječimi metodami	46
B.	Rezultati poskusov z izboljšanimi metodami	54
C.	Rezultati poskusov na praktičnih problemih	59

Povzetek

Na področju induktivnega strojnega učenja so metode za kombiniranje klasifikatorjev deležne precejšnje pozornosti. Načeloma nam omogočajo doseganje boljše klasifikacijske točnosti, saj lahko v primeru, ko nam določen osnovni klasifikator vrne napačno napoved, le-to popravimo z uporabo napovedi ostalih osnovnih klasifikatorjev. Metode za kombiniranje klasifikatorjev lahko razdelimo v dve skupini. V prvi so metode, ki za gradnjo osnovnih klasifikatorjev uporabljajo en sam učni algoritem, njihovo raznolikost pa dosežemo na primer s spreminjanjem učne množice. V drugi skupini so metode, ki kombinirajo klasifikatorje zgrajene z različnimi učnimi algoritmi (heterogene klasifikatorje). Najbolj znana metoda iz te skupine je metoda skladanja klasifikatorjev, pri kateri poteka kombiniranje v dveh korakih. V prvem koraku z različnimi učnimi algoritmi na osnovni učni množici zgradimo osnovne klasifikatorje. Njihove napovedi opišemo s tako imenovanimi metaatributi in jih sestavimo v metaučno množico. V drugem koraku na tako dobljeni metaučni množici zgradimo metaklasifikator, ki združi klasifikacije osnovnih klasifikatorjev v končno klasifikacijo.

V pričujočem magistrskem delu smo s pomočjo enotne metodologije ovrednotili in primerjali večino znanih metod za kombiniranje heterogenih klasifikatorjev. Rezultati so pokazali, da je točnost najboljše metode skladanja klasifikatorjev le malenkost višja od točnosti metode izbire najboljšega osnovnega klasifikatorja s prečnim preverjanjem. Zaradi tega smo raziskali možne izboljšave metod skladanja in odkrili dve novi metodi. Prva metoda uporablja razširjen nabor metaučnih atributov, druga pa za metaučni algoritem uporablja večlična modelska drevesa. Slednja nam nudi signifikantno točnejše napovedi kot dosedaj znane metode za kombiniranje heterogenih klasifikatorjev. Omenjeni metodi smo eksperimentalno ovrednotili tudi na večih praktičnih problemih.

Abstract

Methods for combining classifiers attract considerable attention in the field of inductive machine learning. They enable us, in principle, to achieve better classification accuracy, because we can correct a false prediction of one base level classifier by taking into account predictions of other base level classifiers. We can distinguish two different groups of methods for combining classifiers. Methods of the first group only use one learning algorithm for building base level classifiers and attain their diversity by manipulating the training set. Methods of the second group combine classifiers built by different learning algorithms (heterogeneous classifiers). Among this group of methods, stacking of classifiers is most well-known. Stacking combines classifiers in two steps. In the first step several learning algorithms are used to build base level classifiers. Their predictions are then described by meta level attributes and collected into a meta level learning set. In the second step the meta level learning set is used to build a meta classifier which combines the predictions of the base classifiers into a final prediction.

In the work presented here, we have compared several known methods for combining heterogeneous classifiers using a unified evaluation methodology. The results show, that the accuracy of the best stacking approach is only slightly higher than the accuracy achieved by selection by cross-validation. This has motivated us to search for possible improvements of existing stacking approaches and we have discovered two new methods. The first one uses an extended set of meta level attributes, while the second method uses multi-response model trees as a meta level learning algorithm. The latter performs significantly better than previously known methods for combining heterogeneous classifiers. Both methods were also experimentally evaluated on several practical problems.

Ključne besede

strojno učenje	(machine learning)
klasifikacija	(classification)
ansambel klasifikatorjev	(ensemble of classifiers)
kombiniranje klasifikatorjev	(combining classifiers)
metaučenje	(meta learning)
skladanje	(stacking)

1. Uvod

1.1 Ozadje in motivacija

Klasifikacija je osnovna naloga induktivnega strojnega učenja. Postopek, ki je sposoben klasificirati nove primere, imenujemo klasifikator. Glavni kriterij po katerem ocenjujemo klasifikatorje je največkrat njihova klasifikacijska točnost. V zadnjem času so deležne precejšnje pozornosti metode za kombiniranje večih klasifikatorjev. Njihova osnovna ideja je v tem, da na nek način združijo napovedi večih posameznih (osnovnih) klasifikatorjev. To nam načeloma omogoča natančnejše napovedovanje, saj lahko v primeru, ko nam določen osnovni klasifikator da napačno napoved, le-to popravimo z uporabo napovedi ostalih osnovnih klasifikatorjev. Celoto osnovnih klasifikatorjev, skupaj s postopkom za njihovo kombinacijo, imenujemo tudi *ansambel* klasifikatorjev, osnovne klasifikatorje pa člani ansambla.

Najpreprostejša postopka za kombiniranje sta glasovanje in izbor najboljšega člana ansambla s prečnim preverjanjem. Glasovanje, ki ne zahteva nobenega učenja, poteka tako, da vsak član ansambla prispeva svoj (utežen) glas za neko napoved. Napoved z največ glasovi zmaga. Pri izboru najboljšega člana ansambla s prečnim preverjanjem na učni množici ocenimo klasifikacijsko točnost vsakega člana. Za klasifikacijo novega primera uporabljamo le člana ansambla z najvišjo točnostjo. Opazimo lahko, da v tem primeru težko govorimo o ansamblu, saj nam napovedi za nove primere daje en sam klasifikator.

Poznamo več različnih pristopov k gradnji ansamblov klasifikatorjev, vse pa lahko v grobem razdelimo na dve skupini. V prvi so metode, kjer za gradnjo članov ansambla uporabljamo en sam učni algoritem. Medsebojno raznolikost članov dosežemo s spremi-njanjem učne množice, z vključitvijo naključnosti v učni algoritem ali kako drugače. Iz

te skupine sta verjetno najbolj znani in uporabljani metodi *bagging* in *boosting*. V drugo skupino uvrščamo metode, ki za gradnjo članov ansambla uporabljajo različne učne algoritme (ansambl heterogenih klasifikatorjev). Najbolj razširjena metoda iz te skupine je metoda skladanja klasifikatorjev (ang. *stacking*). Gradnja ansambla poteka v dveh korakih. V prvem koraku z več različnimi učnimi algoritmi na osnovni učni množici zgradimo osnovne klasifikatorje. Njihove napovedi opišemo s tako imenovanimi *metaatributi* ter jih sestavimo v *metaučno množico*. V drugem koraku na tako dobljeni metaučni množici zgradimo *metaklasifikator*, ki naj bi znal smiselno sestaviti klasifikacije posameznih članov ansambla v skupno klasifikacijo ansambla. Ključnega pomena za uspešnost te metode je izbor atributov metaučne množice in izbor učnega algoritma za metaklasifikator. Za ta problem najdemo v literaturi veliko predlaganih rešitev, vendar zaradi različnih pristopov avtorjev k njihovi evalvaciji ni jasno, katera rešitev nam nudi (naj)boljšo klasifikacijsko točnost. Prav tako ni jasno ali lahko z metodo skladanja klasifikatorjev dosežemo boljšo točnost kot z izborom najboljšega člana ansambla s prečnim preverjanjem. V primeru negativnega odgovora skladanje klasifikatorjev ni smiselno zaradi večje računske zahtevnosti in slabše razumljivosti ansambla (v primerjavi z enim samim klasifikatorjem).

1.2 Znanstveni prispevek

Znanstveni prispevek pričujočega magistrskega dela je sestavljen iz dveh delov. Prvi del obsega eksperimentalno ovrednotenje in primerjavo večine znanih metod za kombiniranje klasifikatorjev izvedeno po enotni metodologiji. To nam je omogočilo vpogled v trenutno stanje na področju kombiniranja klasifikatorjev ter hkrati razkrilo možne smeri nadaljnjih raziskav. Rezultati so, nekoliko presenetljivo, pokazali, da je točnost napovedi, ki jih dobimo z najboljšo metodo skladanja klasifikatorjev,¹ le malenkostno višja od točnosti metode izbire najboljšega člana ansambla s prečnim preverjanjem. Zaradi tega smo raziskali možne izboljšave metod skladanja ter odkrili dve novi metodi, ki nam omogočata signifikantno višjo točnost napovedi v primerjavi z metodo izbora najboljšega. Prva metoda uporablja kot metaatribute, poleg verjetnostnih porazdelitev razreda, še dva nova tipa

¹To je metoda *skladanje z MLR*, ki za metaatribute upravlja verjetnostne porazdelitve razreda ter večično linearno regresijo (MLR) kot metaučni algoritem. Podrobno je opisana v poglavju [2.3.7](#).

atributov, druga metoda pa za metaučni algoritem uporablja večlična modelska drevesa. Ti dve metodi tvorita drugi del znanstvenega prispevka pričujočega magistrskega dela. Omenjeni metodi smo temeljito eksperimentalno ovrednotili tudi na večih medicinskih in ekoloških problemih ter tako dokazali njihovo praktično uporabnost.

1.3 Organizacija magistrskega dela

Preostanek magistrskega dela je organiziran na sledeč način. V naslednjem poglavju je opisano kombiniranje klasifikatorjev; po kratkem uvodu v področje sledi pregleden opis dosedaj znanih metod za kombiniranje. Tretje poglavje je posvečeno eksperimentalni primerjavi večine predstavljenih metod. Opišemo evalvacijsko metodologijo in osnovne klasifikatorje, ki jih kombiniramo; poglavje sklenemo z rezultati primerjave ter zaključki, ki iz njih izhajajo. V četrtem poglavju podrobno predstavimo novi metodi skladanja klasifikatorjev ter ju eksperimentalno primerjamo z dosedaj znanimi metodami. Novi metodi preizkusimo tudi na realnih množicah podatkov. V zadnjem poglavju predstavimo sklepe, ki sledijo iz opravljenih eksperimentov ter podamo nekaj možnih smeri za nadaljnje raziskave.

2. Kombiniranje klasifikatorjev

Večina tehnik strojnega učenja se ukvarja s problemom modeliranja podatkov. Model zgradimo na znanih primerih in ga nato uporabimo za napovedovanje neznanih. Eden izmed principov, na katerem temeljijo sistemi za strojno učenje je tudi princip večkratne razlage [25]. Ta trdi, da moramo za optimalno rešitev (model) konkretnega problema upoštevati vse hipoteze, ki so konsistentne z vhodnimi podatki. To je osnovni princip, na katerem temeljijo metode za kombiniranje klasifikatorjev. V nadaljevanju poglavja bomo predstavili področje kombiniranja klasifikatorjev in metaučnja ter podali kratek pregled najbolj znanih metod za kombiniranje klasifikatorjev.

2.1 Ansambli klasifikatorjev

Ansambel klasifikatorjev je množica klasifikatorjev, katerih napovedi na neki način (običajno z glasovanjem) združimo v eno samo napoved. Problem gradnje dobrih ansamblov klasifikatorjev je v raziskavah induktivnega strojnega učenja deležen velike pozornosti [7]. Glavni razlog leži v predpostavki, da so napovedi ansamblov pogosto bolj natančne kot napovedi klasifikatorjev, ki jih sestavljajo [7, 21].

Največ raziskav je bilo opravljenih na področju gradnje ansamblov z uporabo enega samega učnega algoritma [9]; običajno so uporabljeni algoritmi za gradnjo odločitvenih dreves ali učenje nevronske mreže. Različnost klasifikatorjev dosežemo s spreminjanjem sestave učne množice (na primer metodi boosting in bagging), s spreminjanjem atributov učne množice ali z vključitvijo naključnosti v učni algoritem. Napovedi tako dobljenih klasifikatorjev običajno združimo z navadnim ali uteženim glasovanjem.

Pri drugem pristopu za gradnjo klasifikatorjev uporabimo več različnih učnih algoritmov na eni sami učni množici (glej na primer [27]). Tudi v tem primeru lahko za

združevanje napovedi uporabimo glasovanje, vendar običajno z bolj kompleksnimi metodami dosežemo bistveno boljše rezultate. Med temi metodami je najbolj znana metoda skladanja klasifikatorjev (ang. stacked generalization, stacking) [39].

2.2 Metaučenje

Večkrat lahko zasledimo, da se v zvezi s kombiniranjem klasifikatorjev govori o tako imenovanem *metaučnju*. Sam pojem pomeni učenje o učenju, gre pa za to, da na osnovi rezultatov (osnovnega) učenja potegnemo neke splošno veljavne zaključke o (osnovnem) učenju. Na področju strojnega učenja govorimo v tej zvezi o naslednjih nalogah [36]: učenje izbiranja najprimernejšega učnega algoritma, učenje dinamičnega izbiranja primerne pristranskosti učnega algoritma (ang. dynamic bias selection) in učenje kombiniranja napovedi osnovnih klasifikatorjev. V nadaljevanju bomo na kratko opisali vsako nalogo posebej.

- *Učenje izbiranja najprimernejšega učnega algoritma.* Primernost algoritma za dano domeno ocenimo na podlagi nekega kriterija, običajno na podlagi točnosti napovedi. Glavna ideja je, da vsak osnovni učni problem (domeno) opišemo z množico metaatributov od katerih je odvisna točnost napovedi učnih algoritmov. Opisi večih domen z omenjenimi metaatributi, skupaj z natančnostmi različnih učnih algoritmov na teh domenah, sestavljajo metadomeno, nad katero poženemo metaučni algoritem. Rezultat je metaklasifikator, ki naj bi znal izbrati najboljši algoritem za novo domeno, ne da bi algoritem pred tem pognal.
- *Učenje dinamičnega izbiranja primerne pristranskosti učnega algoritma.* Skušamo zgraditi učni algoritem, ki bil sposoben spreminjati svoj prostor hipotez ter tako zagotoviti boljše ujemanje z dano domeno. Kot primer te naloge lahko omenimo problem izbire parametrov učnega algoritma. Pristranskost učnega algoritma je (med drugim) odvisna od izbire njegovih parametrov, ki jih moramo za dano domeno izbrati tako, da bodo rezultati učenja čim boljši.
- *Učenje kombiniranja napovedi osnovnih klasifikatorjev.* Primere iz dane domene opišemo z napovedmi osnovnih klasifikatorjev (ali z njihovimi lastnostmi). Ti opisi,

skupaj s pravimi razredi, tvorijo metaučno množico, nad katero poženemo metaučni algoritem. Tako dobljeni metaklasifikator naj bi znal napovedati razred novega primera na osnovi napovedi osnovnih klasifikatorjev. To je vrsta metaučenja s katero se ukvarjamo v tem magistrskem delu.

2.3 Najbolj znane metode za kombiniranje

Poznamo veliko različnih metod za kombiniranje klasifikatorjev. Glede na to, kako gradimo osnovne klasifikatorje, ki jih nato kombiniramo (člane ansambla), lahko vse metode razdelimo v dve skupini. V prvi so metode, ki za gradnjo članov ansambla uporabljajo en sam učni algoritem, v drugi pa metode, ki kombinirajo klasifikatorje zgrajene z različnimi učnimi algoritmi. V nadaljevanju bomo na kratko opisali najbolj znane predstavnike obeh skupin.

2.3.1 Metoda bagging

Metodo *bagging*¹ je predlagal Breiman [3]. Metoda izkorišča nestabilnost učnih algoritmov. Iz osnovne učne množice sestavimo več novih učnih množic, ki se od prvotne (in med seboj) nekoliko razlikujejo. Uporabimo postopek vzorčenja z zamenjavo (ang. sampling with replacement): iz osnovne učne množice naključno izbiramo primere ter z njimi sestavljamo novo učno množico, pri čemer je pomembno, da je nek primer lahko izbran večkrat. Na ta način sestavimo učno množico z enakim številom primerov kot jih ima prvotna učna množica. Naj opozorimo, da nova učna množica v povprečju vsebuje le 63.2% primerov iz osnovne množice – preostali primeri so kopije. Opisan postopek imenujemo tudi *bootstrap*, ali natančneje *bootstrap 0.632*.

Na osnovi tako dobljenih učnih množic lahko s primernim učnim algoritmom zgradimo množico klasifikatorjev, katerih napovedi združimo v skupno napoved z večinskim glasovanjem. Potreben pogoj za uspešnost metode je, da so zgrajeni klasifikatorji med seboj različni, oziroma da je uporabljeni učni algoritem dovolj nestabilen. (Nestabilnost je lastnost učnega algoritma, da že majhna sprememba učne množice povzroči razliko

¹Ime je okrajšava angleškega izraza “bootstrap aggregating”.

v zgrajenem klasifikatorju.) Znale nestabilne metode so na primer metode za gradnjo odločitvenih dreves in metode za učenje nevronske mreže. Omenimo še, da lahko, če večinsko glasovanje nadomestimo z računanjem povprečja, metodo bagging uporabimo tudi na regresijskih problemih.

2.3.2 Metoda boosting

Metoda *boosting* [19] obsega celo družino metod, ki pa so si v osnovi zelo podobne. Podobno kot pri metodi bagging, tudi metoda boosting z glasovanjem kombinira klasifikatorje zgrajene z enim učnim algoritmom. Razlika je v tem, da smo pri metodi bagging gradnjo med seboj dopolnjujočih se klasifikatorjev (bolj ali manj) prepustili naključju, medtem ko pri metodi boosting z iterativnim postopkom načrtno gradimo klasifikatorje, ki se med seboj čim bolj dopolnjujejo. Postopek začnemo s prvim klasifikatorjem, ki je zgrajen nad celotno množico z enako uteženimi primeri. Za naslednje klasifikatorje v ansamblu želimo, da bi pravilno klasificirali primere, ki jih do sedaj zgrajeni klasifikatorji niso. Zato povečamo težo primerov, ki so bili do sedaj napačno klasificirani in zgradimo nov klasifikator. Postopek gradnje se ustavi, ko je izpolnjen nek ustavitveni kriterij (na primer, ko je natančnost novo zgrajenega klasifikatorja manjša ali enaka 0,5). Napovedi vseh zgrajenih klasifikatorjev združimo z uteženim glasovanjem, pri čemer damo večjo težo glasovom natančnejših klasifikatorjev.

Dodatna posebnost nekaterih metod iz družine boosting je, da nam nudijo teoretično jamstvo za svojo natančnost (glej npr. [23] ali [19]). Pokažemo namreč lahko, da se klasifikacijska napaka ansambla na učnih podatkih hitro zmanjšuje, ko povečujemo število iteracij (število članov ansambla). Edini pogoj za zmanjševanje napake je, da je napaka posameznih članov ansambla manjša od 50%. Vsaj za dvorazredne probleme običajno tega pogoja ni težko izpolniti. Čeprav nam jamstvo o poljubno majhni napaki na učni množici še ne zagotavlja majhne klasifikacijske napake na novih primerih, velja za metodo boosting tudi naslednje verjetnostno jamstvo za napako na novih primerih: napaka bo z veliko verjetnostjo manjša od podane meje, če bo število primerov dovolj veliko ter število iteracij dovolj majhno (glej npr. [30]).

Najbolj znana metoda boosting, za katero tudi velja zgoraj opisana lastnost, je metoda

AdaBoost.M1 [19]. Pogosto se za metode iz družine boosting kot sinonim uporablja tudi ime *arcing*² [4], čeprav naj bi bil izraz boosting rezerviran le za metode, za katere lahko pokažemo, da z njimi lahko zgradimo ansambel s poljubno majhno napako na učni množici.

2.3.3 Naključni gozdovi

Naključni gozdovi [5] so še ena izmed metod za kombiniranje klasifikatorjev, zgrajenih z enim učnim algoritmom, in sicer kot lahko sklepamo iz imena, z algoritmom za gradnjo odločitvenih dreves. Naključni gozdovi so v bistvu nadgradnja metode bagging. Z uporabo postopka bootstrap (glej poglavje 2.3.1) iz osnovne učne množice naredimo izbrano število novih učnih množic. Nad vsako množico nato zgradimo odločitveno drevo, pri čemer je atribut, na osnovi katerega razcepimo vozlišče v drevesu, vsakič izbran izmed določenega števila naključno izbranih atributov. Napovedi vseh zgrajenih dreves združimo z večinskim glasovanjem. Podobno kot bagging, so naključni gozdovi uporabni tudi na regresijskih problemih.

2.3.4 Glasovanje

Z glasovanjem lahko na preprost način združimo napovedi večih klasifikatorjev. Ločimo dve različici: večinsko glasovanje, kjer ima vsak glas enako težo; ter uteženo glasovanje, kjer se utež glasov lahko spreminja (npr. glas točnejšega klasifikatorja ima večjo težo). Kot že rečeno, je glasovanje sestavni del metod bagging, boosting in naključnih gozdov, lahko pa ga uporabljamo tudi za kombiniranje klasifikatorjev zgrajenih z različnimi učnimi algoritmi.

2.3.5 Izberi najboljšega

Namesto da na nek način združimo napovedi večih klasifikatorjev, lahko med njimi izberemo le enega, za katerega pričakujemo, da se bo najbolje obnesel na danem problemu. Tako lahko ocenimo natančnost klasifikatorjev, zgrajenih z različnimi učnimi algoritmi ter izberemo tistega, za katerega je ta ocena najboljša. Če za oceno natančnosti uporabljamo

²Okrajšava za angleški izraz “adaptively resample and combine”.

postopek prečnega preverjanja, lahko v povprečju (za več različnih problemov) pričakujemo večjo natančnost zgrajenih modelov kot če bi uporabljali en sam učni algoritem (glej [29]).

2.3.6 Metoda cascading

Drugačen pristop h kombiniranju uporablja metoda *cascading* [20]. Na začetku na osnovi prvotne učne množice z izbranim učnim algoritmom zgradimo klasifikator, ki nam kot napovedi za primere iz učne množice vrne verjetnostne porazdelitve po razredih. Tako dobljene porazdelitve dodamo v prvotno učno množico kot dodatne attribute ter na novi učni množici zgradimo nov klasifikator. Postopek lahko iterativno ponavljamo, pri čemer lahko v različnih korakih uporabljamo različne učne algoritme. Končni klasifikator je zgrajen iz večih nivojev, kar je nekoliko podobno metodi skladanja klasifikatorjev, ki jo bomo opisali v nadaljevanju.

2.3.7 Skladanje

Skladanje klasifikatorjev (ang. stacked generalization, stacking) je prvi predlagal Wolpert [39]. S skladanjem kombiniramo klasifikatorje, zgrajene z različnimi učnimi algoritmi L_1, L_2, \dots, L_N , na eni sami učni množici S . Učna množica je sestavljena iz primerov oblike $\mathbf{s}_i = (\mathbf{x}_i, y_i)$, kjer \mathbf{x}_i predstavlja vektor atributov, y_i pa razred i -tega primera. Postopek poteka v dveh korakih. V prvem na osnovi napovedi osnovnih klasifikatorjev zgradimo metaučno množico, v drugem koraku pa na tako dobljeni množici zgradimo metaklasifikator, ki zna združiti napovedi osnovnih klasifikatorjev v končno napoved. Prvi korak označujemo kot *nivo 0* ali *osnovni nivo*, drugega pa kot *nivo 1* ali *metanivo*.

V prvem koraku za generiranje metaučne množice uporabimo postopek *izloči enega* ali v splošnem postopek prečnega preverjanja K -tega reda. Učno množico S , ki vsebuje n primerov, razdelimo na K tujih podmnožic S_k približno enake velikosti in s podobnimi porazdelitvami vrednosti razredov (stratificirano). Sedaj za vsako podmnožico zgradimo skupino osnovnih klasifikatorjev $C_1^k, C_2^k, \dots, C_N^k$, pri čemer velja $C_j^k = L_j(S - S_k)$, ($\forall j = 1, \dots, N, \forall k = 1, \dots, K$). Naučene klasifikatorje uporabimo za napovedovanje primerov, ki niso bili vključeni v njihovo učno množico: $\hat{y}_i^j = C_j^k(x_i)$, $x_i \in S_k$. Iz tako dobljenih

Tabela 2.1: Osnovni algoritem metode skladanja klasifikatorjev.

```

function zgradi_ansambel( $S, \{L_1, L_2, \dots, L_N\}, K, L^m$ )
   $\{S_1, S_2, \dots, S_K\} = \text{stratificirano\_razdeli}(S, K)$ 
   $S^m = \{\}$ 
  for  $k = 1$  to  $K$  do
    for  $j = 1$  to  $N$  do
       $C_j^k = L_j(S - S_k)$ 
    endfor
     $S_k^m = \bigcup_{x_i \in S_k} \{(C_1^k(x_i), C_2^k(x_i), \dots, C_N^k(x_i), y_i)\}$ 
  endfor
   $S^m = \bigcup_{k=1}^K S_k^m$ 
  return  $L^m(S^m)$ 
endfunction

```

napovedi sestavimo metaučno množico S^m . Vsakemu primeru iz prvotne učne množice S ustreza v S^m primer oblike $\mathbf{s}_i^m = (\hat{\mathbf{y}}_i, y_i) = ((\hat{y}_i^1, \dots, \hat{y}_i^N), y_i)$. Atributi metaučne množice so torej napovedi osnovnih klasifikatorjev (\hat{y}_i^j), razred pa je pravi razred danega primera (y_i). V drugem koraku na tej metaučni množici zgradimo metaklasifikator z uporabo metaučnega algoritma L^m . Ko sedaj dobimo v klasifikacijo nov primer, morajo njegov razred najprej napovedati osnovni klasifikatorji, nato pa njihove napovedi metaklasifikator sestavi v končno napoved razreda. Celoten algoritem je podan v tabeli 2.1. Dodajmo še, da Wolpert predlaga za gradnjo metaklasifikatorjev uporabo »relativno globalnih in enakomernih« učnih algoritmov.

Kot možno slabost metode skladanja omenimo, da nimamo nobenega teoretičnega jamstva, da se bo klasifikacijska napaka na učni ali testni množici s povečevanjem števila zloženih osnovnih klasifikatorjev zmanjševala, kot ga imamo za nekatere metode iz družine boosting (glej poglavje 2.3.2). Prav tako ne moremo biti prepričani, da bo klasifikacijska napaka ansambla manjša od napak osnovnih klasifikatorjev.

Najpomembnejša dejavnika, ki vplivata na uspešnost metode skladanja, sta verjetno izbira metaatributov in izbira metaučnega algoritma. Na osnovi pravkar opisane metode

je bilo razvitih več izboljšav, ki jih bomo na kratko opisali v nadaljevanju.

Metoda SCANN

Znano je, da ansamblji z raznolikimi osnovnimi klasifikatorji (klasifikatorji, katerih napovedi so šibko korelirane) dosegajo dobro natančnost napovedi. Na tej osnovi je Merz [27] zasnoval metodo SCANN (Stacking, Correspondence Analysis, and Nearest Neighbor). Za metaatribute še vedno uporablja napovedi osnovnih klasifikatorjev, vendar jih nato s pomočjo korespondenčne analize transformira v attribute, ki so med seboj nekorelirani. Nad tako transformirano metaučno množico uporabimo algoritem najbližjih sosedov, ki nam da končno klasifikacijo neznanega primera.

Merz metodo SCANN primerja z več različnimi metodami za kombiniranje klasifikatorjev, med njimi tudi z metodo izbire najboljšega klasifikatorja (glej poglavje 2.3.5). Ugotavlja, da SCANN zagotavlja enako ali boljšo natančnost kot omenjene metode.

Skladanje z MLR

Ting in Witten [33] za metaatribute predlagata uporabo napovedi verjetnostnih porazdelitev razredov namesto napovedi razredov. Metaatributi so tako verjetnosti za vsakega od možnih razredov, ki nam jih vrnejo osnovni klasifikatorji. Na ta način upoševamo tudi zanesljivost napovedi osnovnih klasifikatorjev in ne le njihovih napovedi.

Vsak osnovni klasifikator mora biti zmožen napovedati verjetnost, s katero dani primer pripada vsakemu izmed možnih razredov. Napoved osnovnega klasifikatorja C za neznan primer x je tako verjetnostna porazdelitev:

$$\mathbf{p}^C(x) = (p^C(c_1|x), p^C(c_2|x), \dots, p^C(c_m|x)),$$

kjer je $\{c_1, c_2, \dots, c_m\}$ množica razredov, ki ji primer x lahko pripada, medtem ko $p^C(c_i|x)$ označuje verjetnost, da primer x pripada razredu c_i kot jo je napovedal klasifikator C . Metaatributi so tako verjetnosti za vsak možen razred kot jih napove vsak izmed osnovnih klasifikatorjev:

$$p^{C_j}(c_i|x) \quad \text{za } i=1, \dots, m \quad \text{in } j=1, \dots, N.$$

Poleg pravkar opisanih metaatributov, Ting in Witten predlagata uporabo večlične linearne regresije (ang. multi-response linear regression – MLR) za metaučni algoritem

ter hkrati pokazeta, da nekateri drugi algoritmi (glej [33]) niso primerni za to nalogo. MLR je prireditev linearne regresije za klasifikacijske probleme. Za vsak problem z m različnimi razredi $\{c_1, c_2, \dots, c_m\}$ zastavimo m regresijskih problemov: za vsak razred c_j zgradimo linearno enačbo LR_j , s katero napovedujemo vrednost binarne spremenljivke. Njena vrednost je enaka ena, če je razred enak c_j , in nič, če razred ni enak c_j . Ko hočemo klasificirati nov primer x , izračunamo vrednosti LR_j za vse j in napovemo tisti razred, za katerega je vrednost LR_j največja.

Avtorja poročata, da je natančnost napovedi skladanja z MLR boljša od natančnosti metod glasovanja in izbire najboljšega klasifikatorja s prečnim preverjanjem ter primerljiva z natančnostjo metod bagging in boosting.

Metaodločitvena drevesa

Metaodločitvena drevesa (ang. meta decision trees – MDT) [34] se od navadnih odločitvenih dreves razlikujejo po tem, da imajo namesto napovedi razreda v listih osnovne klasifikatorje. Zgrajena so na metaučni množici, ki ima za attribute le lastnosti napovedanih verjetnostnih porazdelitev (entropija in največja verjetnost) in ne celotnih verjetnostnih porazdelitev (kot na primer skladanje z MLR). Te lastnosti zajemajo tudi zanesljivost napovedi osnovnih klasifikatorjev. Za gradnjo metadreves uporabimo modificirani algoritem za gradnjo navadnih odločitvenih dreves. Zgrajena drevesa so (večinoma) relativno majhna, zato jih lahko (vsaj načeloma) pregledujemo in interpretiramo.

Avtorja poročata [35], da metaodločitvena drevesa dosegajo bistveno boljše napovedno natančnost kot metodi bagging in boosting, a le malenkost boljše v primerjavi z metodo SCANN in metodo izbora najboljšega klasifikatorja s prečnim preverjanjem. Nadalje je bilo v študiji [41] ugotovljeno, da je natančnost metadreves malenkost slabša od natančnosti metode skladanja z MLR.

Metoda grading

Metoda *grading* [32] poskuša zaznati in popraviti napačne napovedi osnovnih klasifikatorjev. Za vsak osnovni klasifikator zgradimo svoj metaklasifikator, katerega naloga je, da napove kdaj bo osnovni klasifikator naredil napako. Razred učne množice za gradnjo

vsakega izmed metaklasifikatorjev je enak ocenjenim (označenim kot pravilna ali nepravilna, ang. graded) napovedim ustreznih osnovnih klasifikatorjev, medtem ko so ostali atributi enaki kot v osnovni učni množici. Napoved razreda za nov primer dobimo z glasovanjem tistih osnovnih klasifikatorjev, katerih metaklasifikatorji so napovedali pravilno klasifikacijo.

Avtorja metode poročata, da je natančnost napovedi njune metode bistveno boljša od natančnosti glasovanja in metode izbire najboljšega klasifikatorja s prečnim preverjanjem ter malenkost slabša od natančnosti metode skladanja z MLR.

Skladanje C

Seewald je predlagal metodo, ki jo je poimenoval *skladanje C* (ang. stacking C³) [31]. Metoda je v osnovi enaka skladanju z MLR, le da uporablja drugačne attribute za opis vsakega izmed m binarnih problemov (glej poglavje 2.3.7). Za problem, ki se nanaša na razred c_j , uporablja za attribute le verjetnosti tega razreda (ki nam jih vrnejo osnovni klasifikatorji) in ne verjetnosti vseh razredov, kot to počne originalna metoda skladanja z MLR. Vsak izmed problemov na metanivoju ima tako N atributov namesto mN . Avtor poroča, da metoda nudi boljše natančnost napovedi kot originalna metoda predvsem na domenah z več razredi.

Pravkar podani pregled metod za kombiniranje klasifikatorjev nikakor ne zajema vseh znanih metod, vendar veliko število metod za kombiniranje in njihovih različic temelji na predstavljenih metodah in principih. V naslednjem poglavju bomo predstavili rezultate primerjalne analize večine omenjenih metod.

³Ime je okrajšava za “stacking with confidences”.

3. Primerjava

obstoječih metod za kombiniranje

Metode, predstavljene v poglavju 2.3, so v originalnih člankih ovrednotene po zelo različnih postopkih. Uporabljene so na različnih učnih množicah, njihova natančnost je ocenjena z različnimi postopki, prav tako tudi signifikantnost razlik med posameznimi metodami. Ni nam bila znana nobena študija iz katere bi bilo razvidno, katera metoda je boljša od druge ter ali je sploh katera metoda za kombiniranje heterogenih klasifikatorjev boljša od izbire najboljšega klasifikatorja s prečnim preverjanjem. Zaradi tega smo naredili primerjavo večine metod za kombiniranje po enotni metodologiji. V nadaljevanju poglavja bomo najprej opisali uporabljen metodologijo primerjanja posameznih metod, nato pa podali rezultate primerjave in njihov komentar.

3.1 Metodologija vrednotenja

Za vrednotenje različnih metod smo izbrali trideset množic podatkov iz zbirke UCI [2]. Množice iz te zbirke se množično uporabljajo v različnih primerjalnih študijah. V izboru smo se izognili množicam z manj kot sto primeri. Izbrane množice, kot tudi njihove lastnosti, so predstavljene v tabeli 3.1.

Pri ocenjevanju posamezne metode smo se osredotočili na oceno klasifikacijske napake, medtem ko smo pri medsebojnem primerjanju dveh metod ugotavljali relativno izboljšanje natančnosti in statistično signifikantnost razlik v natančnosti. V nadaljevanju bomo opisali način računanja vseh treh pokazateljev.

Tabela 3.1: Uporabljene množice podatkov in njihove lastnosti: število primerov, razredov, (diskretnih/zveznih) atributov, verjetnost večinskega razreda in entropija verjetnostne porazdelitve razreda. (Vir: *UCI Repository of machine learning databases* [2].)

MNOŽICA PODATKOV	ŠT. PRIM.	ŠT. RAZ.	(D/Z)	ATR.	VER. VEČ.	ENT.
AUSTRALIAN	690	2	(8/6)	14	0.56	0.99
BALANCE	625	3	(0/4)	4	0.46	1.32
BREAST-CANCER	286	2	(9/0)	9	0.70	0.88
BREAST-W	699	2	(9/0)	9	0.66	0.92
BRIDGES-TD	102	2	(4/3)	7	0.85	0.61
CAR	1728	4	(6/0)	6	0.70	1.21
CHESS	3196	2	(36/0)	36	0.52	0.99
CONTRACEPTIVE	1473	3	(4/5)	9	0.43	1.54
DIABETES	768	2	(0/8)	8	0.65	0.93
DIS	3772	2	(22/6)	28	0.98	0.11
ECHO	131	2	(1/5)	6	0.67	0.91
GERMAN	1000	2	(13/7)	20	0.70	0.88
GLASS	214	6	(0/9)	9	0.36	2.18
HEART-C	303	5	(7/6)	13	0.54	0.99
HEART-H	294	5	(7/6)	13	0.64	0.94
HEART	270	2	(6/7)	13	0.56	0.99
HEPATITIS	155	2	(13/6)	19	0.79	0.74
HYP0	3163	2	(18/7)	25	0.95	0.29
IMAGE	2310	7	(0/19)	19	0.14	2.78
IONOSPHERE	351	2	(0/34)	34	0.64	0.94
IRIS	150	3	(0/4)	4	0.33	1.58
SOLAR-FLARE-C	1389	8	(10/0)	10	0.84	0.88
SOLAR-FLARE-M	1389	6	(10/0)	10	0.95	0.34
SOLAR-FLARE-X	1389	3	(10/0)	10	0.99	0.75
SONAR	208	2	(0/60)	60	0.53	1.00
SOYA	683	19	(35/0)	35	0.13	3.79
TIC-TAC-TOE	958	2	(9/0)	9	0.65	0.93
VOTE	435	2	(16/0)	16	0.61	0.96
WAVEFORM	5000	3	(0/21)	21	0.34	1.58
WINE	178	3	(0/13)	13	0.40	1.56

3.1.1 Klasifikacijska napaka

Klasifikacijsko napako posamezne metode na dani domeni smo ocenjevali z uporabo prečnega preverjanja desetega reda. Pri razdeljevanju učne množice v podmnožice smo poskušali generirati podmnožice s čim bolj podobno porazdelitvijo razredov (ang. stratified cross-validation). Prečno preverjanje smo ponovili desetkrat z različnimi semeni generatorja naključnih števil za vzorčenje množice. Enaka semena so bila uporabljena v vseh poskusih. Oceno klasifikacijske napake učnega algoritma C smo izračunali kot povprečje ocen dobljenih v desetih prečnih preverjanjih:

$$error(C) = \frac{1}{n} \sum_{i=1}^n error_i(C), \quad n = 10.$$

3.1.2 Relativno izboljšanje

Oceno izboljšanja natančnosti na določeni domeni, če smo namesto metode C_2 uporabili metodo C_1 , smo ovrednotili z relativnim izboljšanjem natančnosti:

$$RI(C_1, C_2) = 1 - \frac{error(C_1)}{error(C_2)}.$$

Povprečno relativno izboljšanje prek vseh uporabljenih domen smo izračunali z geometrijskim povprečjem:

$$ARI(C_1, C_2) = 1 - \sqrt[n]{\prod_{i=1}^n \frac{error(C_1^i)}{error(C_2^i)}},$$

kjer n pomeni število uporabljenih domen (v našem primeru trideset). Potrebno je opozoriti, da mera ni simetrična, ker v splošnem $ARI(C_1, C_2)$ ni enako $-ARI(C_1, C_2)$.

3.1.3 Statistična signifikantnost

Statistično signifikantnost razlik v natančnosti dveh metod smo preverjali z uporabo dvojnega t -testa z gotovostjo 95%. S pomočjo t -testa lahko ocenimo ali se (z zahtevano gotovostjo) povprečji dveh skupin meritev razlikujeta, oziroma v našem primeru, ali se razlikujeta povprečji klasifikacijskih napak dveh primerjanih klasifikatorjev. Pri tem zahtevamo, da so posamezne meritve v skupini medseboj neodvisne. Tej zahtevi v strojnem učenju zelo težko popolnoma ustrezemo, saj imamo običajno na razpolago premalo učnih primerov, skušamo pa se ji čim bolj približati. Znanih je več postopkov, s katerimi

dobimo skupino (čimbolj neodvisnih) ocen klasifikacijske napake za dani klasifikator (glej na primer [8]). Uporabljali smo naslednje tri postopke:

- *t-test* 10×10 . Posamezno oceno klasifikacijske napake dobimo s prečnim preverjanjem desetega reda. Prečno preverjanje ponovimo destkrat, vsakič primere učne množice drugače naključno premešamo, oziroma za vsako prečno preverjanje uporabimo drugačno seme naključnega generatorja. Vse uporabljene podmnožice (oziroma semena) v prečnih preverjanjih morajo biti za oba primerjana učna algoritma enake. Težava pri tem postopku je, da se tako učne kot testne množice v desetih ponovitvah prečnega preverjanja močno prekrivajo, zato dobljene ocene napak niso neodvisne. Velika verjetnost je, da bomo tako podcenili varianco klasifikatorjev in zaznali signifikantno razliko, kjer je v resnici ni (napaka tipa I). Kljub omenjeni slabosti je uporaba tega postopka dokaj razširjena.
- *t-test* 1×10 . Deset ocen klasifikacijske napake dobimo iz ene same ponovitve prečnega preverjanja desetega reda. V tem primeru se učne množice sicer prekrivajo, vendar so testne množice neodvisne. Verjetnost napake tipa I je zato manjša.
- *t-test* 5×2 . Postopek je predlagal Dietterich [8]. S petimi ponovitvami (pod enakimi pogoji kot v t-testu 10×10) prečnega preverjanja drugega reda dobimo pet ocen napake. Tako učne kot testne množice se v tem primeru ne prekrivajo, zato je verjetnost napake tipa I mnogo manjša. Slabost metode (poleg, kot bomo videli, slabe občutljivosti) je v tem, da napako klasifikatorja, zgrajenega na določeni učni množici, ocenjujemo z uporabo za polovico manjše učne množice. Postopek zato ni primeren za učne množice z malo primeri.

V tabeli 3.2 je prikazan značilen primer uporabe vseh treh postopkov za primerjavo klasifikatorjev A in B.¹ Zanima nas, če lahko z uporabo metode A namesto metode B, zmanjšamo klasifikacijsko napako. Testi 10×10 , 1×10 in 5×2 nam po vrsti dajo zmanjšanje klasifikacijske napake na 18., 7. in 2. domenah ter povečanje na 2., nobeni in na eni domeni. Vidimo, da je občutljivost testa 10×10 bistveno večja od občutljivosti ostalih

¹Za primer smo vzeli metodo skladanja klasifikatorjev z MLR ter preprosto metodo večinskega glasovanja, obe s tremi osnovnimi klasifikatorji.

dveh testov, vendar pa iz zgoraj povedanega lahko sklepamo, da je v rezultatu 18:2 veliko napak tipa I. Po drugi strani bi lahko iz rezultatov testa 5×2 sklepali, da med obema metodama sploh ni nobene razlike. Zaradi tega se zdi test 1×10 primeren kompromis med občutljivostjo in številom napak tipa I. V nadaljevanju bomo tako navajali le rezultate testa 1×10 .

3.2 Uporabljene metode

S poskusi smo želeli primerjati večino najbolj znanih metod za kombiniranje klasifikatorjev. Zanimala nas je predvsem uspešnost metod za skladanje klasifikatorjev v primerjavi z izborom najboljšega klasifikatorja s prečnim preverjanjem ter v primerjavi z zelo razširjenimi metodami za kombiniranje klasifikatorjev zgrajenih z istim učnim algoritmom. Vse poskuse smo izvedli s programi iz zbirke za rudarjenje podatkov Weka [38]. Zbirka je napisana v programskem jeziku Java in je dostopna pod licenco GNU. Poskuse smo izvedli z naslednjimi metodami za kombiniranje:

- BAGG: Metoda bagging (glej poglavje 2.3.1) z uporabo odločitvenih dreves J4.8 (reimplementacija algoritma C4.5 [28]). Gradili smo ansamble, sestavljene iz desetih odločitvenih dreves.
- BOOS: Metoda boosting (glej poglavje 2.3.2), oziroma njena najbolj razširjena različica AdaBoost.M1 [19]. Enako kot pri metodi bagging smo gradili ansamble z desetimi odločitvenimi drevesi.
- RFOR: Naključni gozdovi (glej poglavje 2.3.3), sestavljeni iz stotih odločitvenih dreves.
- VOTE: Preprosto večinsko glasovanje (glej poglavje 2.3.4).
- SELB: Izbiranje najboljšega klasifikatorja s prečnim preverjanjem (glej poglavje 2.3.5). Uporabili smo prečno preverjanje desetega reda.
- SMDT: Metaodločitvena drevesa (glej poglavje 2.3.7). V zbirki Weka metaodločitvenih dreves ni, zato smo jih reimplementirali v programskem jeziku Java ter vključili v zbirko.

Tabela 3.2: Primer uporabe različnih načinov testiranja signifikantnosti razlik v natančnosti dveh klasifikatorjev. Za vsako množico podatkov so podane klasifikacijske napake za metodi A in B (v %); relativno izboljšanje natančnosti (v %), če namesto metode B uporabimo metodo A; ter signifikantnost razlik ob uporabi testov 10×10 , 1×10 in 5×2 . »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA PODATKOV	NAPAKA A	NAPAKA B	REL. IZB.	10×10	1×10	5×2
AUSTRALIAN	14.28	14.17	-0.72	.	.	.
BALANCE	10.16	13.41	24.22	+	+	.
BREAST-CANCER	27.73	25.31	-9.53	-	.	.
BREAST-W	2.73	3.46	21.07	+	.	.
BRIDGES-TD	14.12	15.29	7.69	+	.	.
CAR	5.61	6.49	13.55	+	.	.
CHESSE	0.60	1.43	58.20	+	+	.
CONTRACEPTIVE	45.91	47.78	3.91	+	.	.
DIABETES	23.71	24.22	2.10	+	.	.
DIS	0.96	1.33	27.74	+	+	.
ECHO	28.32	31.53	10.17	+	.	.
GERMAN	24.63	24.92	1.16	.	.	-
GLASS	30.84	29.21	-5.60	.	.	.
HEART-C	17.59	18.28	3.79	.	.	.
HEART-H	16.80	16.63	-1.02	.	.	.
HEART	15.74	18.22	13.62	+	.	.
HEPATITIS	15.81	17.94	11.87	+	.	.
HYP0	0.72	1.36	46.62	+	+	.
IMAGE	2.86	2.94	2.65	.	.	.
IONOSPHERE	7.35	7.15	-2.79	.	.	.
IRIS	4.73	4.40	-7.58	.	.	.
SOLAR-FLARE-C	15.91	16.16	1.56	+	.	.
SOLAR-FLARE-M	5.17	5.13	-0.84	.	.	.
SOLAR-FLARE-X	0.90	1.00	10.07	+	.	.
SONAR	13.51	17.31	21.94	+	.	.
SOYA	7.07	6.71	-5.46	.	.	.
TIC-TAC-TOE	0.58	9.24	93.67	+	+	+
VOTE	3.54	6.90	48.67	+	+	.
WAVEFORM	16.97	18.42	7.87	+	+	+
WINE	2.87	1.74	-64.52	-	.	.
POVPREČJE/SKUPAJ	12.59	13.60	18.50	18:2	7:0	2:1

- SMLR: Skladanje klasifikatorjev z MLR (glej poglavje 2.3.7).
- SCMLR: Metoda skladanje C (glej poglavje 2.3.7).

Zanimal nas je tudi vpliv števila osnovnih klasifikatorjev na uspešnost skladanja klasifikatorjev, zato smo vse poskuse z metodami za skladanje heterogenih klasifikatorjev opravili na dveh skupinah osnovnih klasifikatorjev. V prvi skupini so naslednji trije osnovni klasifikatorji:

- J4.8: Reimplementacija znanega algoritma za gradnjo odločitvenih dreves C4.5 [28].
- IBK: Metoda k -najbližjih sosedov [1], število sosedov je enako 1.
- NB: Naivni Bayesov klasifikator [22].

Drugo skupino sestavlja sedem osnovnih klasifikatorjev, poleg zgornjih treh še naslednji štirje:

- K*: Algoritem, ki nove primere klasificira na osnovi množice že klasificiranih primerov. Pri tem uporablja razdaljo na osnovi entropije [6].
- KDE: Algoritem za klasifikacijo na osnovi ocenjevanja gostote jedra.
- DT: Algoritem za klasifikacijo na osnovi odločitvene tabele [24].
- MLR: Algoritem večlične linearne regresije kot je predlagan v [33] in opisan v poglavju 2.3.7.

V vseh poskusih smo algoritme uporabili z njihovimi privzetimi nastavitvami, razen algoritem NB, kjer smo uporabili ocenjevanje porazdelitev numeričnih atributov z gostoto jedra (ang. kernel density estimation). Rezultati poskusov so predstavljeni v naslednjem razdelku.

3.3 Rezultati primerjave

Rezultati poskusov so v celoti podani v prilogi A. V tabeli 1 so podane klasifikacijske napake osnovnih klasifikatorjev, v tabelah 2 in 3 so podane klasifikacijske napake ansamblov

s tremi in s sedmimi osnovnimi klasifikatorji. Na smiselnost uporabe ansamblov v primerjavi s posameznimi klasifikatorji, lahko sklepamo iz tabel 4 in 5, kjer so podana relativna izboljšanja natančnosti, če namesto posameznega klasifikatorja uporabimo metodo skladanja klasifikatorjev z MLR. V tabelah 6 in 7 so podana relativna izboljšanja natančnosti ob uporabi skladanja klasifikatorjev z MLR, v primerjavi z ostalimi metodami za kombiniranje (s tremi in sedmimi osnovnimi klasifikatorji). Nazadnje so v tabeli 8 podana relativna izboljšanja natančnosti ob uporabi sedmih osnovnih klasifikatorjev namesto le treh.

Preden se lotimo medsebojne primerjave posameznih metod za kombiniranje klasifikatorjev, moramo odgovoriti na vprašanje ali se kombiniranje sploh izplača, oziroma ali s kombiniranjem lahko dosežemo občutno večjo natančnost klasifikacije. Iz tabel 4 in 5 lahko vidimo, da z uporabo metode skladanja z MLR na testnih domenah dosežemo povprečno relativno izboljšanje med 26% in 43%. Ob tem so izboljšanja natančnosti statistično signifikantna na eni do dveh tretjinah vseh testnih domen. Na osnovi teh rezultatov lahko z veliko gotovostjo trdimo, da je uporaba metod za kombiniranje klasifikatorjev smiselna, ko želimo doseči čim boljšo točnost napovedovanja.

Za medsebojno primerjavo posameznih metod sta tabeli 3.3 in 3.4 bolj primerni kot tabele 1–7. Podajata namreč povprečno relativno izboljšanje metode X v primerjavi z metodo Y za vsak par X in Y ter tudi število statistično signifikantnih zmag in porazov (dobljenih z uporabo t-testa 1×10). V tabeli 3.3 so podatki za ansamble s tremi osnovnimi klasifikatorji, medtem ko so v tabeli 3.4 podatki za ansamble s sedmimi klasifikatorji (podatki za metode, ki ne kombinirajo heterogenih klasifikatorjev, so v obeh tabelah enaki).

V tabeli 3.3 lahko opazimo, da se metode skladanja klasifikatorjev (SMDT, SMLR in SCMLR) ter metoda izbire najboljšega klasifikatorja s prečnim preverjanjem (SELB) veliko bolje obnesejo, kot metode bagging (BAGG), boosting (BOOS), naključna drevesa (RFOR) in glasovanje (VOTE). Povprečno relativno izboljšanje natančnosti (ARI) ob uporabi metod iz druge skupine v primerjavi z metodami iz prve skupine je med 15% in 20%, z izjemo naključnih dreves, kjer je ARI okoli 10%. Tudi če primerjamo število zmag in porazov je slika podobna: metode iz druge skupine imajo od 6 do 11 zmag več kot porazov, spet z izjemo naključnih dreves, kjer je razlika med 3 in 5. Če primerjamo

Tabela 3.3: Relativno izboljšanje natančnosti ansamblov, zgrajenih z različnimi metodami. Prve tri metode kombinirajo odločivna drevesa, preostale pa **tri različne osnovne klasifikatorje**. Številka v vrstici X in stolpcu Y pomeni relativno izboljšanje natančnosti, če uporabimo metodo X namesto metode Y. Poleg tega je podano število zmag:porazov metode X nad metodo Y (glede na 1×10 t-test). V zadnjem stolpcu je seštevek vseh zmag in porazov v vrstici.

	BAGG	BOOS	RFOR	VOTE	SELB	SMDT	SMLR	SCMLR	SKUPAJ
BAGG		-0.95 5:5	-9.13 2:4	0.44 6:3	-19.74 1:7	-17.62 1:7	-22.16 0:6	-23.91 0:7	15:39
BOOS	0.94 5:5		-8.11 0:2	1.38 6:6	-18.62 3:8	-16.51 3:8	-21.02 3:12	-22.75 2:13	22:54
RFOR	8.37 4:2	7.50 2:0		8.77 4:3	-9.73 2:5	-7.77 1:5	-11.94 1:4	-13.55 0:5	14:24
VOTE	-0.44 3:6	-1.40 6:6	-9.61 3:4		-20.28 1:8	-18.14 2:8	-22.71 0:7	-24.46 0:8	15:47
SELB	16.49 7:1	15.70 8:3	8.86 5:2	16.86 8:1		1.78 1:1	-2.02 1:1	-3.48 2:1	32:10
SMDT	14.98 7:1	14.17 8:3	7.21 5:1	15.35 8:2	-1.81 1:1		-3.87 2:1	-5.35 2:2	33:11
SMLR	18.14 6:0	17.37 12:3	10.67 4:1	18.50 7:0	1.98 1:1	3.72 1:2		-1.43 1:1	32:8
SCMLR	19.30 7:0	18.53 13:2	11.93 5:0	19.65 8:0	3.36 1:2	5.08 2:2	1.41 1:1		37:7

Tabela 3.4: Relativno izboljšanje natančnosti ansamblov, zgrajenih z različnimi metodami. Tabela je enaka zgornji, le da ansambli heterogenih klasifikatorjev (zadnjih pet metod) uporabljajo **sedem različnih osnovnih klasifikatorjev**.

	BAGG	BOOS	RFOR	VOTE	SELB	SMDT	SMLR	SCMLR	SKUPAJ
BAGG		-0.95 5:5	-9.13 2:4	-5.78 4:6	-22.38 0:9	-21.47 1:9	-25.00 0:10	-28.75 0:11	12:54
BOOS	0.94 5:5		-8.11 0:2	-4.79 4:5	-21.23 3:10	-20.33 1:8	-23.82 0:9	-27.54 0:11	13:50
RFOR	8.37 4:2	7.50 2:0		3.07 2:2	-12.14 0:6	-11.31 0:5	-14.54 1:3	-17.98 0:4	9:22
VOTE	5.46 6:4	4.57 5:4	-3.17 2:2		-15.70 2:10	-14.84 0:9	-18.17 0:11	-21.72 0:10	15:50
SELB	18.29 9:0	17.52 10:3	10.83 6:0	13.57 10:2		0.74 1:2	-2.13 3:2	-5.21 2:1	41:10
SMDT	17.68 9:1	16.90 8:1	10.16 5:0	12.92 9:0	-0.75 2:1		-2.90 4:1	-5.99 3:4	40:8
SMLR	20.00 10:0	19.24 9:0	12.69 3:1	15.37 11:0	2.09 2:3	2.82 1:4		-3.01 1:2	37:10
SCMLR	22.33 11:0	21.59 11:0	15.24 4:0	17.84 10:0	4.95 1:2	5.65 4:3	2.92 2:1		43:6

metode iz druge skupine medseboj, opazimo da so razlike majhne: glede na ARI med -5% in 5% ter z minimalno razliko števila zmag in porazov.

Razmere se bistveno ne spremenijo, če namesto ansamblov s tremi, primerjamo ansamble s sedmimi različnimi osnovnimi klasifikatorji (tabela 3.4). Povprečno relativno izboljšanje natančnosti metod iz druge skupine v primerjavi z metodami iz druge skupine se za nekaj odstotkov poveča (z izjemo glasovanja, kjer se nekoliko zmanjša), prav tako se poveča razlika med številom zmag in porazov. Razlike med metodami druge skupine so tudi v tem primeru minimalne. Poveča se edino število zmag metaodločitvenih dreves proti metodi skladanja z MLR (na 4:1).

Ob upoštevanju povprečnega relativnega izboljšanja za ansamble s tremi in sedmimi osnovnimi klasifikatorji lahko sklenemo, da z metodo SCMLR dobimo nekoliko boljše rezultate kot z metodo SMLR, s katero spet dobimo nekoliko boljše rezultate kot z metodama SELB in SMDT. S to ugotovitvijo se večinoma skladajo tudi razlike v številu zmag in porazov, vendar so razlike majhne in lahko sklenemo, da so vse tri metode skladanja klasifikatorjev med seboj primerljive ter da njihova natančnost ni bistveno boljša od natančnosti metode izbire najboljšega klasifikatorja s prečnim preverjanjem.

Zaradi večje računske zahtevnosti ter zanemarljivo boljše natančnosti metod za skladanje, v primerjavi z metodo SELB, njihovo uporabo težko upravičimo. To nas je vzbudilo k iskanju novih metod skladanja klasifikatorjev, katerih natančnost bi bila bistveno boljša kot natančnost metode SELB. Odkriti metodi sta predstavljeni v naslednjem poglavju.

4. Izboljšano skladanje klasifikatorjev

V prejšnjem poglavju smo predstavili rezultate poskusov, iz katerih sledi, da nam različne metode za skladanje klasifikatorjev nudijo medseboj primerljivo natančnost klasifikacije. Hkrati je njihova natančnost tudi primerljiva z natančnostjo metode izbire najboljšega klasifikatorja s prečnim preverjanjem. Primerjava je vseeno pokazala, da je metoda skladanja z MLR malenkost boljša od ostalih metod, zato se nam je zdelo smiselno vzeti omenjeno metodo kot osnovo pri iskanju boljših metod za skladanje.¹ Razvili smo dve izboljšavi skladanja z MLR; prva uporablja razširjen nabor metaatributov, druga pa drugačen učni algoritem na metanivoju. Obe metodi bomo opisali v nadaljevanju.

4.1 Razširitev nabora metaatributov

Spomnimo se (glej poglavje 2.3.7 ali [33]), da metoda skladanja z MLR za metaatribute uporablja napovedi verjetnosti vseh osnovnih klasifikatorjev za vsak možen razred:

$$p^{C_j}(c_i|x) \quad \text{za } i = 1, \dots, m \text{ in } j = 1, \dots, N,$$

kjer N pomeni število osnovnih klasifikatorjev in m število razredov v dani množici podatkov. Kot metaučni algoritem metoda uporablja večlično linearno regresijo (ang. multi-response linear regression – MLR).

Naša prva metoda za metaučni algoritem ravno tako uporablja MLR. Razlika je v uporabljenih metaatributih. Poleg zgoraj omenjenih napovedi verjetnosti $p^{C_j}(c_i|x)$, smo uporabili še dve dodatni množici atributov [40]:

¹Metoda skladanje C, katere rezultati so še nekoliko boljši od metode skladanja z MLR, v tem času še ni bila predstavljena.

- verjetnostne porazdelitve, pomnožene z največjo verjetnostjo

$$P_{C_j} = p^{C_j}(c_i|x) \times M_{C_j(x)} = p^{C_j}(c_i|x) \times \max_{i=1}^m (p^{C_j}(c_i|x))$$

za $i = 1, \dots, m$ in $j = 1, \dots, N$ ter

- entropije verjetnostnih porazdelitev

$$E_{C_j}(x) = - \sum_{i=1}^m p_{C_j}(c_i|x) \cdot \log_2 p_{C_j}(c_i|x).$$

Skupno število metaatributov, ki jih metoda uporablja je enako $N(2m + 1)$.

Pojasnimo še, zakaj smo se odločili za tako razširjen nabor metaatributov. Že Ting in Witten [33] sta kot razlog za uporabo verjetnostnih porazdelitev v metodi skladanja z MLR navedla, da verjetnostne porazdelitve vsebujejo ne samo napovedi osnovnih klasifikatorjev, pač pa tudi njihovo zanesljivost. Naši dodatni atributi poskušajo bolj določeno zajeti zanesljivost napovedi. Omenimo, da so atributi M_{C_j} in E_{C_j} edini metaatributi, ki jih uporabljajo metaodločitvena drevesa (glej poglavje 2.3.7); od njih je odvisna dobra klasifikacijska točnost metaodločitvenih dreves in so zato gotovo uporabni za učenje na metanivoju.

Entropija (atributi E_{C_j}) je mera za negotovost; večja kot je entropija napovedane verjetnostne porazdelitve, večja je negotovost napovedi. Največja verjetnost v napovedani verjetnostni porazdelitvi M_{C_j} prav tako vsebuje informacijo o zanesljivosti napovedi; večja vrednost M_{C_j} pomeni, da je napoved zanesljiva, medtem ko manjša vrednost M_{C_j} pomeni, da je napoved nezanesljiva. Atributi P_{C_j} združujejo napovedi same (posamezne verjetnosti) in zanesljivost napovedi, ki jih vsebujejo največje verjetnosti M_{C_j} v napovedani porazdelitvi. Opisano kombinacijo atributov smo dodali v upanju, da bo metaučni algoritem lažje uporabil tako sestavljeno informacijo, ter se mu ne bo treba učiti, da je potrebno informacijo o napovedih uporabljati v kombinaciji z njihovo zanesljivosti.

Omeniti moramo, da smo predhodne poskuse opravili le z uporabo atributov P_{C_j} in E_{C_j} (brez originalnih verjetnostnih porazdelitev). Rezultati niso pokazali signifikantnega izboljšanja glede na uporabo samih verjetnostnih porazdelitev. Morebitno izboljšanje klasifikacijske točnosti pravkar opisane metode tako temelji na sinergiji uporabe vseh treh tipov atributov hkrati.

4.2 Skladanje z večličnimi modelskimi drevesi

V naši drugi izboljšavi metode skladanja klasifikatorjev z MLR [33] smo uporabili nespremenjeno množico metaatributov, to je, verjetnostne porazdelitve kot jih napovedo osnovni klasifikatorji. Usmerili smo se v iskanje algoritma za učenje na metanivoju, ki bi boljši kot je algoritem večlične linearne regresije (MLR).

Skladanje z MLR uporablja linearno regresijo za klasifikacijo. Modelska drevesa lahko obravnavamo kot razširitev linearnih modelov (linearne regresije) na odsekoma linearne modele. V iskanju alternative algoritmu MLR, je bilo zato logično raziskati možnost uporabe modelskih dreves namesto linearne regresije. Poleg tega so Frank in sodelavci [18] pokazali, da z uporabo modelskih dreves za klasifikacijo dosežemo boljšo klasifikacijsko točnost kot z uporabo MLR.

Spomnimo se (glej poglavje 2.3.7), da MLR zastavi po eno binarno klasifikacijsko nalogo za vsako različno vrednost razreda dane domene. Za vsak razred c_j tako zgradimo linearno enačbo LR_j , s katero napovedujemo vrednost binarne spremenljivke. Ko dobimo v klasifikacijo nov primer x , izračunamo $LR_j(x)$ za vse vrednosti j . Napovedani razred novega primera c_k je tisti, pri katerem je vrednost $LR_k(x)$ največja.

Naša metoda namesto linearne regresije uporablja modelska drevesa, vse ostalo pa je enako kot pri metodi skladanja z MLR. Namesto m linearnih enačb LR_j zgradimo m modelskih dreves MT_j . Za klasifikacijo novega primera x izračunamo vrednosti $MT_j(x)$ za vse j ter napovemo razred c_k , katerega vrednost $MT_k(x)$ je največja. Za gradnjo modelskih dreves smo uporabili algoritem M5' [37]. Glede na sorodnost naše metode z metodo skladanja z MLR, smo našo metodo poimenovali skladanje klasifikatorjev z večličnimi modelskimi drevesi [17].

4.3 Primerjava z obstoječimi metodami

Obe pravkar predstavljeni metodi skladanja klasifikatorjev smo ovrednotili in primerjali z že znanimi metodami predstavljenimi v poglavjih 2 in 3; tokrat smo med seboj primerjali le metode za kombiniranje heterogenih klasifikatorjev:

- VOTE: Preprosto večinsko glasovanje (glej poglavje 2.3.4).

- SELB: Izbiranje najboljšega klasifikatorja s prečnim preverjanjem (glej poglavje 2.3.5).
- SMDT: Metaodločitvena drevesa (glej poglavje 2.3.7).
- SMLR: Skladanje klasifikatorjev z MLR (glej poglavje 2.3.7).
- SCMLR: Metoda skladanje C (glej poglavje 2.3.7).
- SMLRE: Skladanje klasifikatorjev z razširjenim naborom metaatributov in MLR (glej poglavje 4.1).
- SMM5: Skladanje klasifikatorjev z večličnimi modelskimi drevesi (glej poglavje 4.2).

Uporabili smo natanko enako metodologijo, opisano v poglavju 3.1: iste učne množice (glej tabelo 3.1), enak način ocenjevanja točnosti (povprečje desetih ponovitev prečnega preverjanja desetega reda) in enak postopek ocenjevanja signifikantnosti morebitnih razlik (t-test 1×10).

Rezultati poskusov so v celoti podani v prilogi B. V tabelah 9 in 10 so podane klasifikacijske napake različnih ansamblov s tremi in sedmimi osnovnimi klasifikatorji. V tabelah 11 in 12 so podana relativna izboljšanja natančnosti ob uporabi skladanja klasifikatorjev z večličnimi modelskimi drevesi, v primerjavi z ostalimi metodami za kombiniranje (s tremi in sedmimi osnovnimi klasifikatorji). Nazadnje so v tabeli 13 podana relativna izboljšanja natančnosti ob uporabi sedmih osnovnih klasifikatorjev namesto le treh.

Posamezne metode lažje kot s pomočjo tabel 11 in 12, medseboj primerjamo na osnovi tabel 4.1 in 4.2. Tabeli sta analogni tabelam 3.3 in 3.4 iz poglavja 3. Podajata povprečno relativno izboljšanje metode X v primerjavi z metodo Y za vsak par X in Y ter tudi število statistično signifikantnih zmag in porazov (dobljenih z uporabo t-testa 1×10). V tabeli 4.1 so podatki za ansamble s tremi osnovnimi klasifikatorji, medtem ko so v tabeli 4.2 podatki za ansamble s sedmimi klasifikatorji.

Skladanje z razširjenim naborom metaatributov

Oglejmo si najprej metodo skladanja z razširjenim naborom metaatributov in MLR (SMLRE). Z uporabo treh osnovnih klasifikatorjev (tabela 4.1) je, glede na razliko števila signifikantnih zmag in porazov, metoda uspešnejša od ostalih metod za kombiniranje (z izjemo

metode SMM5, o kateri bomo govorili v naslednjem razdelku). Ima le po en poraz proti metodama SMLR in SCMLR. Skupna razlika med zmagami in porazi je +10. Slabše se metoda obnese glede na povprečno relativno izboljšanje: boljša je le od metod SMDT in VOTE. Vseeno so razlike majhne, razen v primerjavi z metodo VOTE. Ob uporabi sedmih osnovnih klasifikatorjev (tabela 4.2) je glede na povprečno relativno izboljšanje metoda slabša od metod za skladanje in metode SELB, po številu zmag in porazov pa je od njih nekoliko boljša. Sklenemo lahko, da nam metoda SMLRE nudi malenkost boljšo klasifikacijsko točnost kot metoda SELB in ostale metode za skladanje, še posebej, če kombiniramo manjše število osnovnih klasifikatorjev.

Skladanje z večličnimi modelskimi drevesi

Sedaj si oglejmo še metodo skladanja z večličnimi modelskimi drevesi (SMM5). Iz tabele 4.1 lahko razberemo, da je metoda SMM5 razred zase, saj je bistveno boljša od vseh ostalih metod. Nima nobenega signifikantnega poraza ter vsaj tri zmage proti ostalim metodam. Povprečno relativno izboljšanje je vsaj 10% (najmanjše je v primerjavi z metodo SCMLR). Z uporabo sedmih osnovnih klasifikatorjev (tabela 4.1) ima metoda SMM5 še vedno povprečno relativno izboljšanje nad ostalimi metodami večje od 7%. Vendar se razlika v številu zmag in porazov, glede na metodi SCMLR in SMLRE, zmanjša na ena (ob uporabi t-testa 1×10). Po drugi strani znaša razmerje v številu zmag in porazov za omenjeni metodi z uporabo t-testa 5×2 4:0, oziroma 2:0. Ob upoštevanju teh rezultatov lahko zaključimo, da nam metoda SMM5 nudi boljšo klasifikacijsko točnost kot metoda SELB in kot ostale metode za skladanje klasifikatorjev.

Vpliv števila osnovnih klasifikatorjev

Ob pregledu tabel 4.1 in 4.2 lahko opazimo, da ostaja metoda glasovanja (VOTE) najslabša ne glede na to ali uporabimo tri ali sedem osnovnih klasifikatorjev. Relativna uspešnost ostalih metod za kombiniranje je bolj odvisna od števila osnovnih klasifikatorjev. Z upoštevanjem skupne razlike med številom zmag in porazov lahko sklenemo, da najbolj poveča uspešnost metod SMDT (z 0 na +8) in SCMLR (z +2 na +10). Zmanjša se relativna uspešnost metode SMM5 (z +28 na +22), medtem ko se rezultat ostalih metod

praktično ne spremeni.

Zanimiv je tudi pogled na tabelo 13 v prilogi B. V njej so relativna izboljšanja natančnosti, če posamezna metoda namesto treh uporablja sedem osnovnih klasifikatorjev. Opazimo lahko, da je povprečno relativno izboljšanje natančnosti posameznih metod za skladanje dokaj majhno (največje je za glasovanje – 6%, za metode SCMLR, SMDT, SMLR in SELB je med 2 in 4%, za metodi SMLRE in SMM5 pa je manjše od 1%). Ob upoštevanju števila zmag in porazov se najbolj poveča uspešnost metod VOTE, SELB, SCMLR in SMLRE, najmanj pa uspešnost metode SMM5 (rezultat 1:0). Ob tem je zanimivo, da je metoda SMM5 s tremi osnovnimi klasifikatorji boljša od ostalih metod, ki uporabljajo sedem osnovnih klasifikatorjev. Razmerje v številu zmag in porazov je skoraj enako kot za metodo SMM5 s sedmimi osnovnimi klasifikatorji.

4.4 Komentar rezultatov

V večina primerjanih metod temelji na metodi skladanja z MLR (SMLR). Seewald [31] na osnovi empiričnih rezultatov trdi, da je uspešnost metode SMLR slabša na domenah z več razredi, kot na domenah z dvema razredoma. Kot verjeten razlog navaja število metaatributov, ki je pri tej metodi premosorazmerno s številom razredov ter predlaga zmanjšanje njihovega števila (metoda SCMLR).

V naših poskusih smo na število atributov metaučne množice vplivali na dva načina. Pri prvem smo povečali število osnovnih klasifikatorjev. Opazimo lahko, da se (relativna) uspešnost metode SCMLR s povečanjem števila osnovnih klasifikatorjev poveča. To je skladno s trditvijo Seewalda o problematičnosti velikega števila atributov metaučne množice.

Drug način povečanja števila metaatributov je uporaba metode SMLRE. Ob uporabi majhnega števila osnovnih klasifikatorjev prevlada vpliv dodatnih informacij o zanesljivosti napovedi. S povečanjem števila osnovnih klasifikatorjev pa se poveča negativni vpliv velikega števila metaatributov. Zaradi tega je prednost metode SMLRE pred metodo SMLR relativno majhna.

Po pričakovanju je metoda skladanja z večličnimi modelskimi drevesi (SMM5) uspešnejša od metode skladanja z večlično linearno regresijo (SMLR). Že Frank in sode-

lavci [18], ki so raziskali uporabo regresijskih metod za klasifikacijo, so ugotovili, da je metoda klasifikacije z modelskimi drevesi zelo dobra; boljša na primer kot klasifikacija z linearno regresijo in boljša kot odločitvena drevesa C5.0 (izboljšana različica C4.5 [28]), predvsem na domenah z zveznimi atributi. Glede na to, da so metaatributi verjetnosti in torej zvezni, so večlična modelska drevesa zelo primerna za učenje na metanivoju. To so potrdili tudi rezultati naših poskusov.

Rezultati kažejo, da je metoda SMM5 boljša, ne le od metode SMLR, ampak tudi od metode SCMLR. Vzrok prednosti metode SCMLR pred SMLR je v zmanjšanju števila metaatributov, vendar pa pri tem nekaj potencialno uporabne informacije zavržemo (ohranimo le verjetnosti za en razred). Kot kaže, predstavlja za modelska drevesa (M5') metode SMM5 veliko število metaatributov manjši problem kot za linearno regresijo metode SMLR. Ob tem pa modelska drevesa znajo koristno uporabiti informacijo, ki jo metoda SCMLR zavrže. Zaradi tega je klasifikacijska točnost metode SMM5 boljša od točnosti ostalih obravnavanih metod, vključno z metodo SMLRE.

4.5 Uporaba metod na praktičnih problemih

Do sedaj smo klasifikacijsko točnost obravnavanih metod ocenjevali le na množicah podatkov iz zbirke UCI [2]. Vendar smo želeli uspešnost metod za kombiniranje klasifikatorjev preizkusiti še na nekaj praktičnih problemih. Zbrali smo deset različic učnih množic predstavljenih v tabeli 4.3; njihovo problematiko bomo na kratko opisali v nadaljevanju.

4.5.1 Množice podatkov

- ACUTE_ABDOMINAL_PAIN_IN_CHILDREN. Po pregledu pacienta s simptomi akutne trebušne bolečine (ang. acute abdominal pain in children – AAPC) mora zdravnik sprejeti odločitev o nadaljnjem zdravljenju. Otroka lahko odpusti (če vzrok bolečine ni patološki), lahko ga napoti na operacijo ali pa se odloči za nadaljnje opazovanje, po izteku katerega bo pacientovo stanje ponovno ocenil. Množica podatkov je povzeta po [15] in vsebuje podatke o 335 pacientih. Stanje vsakega pacienta je opisano s 84. atributi, ki natančneje opisujejo simptome in rezultate laboratorijskih

Tabela 4.3: Uporabljene realne množice podatkov in njihove lastnosti: število primerov, razredov, (diskretnih/zveznih) atributov, verjetnost večinskega razreda in entropija verjetnostne porazdelitve razreda.

MNOŽICA PODATKOV	ŠT. PRIM.	ŠT. RAZ.	(D/Z)	ATR.	VER. VEČ.	ENT.
ACUTE_ABDOMINAL_PAIN	335	3	(83/1)	84	0.47	1.52
WATER	292	5	(0/80)	80	0.36	2.16
WATER+DIVERSITY	292	5	(0/86)	86	0.36	2.16
SEA_CUCUMBER	128	2	(1/9)	10	0.81	0.70
REUMA+BKGRD_KNOW	462	8	(16/6)	22	0.34	2.51
REUMA	462	8	(10/6)	16	0.34	2.51
BIODEG-FUNCTION_GROUP	328	4	(0/31)	31	0.37	1.90
BIODEG-FRAGMENTS	328	4	(0/61)	61	0.37	1.90
DIATOMA_VULGARE	1060	2	(0/9)	9	0.61	0.97
TRAFFIC	256	3	(1/3)	4	0.50	1.50

preiskav.

- WATER in WATER+DIVERSITY. Reke predstavljajo enega najpomembnejših vodnih virov, zato so za upravljanje z rečnimi vodami potrebni zanesljivi podatki o njihovi kvaliteti. Kvaliteta vode je odvisna od številnih fizikalnih, kemijskih in bioloških lastnosti. Slednje se odražajo v prisotnosti določenih vrst organizmov ter njihovi številčnosti. Na osnovi teh lastnosti lahko rečne vode razvrstimo v enega izmed petih kakovostnih razredov. Množici podatkov sta povzeti po [13] in vsebujeta podatke o 292 vodnih vzorcih iz angleških rek. Za vsak vzorec je podana številčna prisotnost (ocenjena s številom od 0 do 5) vsakega izmed 80 vodnih organizmov. Druga množica vsebuje še dodatnih šest atributov, ki opisujejo raznolikost prisotnih organizmov (npr. skupno število prisotnih družin organizmov).
- SEA_CUCUMBER. Tradicionalno predstavljajo morske kumare pomemben vir prehrane na tihoceanskih otokih. Naraščajoči vplivi človekovih dejavnosti na ekosisteme koralnih grebenov ter prekomeren lov lahko povzročijo zmanjšanje njihovega števila ali celo njihovo izumrtje. Eden pomembnejših pogojev za njihovo ohranitev je gotovo poznavanje njihovega habitata. Množica je povzeta po [11] in vsebuje podatke izmerjene na 128 lokacijah. Vsaka lokacija je opisana z desetimi spremenljivkami, kot so: lega na privetrni ali zavetrni strani otoka, odstotek površine, ki je

Tabela 4.4: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega ansambla uporabimo metodo skladanja klasifikatorjev z večličnimi modelskimi drevesi **s tremi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE
ACUTE_ABDOMINAL_PAIN	0.49 .	2.64 .	3.10 .	-2.01 .	-4.91 .	-0.25 .
WATER	6.41 +	7.83 +	5.62 +	1.13 .	1.99 .	-3.00 .
WATER+DIVERSITY	2.84 .	12.05 +	4.45 .	0.92 .	2.59 .	10.79 .
SEA_CUCUMBER	10.50 .	15.15 .	16.60 .	4.39 .	6.67 .	5.77 .
REUMA+BKGRD_KNOW	-1.33 .	-0.79 .	0.26 .	-1.28 .	-3.71 .	10.22 .
REUMA	-3.29 .	3.21 +	1.47 .	0.55 .	-2.04 .	6.70 .
BIODEG-FUNCTION_GROUP	-1.43 .	4.07 .	5.35 .	-1.22 .	-3.74 -	1.05 .
BIODEG-FRAGMENTS	1.36 .	4.04 .	2.61 .	-2.38 .	-4.71 .	0.43 .
DIATOMA_VULGARE	-1.99 .	2.65 .	1.41 .	-0.55 .	-0.55 .	-1.74 .
TRAFFIC	-5.04 .	4.89 .	5.19 .	-2.82 .	-3.91 .	0.00 .
POVPREČJE	0.95	5.69	4.71	-0.31	-1.16	3.11
ZMAGE/PORAZI	1:0	3:0	1:0	0:0	0:1	0:0

pokrita s peskom, s prodrom, odstotek živih in odstotek mrtvih koral in podobno. Za vsako lokacijo napovedujemo njeno primernost oziroma neprimernost za življenje morskih kumar vrste *Holothria leucospilota*, ki je določena glede na opaženo število teh živali.

- REUMA in REUMA+BACKGROUND_KNOWLEDGE. Postavitev pravilne diagnoze na začetni stopnji revmatičnega obolenja je zahteven problem, ker so si simptomi in laboratorijski izsledki različnih revmatičnih bolezni med seboj zelo podobni. Več kot dvesto različnih revmatičnih bolezni lahko razvrstimo v osem razredov. Množica podatkov je bila uporabljena v [26] in vsebuje podatke o 462 pacientih. Vsak pacient je opisan s 16. atributi, ki opisujejo njegovo stanje ter morebitno dosedanje zdravljenje. Prva množica vsebuje še dodatnih šest atributov, ki zajemajo zdravnikovo predznanje. Klasifikacijska naloga je, da na osnovi simptomov za diagnozo določimo eno od osmih skupin revmatičnih bolezni.
- BIODEG-FUNCTION_GROUP in BIODEG-FRAGMENTS. Za uporabo posameznih kemijskih spojin je zelo pomembna njihova biorazgradljivost, to je lastnost, da se

Tabela 4.5: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega ansambla uporabimo metodo skladanja klasifikatorjev z večličnimi modelskimi drevesi s **sedmimi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE
ACUTE_ABDOMINAL_PAIN	-11.80 .	16.15 .	14.25 .	-5.28 .	-5.87 .	4.05 .
WATER	12.34 .	11.21 .	12.14 +	1.55 .	2.31 .	5.00 .
WATER+DIVERSITY	8.23 .	14.10 .	12.76 .	0.41 .	-3.42 .	11.37 .
SEA_CUCUMBER	-2.11 .	-9.60 .	-7.78 .	3.48 .	2.02 -	3.00 .
REUMA+BKGRD_KNOW	-2.55 .	-1.92 .	0.38 .	-0.56 .	-5.81 .	13.62 .
REUMA	-2.79 .	2.07 .	1.09 .	-0.13 .	-2.74 .	13.33 .
BIODEG-FUNCTION_GROUP	-8.23 .	2.30 .	4.49 .	-2.77 .	-5.09 .	3.41 .
BIODEG-FRAGMENTS	-7.56 .	1.88 .	0.98 .	-1.22 .	-5.78 .	3.56 .
DIATOMA_VULGARE	-0.99 .	6.43 +	3.59 +	-2.12 .	-2.15 .	-2.67 .
TRAFFIC	-17.20 .	0.68 .	-1.74 .	-3.53 .	-4.64 .	6.39 +
POVPREČJE	-2.92	4.62	4.25	-0.99	-3.07	6.24
ZMAGE/PORAZI	0:0	1:0	2:0	0:0	0:1	1:0

(v interakciji z okoljem) razgradijo na okolju neškodljive snovi. Biorazgradljivost določene spojine je odvisna od njene strukture, podrobneje pa lahko te povezave analiziramo s pomočjo tako imenovanega QSAR (Quantitative Structure–Activity Relationships) modeliranja. Množici podatkov, ki smo ju uporabili, sta povzeti po [10] in vsebujeta 328 različnih spojin. Vsaka spojina je razvrščena v enega od štirih razredov glede na stopnjo biorazgradljivosti. Struktura spojin je v prvi množici opisana s številom vsake izmed 29 izbranih funkcijskih skupin, v drugi pa s številom linearnih fragmentov z dvema ali tremi atomi ter podstruktur zvezdaste topologije s štirimi atomi. Obema množicama sta dodana še atributa hidrofobičnost in molekulska teža spojine.

- DIATOMA_VULGARE. Podobno kot zgoraj omenjeni množici WATER imamo tudi pri tej množici opravka s kvaliteto rečne vode. V tem primeru je kvaliteta vode opisana z devetimi fizikalnimi in kemijskimi parametri (trdota vode, količina nekaterih dušikovih spojin in podobno), zanima pa nas prisotnost oziroma odsotnost alge *Diatoma vulgare*. Množica je bila uporabljena v [12] in vsebuje podatke o 1060

vzorcih rečne vode zbranih tekom šestih let.

- **TRAFFIC.** Za uspešno upravljanje cestnega prometa morajo imeti upravljalci na voljo čim več informacij o dejanskem stanju prometa. Na cestah so nameščeni senzorji, ki merijo hitrost vozil, pretok vozil in odstotek časa, ko vozila zasedajo senzor. Za vsak senzor tudi vemo kje, oziroma na kakšnem tipu ceste (avtocesta, dovoz ali izvoz) je nameščen. Na osnovi teh podatkov lahko napovemo, kakšno je stanje prometa na določenem cestnem odseku. Množica podatkov je povzeta po [14] in vsebuje 256 primerov opisanih z zgoraj omenjenimi atributi ter razvrščenih v enega od treh razredov (nesreča, zastoj ali nekritičen odsek).

4.5.2 Rezultati

Rezultati poskusov na omenjenih množicah so predstavljeni v tabelah 4.4 in 4.5 ter v prilogi C. Tabeli 4.4 in 4.5 tako prikazujeta relativno izboljšanje natančnosti (ARI), ki ga dobimo z uporabo metode skladanja z večličnimi modelskimi drevesi (SMM5) za tri in sedem osnovnih klasifikatorjev. V tabeli 14 so podane klasifikacijske napake osnovnih klasifikatorjev, v tabelah 15 in 16 pa klasifikacijske napake različnih ansamblov s tremi in s sedmimi osnovnimi klasifikatorji.

Za metodo skladanja z večličnimi modelskimi drevesi (SMM5) lahko ob uporabi treh osnovnih klasifikatorjev ugotovimo (glej tabelo 4.4), da je tudi na teh množicah podatkov boljša od metod SELB in SMDT (ARI okrog 5% ter tri, oziroma ena signifikantna zmaga). Prednost pred ostalimi metodami pa je zanemarljiva, za metodo SCMLR pa celo rahlo zaostaja (en poraz in ARI -1%). Pri kombiniranju sedmih osnovnih klasifikatorjev (tabela 4.5) je metoda (SMM5) še vedno boljša od metod SELB in SMDT in slabša od metode SCMLR (en poraz in ARI -3%). Z uporabo večjega števila osnovnih klasifikatorjev se nekoliko zmanjša relativna uspešnost metode SMLRE (prednost metode SMM5 glede na ARI se poveča s treh na šest odstotkov), kar je skladno z ugotovitvami, predstavljenimi na začetku tega poglavja. Zanimivo je, da je uspešnost kombiniranja z glasovanjem na teh množicah podatkov povsem primerljiva z ostalimi metodami.

Na osnovi pravkar predstavljenih rezultatov opazimo, da je pri izbranih množicah podatkov razlika med posameznimi metodami za kombiniranje manjša. Vzroke za to

lahko verjetno iščemo v lastnostih desetih uporabljenih množic podatkov, ki se precej razlikujejo od izbranih množic podatkov iz zbirke UCI. Tako je na prvi pogled opazna razlika v številu atributov in v številu razredov. Poskusi na teh desetih domenah potrjujejo odvisnost uspešnosti posameznih metod za kombiniranje od lastnosti uporabljene množice podatkov in izpostavljajo potrebo po dodatnih analizah teh povezav.

5. Sklep

V pričujočem magistrskem delu smo eksperimentalno ovrednotili in primerjali več metod za kombiniranje klasifikatorjev. Osredotočili smo se predvsem na metode za skladanje heterogenih klasifikatorjev in pokazali, da je njihova klasifikacijska točnost kvečjemu primerljiva s točnostjo metode izbire najboljšega klasifikatorja s prečnim preverjanjem. Na osnovi metode skladanja klasifikatorjev z verjetnostnimi porazdelitvami in večlično linearno regresijo, smo razvili dve novi metodi za skladanje klasifikatorjev. Prva uporablja razširjeni nabor metaatributov, druga pa kot učni algoritem na metanivoju uporablja večlična modelska drevesa. S poskusi na večjem številu klasifikacijskih problemov smo pokazali, da je klasifikacijska točnost druge metode boljša od točnosti ostalih metod skladanja klasifikatorjev, prav tako je boljša tudi od točnosti metode izbire najboljšega klasifikatorja s prečnim preverjanjem.

Poudariti moramo, da je predstavljeni pristop h kombiniranju klasifikatorjev namenjen kombiniranju *heterogenih* (zgrajenih z različnimi učnimi algoritmi) in *močnih* (točnost vsakega posameznega klasifikatorja je relativno dobra) klasifikatorjev. Nasprotno pa metoda skladanja klasifikatorjev, na primer, ni primerna za kombiniranje večjega števila klasifikatorjev zgrajenih z istim učnim algoritmom nad različnimi vzorci prvotne učne množice. Za grajenje takih ansamblov sta primernejši, na primer, metodi bagging in boosting. Razlika med metodami za kombiniranje heterogenih in metodami za kombiniranje homogenih klasifikatorjev je tudi v tem, da je teoretično ozadje slednjih precej bolj raziskano. Teoretično razumevanje delovanja metod za skladanje klasifikatorjev je še precej slabo, kar je gotovo ovira za nadaljnje izboljševanje teh metod, hkrati pa tudi izziv za nadaljnje delo na tem področju. Kljub temu lahko na osnovi predstavljenih poskusov sklenemo, da je metoda skladanja klasifikatorjev z večličnimi modelskimi drevesi dobra

izbira, ko hočemo doseči kar najboljšo klasifikacijsko točnost.

Med raziskavami, predstavljenimi na tem mestu in v dveh drugih prispevkih [41, 35], smo pregledali tudi literaturo s področja kombiniranja klasifikatorjev. Pri tem smo našli na nekaj trditev, ki so v nasprotju z rezultati naših poskusov. Vzroke teh razlik lahko najverjetneje iščemo v različnih metodologijah opravljanja poskusov ter mogoče tudi v uporabi različnih učnih množic. V naših poskusih smo uporabili veliko število množic podatkov ter skrbno izbrano metodologijo vrednotenja opravljenih poskusov. Predstavljeni rezultati nam tako nudijo boljši pregled nad in primerjavo med različnimi metodami skladenja klasifikatorjev.

Literatura

- [1] D. Aha, D. W. Kibler in M. K. Albert. Instance-based learning algorithms. *Machine Learning*, 6:37–66, 1991. [20](#)
- [2] C. L. Blake in C. J. Merz. UCI repository of machine learning databases, 1998. [14](#), [15](#), [31](#)
- [3] L. Breiman. Bagging predictors. *Machine Learning*, 24(2):123–140, 1996. [6](#)
- [4] L. Breiman. Arcing classifiers. *The Annals of Statistics*, 26(3):801–849, 1998. [8](#)
- [5] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001. [8](#)
- [6] J. G. Cleary in L. E. Trigg. K*: An instance-based learner using an entropic distance measure. V *Proceedings of the 12th International Conference on Machine Learning*, strani 108–114, San Francisco, 1995. Morgan Kaufmann. [20](#)
- [7] T. G. Dietterich. Machine-learning research: Four current directions. *AI Magazine*, 18(4):97–136, 1997. [4](#)
- [8] T. G. Dietterich. Approximate statistical test for comparing supervised classification learning algorithms. *Neural Computation*, 10(7):1895–1923, 1998. [17](#)
- [9] T. G. Dietterich. Ensemble methods in machine learning. V *Proceedings of the First International Workshop on Multiple Classifier Systems*, strani 1–15, Berlin, 2000. Springer. [4](#)
- [10] S. Džeroski, H. Blockeel, B. Kompare, S. Kramer, B. Pfahringer in W. Van Laer. Experiments in predicting biodegradability. V *Proceedings of the Ninth International Workshop on Inductive Logic Programming*, strani 80–91, Berlin, 1999. Springer. [34](#)

- [11] S. Džeroski in D. Drumm. Using machine learning to identify the habitant preference of the sea cucumber, holothria leucospilota on rarotonga, cook islands. V *Book of Abstracts of the Third European Ecological Modelling Conference*, stran 10, Zagreb, 2001. International Society for Ecological Modelling, European chapter. [32](#)
- [12] S. Džeroski, J. Grbović in H. Blockeel. Predicting river water comunicaties with logical decision trees. V *Book of Abstracts of the Third European Ecological Modelling Conference*, stran 11, Zagreb, 2001. International Society for Ecological Modelling, European chapter. [34](#)
- [13] S. Džeroski, J. Grbović in W. J. Walley. Machine learning applications in biological classification of river water quality. V R. S. Michalski, I. Bratko in M. Kubat, uredniki, *Machine Learning, Data Mining and Knowledge Discovery: Methods and Applications*, strani 429–448. John Wiley and Sons, Chichester, 1998. [32](#)
- [14] S. Džeroski, N. Jacobs, M. Molina, C. Moure, S. Muggleton in W. Van Laer. Detecting traffic problems with ILP. V *Proceedings of the Eighth International Conference on Inductive Logic Programming*, strani 281–290, Berlin, 1998. Springer. [35](#)
- [15] S. Džeroski, G. Potamias, V. Moustakis in G. Charissis. Automated revision of expert rules for treating acute abdominal pain in children. V *Proceedings of the Sixth European Conference on Artificial Intelligence in Medicine*, strani 98–109, Berlin, 1997. Springer. [31](#)
- [16] S. Džeroski in B. Ženko. Is combining classifiers better than selecting the best one? V *Proceedings of the Nineteenth International Conference on Machine Learning*, San Francisco, 2002. Morgan Kaufmann.
- [17] S. Džeroski in B. Ženko. Stacking with multi-response model trees. V *Multiple Classifiers Systems, Proceedings of the Third International Workshop*, Berlin, 2002. Springer. [26](#)
- [18] E. Frank, Y. Wang, S. Inglis, G. Holmes in I. H. Witten. Using model trees for classification. *Machine Learning*, 32(1):63–76, 1998. [26](#), [31](#)

- [19] Y. Freund in R. E. Schapire. Experiments with a new boosting algorithm. V *Proceedings of the Thirteenth International Conference on Machine Learning*, strani 148–156, San Francisco, 1996. Morgan Kaufmann. [7](#), [8](#), [18](#)
- [20] J. Gama. Combining classifiers by constructive induction. V *Proceedings of the Ninth European Conference on Machine Learning*, strani 178–189, 1998. [9](#)
- [21] M. Gams, M. Bohanec in B. Cestnik. A schema for using multiple knowledge. V S. J. Hanson, T. Petsche, M. Kearns in R. L. Rivest, uredniki, *Computational Learning Theory and Natural Learning Systems*, zvezek II, strani 157–170. MIT Press, Cambridge, Massachusetts, 1994. [4](#)
- [22] G. H. John in P. Langley. Estimating continuous distributions in bayesian classifiers. V *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence*, strani 338–345, San Francisco, 1995. Morgan Kaufmann. [20](#)
- [23] M. J. Kearns in U. V. Vazirani. *An introduction to computational learning theory*. MIT Press, Cambridge, 1994. [7](#)
- [24] R. Kohavi. The power of decision tables. V *Proceedings of the Eighth European Conference on Machine Learning*, strani 174–189, 1995. [20](#)
- [25] I. Kononenko. *Strojno učenje*. Založba FE in FRI, Ljubljana, 1997. [4](#)
- [26] N. Lavrač, S. Džeroski, V. Pirnat in V. Križman. The use of background knowledge in learning medical diagnostic rules. *Applied Artificial Intelligence*, 7(3):273–293, 1993. [33](#)
- [27] C. J. Merz. Using correspondence analysis to combine classifiers. *Machine Learning*, 36(1/2):33–58, 1999. [4](#), [11](#)
- [28] J. R. Quinlan. *C4.5: Programs for Machine Learning*. Morgan Kaufmann, San Francisco, 1993. [18](#), [20](#), [31](#)
- [29] C. Schaffer. Selecting a classification method by cross-validation. V *Proceedings of the Fourth International Workshop on Artificial Intelligence & Statistics*, strani 15–25, 1993. [9](#)

- [30] R. Schapire. The boosting approach to machine learning: An overview. V *MSRI Workshop on Nonlinear Estimation and Classification*, Berkeley, 2001. 7
- [31] A. K. Seewald. How to make stacking better and faster while also taking care of an unknown weakness. V *Proceedings of the Nineteenth International Conference on Machine Learning*, San Francisco, 2002. Morgan Kaufmann. 13, 30
- [32] A. K. Seewald in J. Fürnkranz. An evaluation of grading classifiers. V *Advances in Intelligent Data Analysis: Proceedings of the Fourth International Symposium (IDA-01)*, Berlin, 2001. Springer. 12
- [33] K. M. Ting in I. H. Witten. Issues in stacked generalization. *Journal of Artificial Intelligence Research*, 10:271–289, 1999. 11, 12, 20, 24, 25, 26
- [34] L. Todorovski in S. Džeroski. Combining multiple models with meta decision trees. V *Proceedings of the Fourth European Conference on Principles of Data Mining and Knowledge Discovery*, strani 54–64, Berlin, 2000. Springer. 12
- [35] L. Todorovski in S. Džeroski. Combining classifiers with meta decision trees. *Machine Learning*, 50(3):223–249, 2002. 12, 38
- [36] R. Vilalta in Y. Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2):77–95, 2002. 5
- [37] Y. Wang in I. H. Witten. Induction of model trees for predicting continuous classes. V *Proceedings of the Poster Papers of the European Conference on Machine Learning*, Prague, 1997. University of Economics, Faculty of Informatics and Statistics. 26
- [38] I. H. Witten in E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations*. Morgan Kaufmann, San Francisco, 1999. 18
- [39] D. Wolpert. Stacked generalization. *Neural Networks*, 5(2):241–260, 1992. 5, 9
- [40] B. Ženko in S. Džeroski. Stacking with an extended set of meta-level attributes and MLR. V *Proceedings of the Thirteenth European Conference on Machine Learning*, Berlin, 2002. Springer. 24

-
- [41] B. Ženko, L. Todorovski in S. Džeroski. A comparison of stacking with MDTs to bagging, boosting, and other stacking methods. V *Proceedings of the First IEEE International Conference on Data Mining*, strani 669–670, Los Alamitos, 2001. IEEE Computer Society. [12](#), [38](#)

Izjava

Izjavljam, da sem magistrsko delo izdelal samostojno na Odseku za inteligentne sisteme Instituta Jožef Stefan pod vodstvom somentorja doc. dr. Saša Džeroskega ter s pomočjo mentorja prof. dr. Ivana Bratka. Izkazano pomoč drugih sodelavcev sem v celoti navedel v zahvali.

Bernard Ženko

Priloge

A. Rezultati poskusov z obstoječimi metodami

Tabela 1: Klasifikacijske napake (v %) osnovnih klasifikatorjev.

MNOŽICA POD.	J4.8	IBK	NB	K*	KDE	DT	MLR
AUSTRALIAN	14.54	18.39	18.65	20.93	17.94	16.04	14.51
BALANCE	22.43	15.10	8.48	11.06	11.38	24.82	13.38
BREAST-CANCER	26.12	26.96	26.89	25.56	27.31	28.32	28.81
BREAST-W	5.39	4.28	2.69	4.33	4.43	9.10	3.86
BRIDGES-TD	14.71	16.86	14.02	15.20	16.67	17.75	18.63
CAR	7.44	5.83	14.40	12.55	5.79	5.57	16.26
CHESS	0.60	3.34	12.16	3.16	3.04	2.43	5.95
CONTRACEPTIVE	46.94	57.37	48.67	50.43	56.17	45.28	49.74
DIABETES	26.26	29.40	24.70	29.77	28.57	25.68	23.03
DIS	0.95	1.71	3.19	1.56	1.66	1.54	1.54
ECHO	34.12	42.29	27.33	36.18	36.41	36.41	28.47
GERMAN	28.82	27.62	25.43	29.76	30.14	28.30	24.23
GLASS	32.24	29.81	49.86	24.58	29.81	30.33	43.55
HEART-C	23.37	24.09	15.61	24.82	24.06	21.39	15.41
HEART-H	20.34	21.70	15.03	22.28	21.33	20.27	13.61
HEART	22.19	23.22	15.67	24.26	23.11	17.78	16.85
HEPATITIS	20.77	18.97	15.35	19.81	19.94	19.81	16.26
HYP0	0.72	2.93	1.81	2.07	2.55	0.86	3.62
IMAGE	3.18	2.84	14.29	2.88	2.72	7.77	16.76
IONOSPHERE	10.26	13.19	8.15	15.93	11.03	10.20	13.45
IRIS	5.33	4.80	4.07	5.33	4.73	7.00	15.73
SOLAR-FLARE-C	15.69	17.69	21.12	16.44	17.62	15.98	15.83
SOLAR-FLARE-M	4.90	5.70	9.16	5.46	5.73	4.95	4.90
SOLAR-FLARE-X	0.86	1.06	3.96	1.00	1.07	0.86	0.86
SONAR	26.68	13.51	27.84	15.19	14.13	27.74	27.55
SOYA	7.64	8.96	7.12	9.93	9.05	13.63	6.22
TIC-TAC-TOE	15.11	0.96	30.22	2.97	0.97	20.33	20.25
VOTE	3.54	7.36	9.82	6.67	7.45	5.43	4.37
WAVEFORM	23.62	22.63	19.24	23.28	25.58	27.22	14.04
WINE	6.57	4.66	2.64	1.29	4.66	6.40	1.24
POVPREČJE	15.71	15.77	16.59	15.49	15.50	16.64	15.96

Tabela 2: Klasifikacijske napake (v %) ansamblov klasifikatorjev. Metode za kombiniranje heterogenih klasifikatorjev uporabljajo **tri različne osnovne klasifikatorje**.

MNOŽICA POD.	BAGG	BOOS	RFOR	VOTE	SELB	SMDT	SMLR	SCMLR
AUSTRALIAN	13.67	15.58	13.13	14.17	14.74	14.30	14.28	14.28
BALANCE	17.31	21.49	18.32	13.41	8.48	8.48	10.16	9.46
BREAST-CANCER	26.89	33.25	30.35	25.31	28.15	28.04	27.73	27.45
BREAST-W	4.98	3.72	3.52	3.46	2.69	2.69	2.73	2.73
BRIDGES-TD	14.90	19.41	15.29	15.29	15.78	16.08	14.12	14.12
CAR	6.78	4.16	5.38	6.49	5.83	5.02	5.61	6.41
CHESS	0.61	0.38	0.78	1.43	0.60	0.60	0.60	0.60
CONTRACEPTIVE	46.90	49.33	48.80	47.78	47.33	48.19	45.91	45.93
DIABETES	24.62	28.53	24.05	24.22	25.09	24.87	23.71	23.70
DIS	1.09	1.04	1.33	1.33	0.95	0.95	0.96	0.96
ECHO	31.60	34.35	30.61	31.53	27.48	29.31	28.32	28.93
GERMAN	26.37	29.23	23.86	24.92	25.43	25.13	24.63	24.65
GLASS	26.03	23.18	20.75	29.21	31.92	31.03	30.84	29.49
HEART-C	20.07	20.83	18.51	18.28	15.61	15.64	17.59	16.30
HEART-H	20.58	20.82	19.22	16.63	15.03	15.27	16.80	14.73
HEART	19.78	21.78	17.85	18.22	15.89	15.89	15.74	15.48
HEPATITIS	17.74	18.45	16.26	17.94	15.87	16.45	15.81	16.13
HYP0	0.78	1.07	1.00	1.36	0.72	1.33	0.72	0.72
IMAGE	2.55	1.84	1.90	2.94	2.85	2.84	2.86	2.78
IONOSPHERE	7.83	6.41	6.41	7.15	8.40	8.60	7.35	7.35
IRIS	5.73	5.80	5.40	4.40	4.80	4.80	4.73	4.60
SOLAR-FLARE-C	15.70	16.36	18.29	16.16	15.69	15.69	15.91	15.69
SOLAR-FLARE-M	4.91	6.36	5.87	5.13	4.90	4.95	5.17	4.90
SOLAR-FLARE-X	0.86	1.31	1.15	1.00	0.87	0.87	0.90	0.86
SONAR	22.02	20.19	16.30	17.31	13.51	14.52	13.51	13.51
SOYA	7.22	7.19	6.73	6.71	7.22	7.13	7.07	6.75
TIC-TAC-TOE	6.80	3.43	3.27	9.24	0.96	0.96	0.58	0.58
VOTE	3.93	4.76	3.52	6.90	3.54	3.54	3.54	3.54
WAVEFORM	18.00	18.58	15.13	18.42	19.24	17.64	16.97	16.97
WINE	5.11	4.04	2.08	1.74	2.92	2.92	2.87	2.75
POVPREČJE	14.05	14.76	13.17	13.60	12.75	12.79	12.59	12.41

Tabela 3: Klasifikacijske napake (v %) ansamblov klasifikatorjev. Metode za kombiniranje heterogenih klasifikatorjev uporabljajo **sedem različnih osnovnih klasifikatorjev**.

MNOŽICA POD.	BAGG	BOOS	RFOR	VOTE	SELB	SMDT	SMLR	SCMLR
AUSTRALIAN	13.67	15.58	13.13	13.99	14.84	15.06	13.97	14.03
BALANCE	17.31	21.49	18.32	10.14	8.48	8.45	10.51	9.84
BREAST-CANCER	26.89	33.25	30.35	25.77	28.08	28.29	27.94	27.97
BREAST-W	4.98	3.72	3.52	3.65	2.69	2.69	2.72	2.72
BRIDGES-TD	14.90	19.41	15.29	15.39	16.37	17.25	15.59	15.49
CAR	6.78	4.16	5.38	6.73	5.69	3.72	4.24	5.00
CHESS	0.61	0.38	0.78	1.59	0.60	0.60	0.60	0.60
CONTRACEPTIVE	46.90	49.33	48.80	47.26	45.95	47.64	44.94	44.81
DIABETES	24.62	28.53	24.05	24.10	23.11	24.27	23.70	23.61
DIS	1.09	1.04	1.33	1.33	0.96	0.97	1.00	0.99
ECHO	31.60	34.35	30.61	30.92	28.63	30.46	29.54	27.94
GERMAN	26.37	29.23	23.86	24.08	24.67	24.29	23.20	23.13
GLASS	26.03	23.18	20.75	25.79	25.19	25.56	24.63	22.71
HEART-C	20.07	20.83	18.51	18.22	15.94	16.01	18.25	16.04
HEART-H	20.58	20.82	19.22	16.60	14.39	14.25	16.60	15.20
HEART	19.78	21.78	17.85	17.26	16.15	16.63	16.04	15.96
HEPATITIS	17.74	18.45	16.26	16.39	16.39	16.77	16.84	17.29
HYPO	0.78	1.07	1.00	1.56	0.76	1.35	0.77	0.77
IMAGE	2.55	1.84	1.90	1.92	3.03	2.46	2.02	1.94
IONOSPHERE	7.83	6.41	6.41	8.52	8.52	8.77	7.12	7.09
IRIS	5.73	5.80	5.40	5.00	4.80	4.80	4.93	4.53
SOLAR-FLARE-C	15.70	16.36	18.29	16.10	15.75	15.76	16.34	15.73
SOLAR-FLARE-M	4.91	6.36	5.87	5.28	4.93	4.99	5.20	4.92
SOLAR-FLARE-X	0.86	1.31	1.15	1.00	0.88	0.88	0.89	0.86
SONAR	22.02	20.19	16.30	15.29	14.71	14.86	14.52	14.28
SOYA	7.22	7.19	6.73	6.71	6.22	6.34	7.36	6.66
TIC-TAC-TOE	6.80	3.43	3.27	3.58	0.96	0.96	0.64	0.49
VOTE	3.93	4.76	3.52	6.25	3.93	3.93	3.75	3.72
WAVEFORM	18.00	18.58	15.13	16.64	14.04	13.85	15.65	16.08
WINE	5.11	4.04	2.08	1.46	2.30	2.19	2.08	2.02
POVPREČJE	14.05	14.76	13.17	12.95	12.30	12.47	12.39	12.08

Tabela 4: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega klasifikatorja uporabimo metodo skladanja klasifikatorjev z MLR s **tremi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	J4.8	IBK	NB	K*	KDE	DT	MLR
AUSTRALIAN	1.79 .	22.38 +	23.47 .	31.79 +	20.44 .	11.02 .	1.60 .
BALANCE	54.71 +	32.73 +	-19.81 -	8.10 .	10.69 .	59.06 +	24.04 +
BREAST-CANCER	-6.16 .	-2.85 .	-3.12 .	-8.48 .	-1.54 .	2.10 .	3.76 .
BREAST-W	49.34 +	36.12 +	-1.60 .	36.96 +	38.39 +	69.97 +	29.26 .
BRIDGES-TD	4.00 .	16.28 .	-0.70 .	7.10 .	15.29 .	20.44 +	24.21 .
CAR	24.57 +	3.77 .	61.01 +	55.28 +	3.00 .	-0.73 .	65.47 +
CHESS	0.00 .	82.10 +	95.08 +	81.11 +	80.35 +	75.45 +	89.96 +
CONTRACEPTIVE	2.18 .	19.97 +	5.66 .	8.96 +	18.26 +	-1.39 .	7.70 +
DIABETES	9.72 .	19.35 +	4.01 .	20.34 +	17.00 +	7.66 +	-2.94 .
DIS	-1.40 .	43.79 +	69.91 +	38.54 +	42.08 +	37.80 +	37.58 +
ECHO	17.00 .	33.03 +	-3.63 .	21.73 +	22.22 +	22.22 .	0.54 .
GERMAN	14.54 .	10.83 +	3.15 .	17.24 +	18.28 +	12.97 .	-1.65 .
GLASS	4.35 .	-3.45 .	38.14 +	-25.48 .	-3.45 .	-1.69 .	29.18 +
HEART-C	24.72 .	26.99 .	-12.68 .	29.12 .	26.89 .	17.75 .	-14.13 .
HEART-H	17.39 +	22.57 +	-11.76 .	24.58 +	21.21 +	17.11 +	-23.50 -
HEART	29.05 .	32.22 +	-0.47 .	35.11 +	31.89 +	11.46 .	6.59 .
HEPATITIS	23.91 +	16.67 .	-2.94 .	20.20 .	20.71 .	20.20 +	2.78 .
HYPO	0.00 .	75.32 +	60.10 +	64.98 +	71.62 +	15.81 +	79.98 +
IMAGE	10.20 .	-0.76 .	80.01 +	0.75 .	-5.10 .	63.23 +	82.95 +
IONOSPHERE	28.33 +	44.28 +	9.79 .	53.85 +	33.33 .	27.93 .	45.34 +
IRIS	11.25 .	1.39 .	-16.39 .	11.25 .	0.00 .	32.38 .	69.92 +
SOLAR-FLARE-C	-1.38 .	10.05 +	24.65 +	3.24 +	9.69 +	0.45 .	-0.50 .
SOLAR-FLARE-M	-5.59 .	9.34 +	43.55 +	5.28 +	9.80 +	-4.51 .	-5.59 .
SOLAR-FLARE-X	-4.17 .	14.97 +	77.27 +	10.07 .	16.11 +	-4.17 .	-4.17 .
SONAR	49.37 +	0.00 .	51.47 +	11.08 .	4.42 .	51.30 +	50.96 +
SOYA	7.47 .	21.08 +	0.62 .	28.76 +	21.84 +	48.12 +	-13.65 .
TIC-TAC-TOE	96.13 +	39.13 .	98.07 +	80.35 +	39.79 .	97.13 +	97.11 +
VOTE	0.00 .	51.87 +	63.93 +	46.90 +	52.47 +	34.75 +	18.95 .
WAVEFORM	28.18 +	25.02 +	11.83 +	27.13 +	33.66 +	37.68 +	-20.81 -
WINE	56.41 +	38.55 .	-8.51 .	-121.74 .	38.55 .	55.26 +	-131.82 .
POVPREČJE	25.83	28.92	41.45	28.52	27.36	37.18	35.84
ZMAGE/PORAZI	10:0	19:0	13:1	19:0	16:0	16:0	12:2

Tabela 5: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega klasifikatorja uporabimo metodo skladanja klasifikatorjev z MLR s **sedmimi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	J4.8	IBK	NB	K*	KDE	DT	MLR
AUSTRALIAN	3.89 .	24.03 +	25.10 +	33.24 +	22.13 .	12.92 .	3.70 .
BALANCE	53.14 +	30.40 +	-23.96 -	4.92 .	7.59 .	57.64 +	21.41 .
BREAST-CANCER	-6.96 .	-3.63 .	-3.90 .	-9.30 -	-2.30 .	1.36 .	3.03 .
BREAST-W	49.60 +	36.45 +	-1.06 .	37.29 +	38.71 +	70.13 +	29.63 +
BRIDGES-TD	-6.00 .	7.56 .	-11.19 .	-2.58 .	6.47 .	12.15 .	16.32 .
CAR	43.08 +	27.38 .	70.58 +	66.25 +	26.80 .	23.99 +	73.94 +
CHESS	0.00 .	82.10 +	95.08 +	81.11 +	80.35 +	75.45 +	89.96 +
CONTRACEPTIVE	4.25 +	21.67 +	7.66 +	10.89 +	19.99 +	0.75 .	9.65 +
DIABETES	9.77 +	19.40 +	4.06 .	20.38 +	17.05 +	7.71 +	-2.88 .
DIS	-5.32 .	41.61 +	68.75 +	36.16 +	39.84 +	35.40 +	35.17 +
ECHO	13.42 .	30.14 +	-8.10 .	18.35 +	18.87 +	18.87 .	-3.75 .
GERMAN	19.50 +	16.00 +	8.77 +	22.04 +	23.03 +	18.02 .	4.25 .
GLASS	23.62 .	17.40 .	50.61 +	-0.19 .	17.40 +	18.80 .	43.46 +
HEART-C	21.89 .	24.25 .	-16.91 .	26.46 .	24.14 .	14.66 .	-18.42 .
HEART-H	18.39 +	23.51 +	-10.41 .	25.50 +	22.17 +	18.12 +	-22.00 .
HEART	27.71 .	30.94 +	-2.36 .	33.89 +	30.61 +	9.79 .	4.84 .
HEPATITIS	18.94 .	11.22 .	-9.66 .	14.98 .	15.53 .	14.98 .	-3.57 .
HYPO	-6.11 .	73.81 +	57.67 +	62.84 +	69.89 +	10.66 .	78.76 +
IMAGE	36.46 +	28.70 +	85.86 +	29.77 +	25.64 +	73.98 +	87.94 +
IONOSPHERE	30.56 +	46.00 +	12.59 .	55.28 +	35.40 +	30.17 .	47.03 +
IRIS	7.50 .	-2.78 .	-21.31 .	7.50 .	-4.23 .	29.52 +	68.64 +
SOLAR-FLARE-C	-4.13 .	7.61 +	22.60 +	0.61 .	7.23 +	-2.25 .	-3.23 .
SOLAR-FLARE-M	-6.18 .	8.84 .	43.24 +	4.75 .	9.30 .	-5.10 .	-6.18 .
SOLAR-FLARE-X	-3.33 .	15.65 +	77.46 +	10.79 .	16.78 +	-3.33 .	-3.33 .
SONAR	45.59 +	-7.47 .	47.84 +	4.43 .	-2.72 .	47.66 +	47.29 +
SOYA	3.64 .	17.81 .	-3.50 .	25.81 +	18.61 .	45.97 +	-18.35 -
TIC-TAC-TOE	95.79 +	33.70 .	97.89 +	78.60 +	34.41 .	96.87 +	96.86 +
VOTE	-5.84 .	49.06 +	61.83 +	43.79 +	49.69 +	30.93 .	14.21 .
WAVEFORM	33.77 +	30.86 +	18.69 +	32.80 +	38.83 +	42.52 +	-11.41 -
WINE	68.38 +	55.42 .	21.28 .	-60.87 .	55.42 .	67.54 +	-68.18 .
POVPREČJE	27.52	30.53	42.77	30.13	29.01	38.60	37.30
ZMAGE/PORAZI	13:0	18:0	16:1	18:1	17:0	16:0	12:2

Tabela 6: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega ansambla uporabimo metodo skladanja klasifikatorjev z MLR s **tremi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	BAGG	BOOS	RFOR	VOTE	SELB	SMDT	SCMLR
AUSTRALIAN	-4.45 .	8.37 .	-8.72 .	-0.72 .	3.15 .	0.20 .	0.00 .
BALANCE	41.31 +	52.72 +	44.54 +	24.22 +	-19.81 -	-19.81 -	-7.45 -
BREAST-CANCER	-3.12 .	16.61 .	8.64 .	-9.53 .	1.49 .	1.12 .	-1.02 .
BREAST-W	45.11 +	26.54 .	22.36 .	21.07 .	-1.60 .	-1.60 .	-0.00 .
BRIDGES-TD	5.26 .	27.27 .	7.69 .	7.69 .	10.56 .	12.20 .	0.00 .
CAR	17.17 .	-35.10 -	-4.41 .	13.55 .	3.77 .	-11.75 -	12.46 +
CHESS	1.55 .	-56.55 .	23.60 .	58.20 +	0.00 .	0.00 .	0.00 .
CONTRACEPTIVE	2.11 .	6.92 +	5.91 .	3.91 .	2.98 .	4.73 .	0.03 .
DIABETES	3.70 .	16.89 +	1.41 .	2.10 .	5.50 .	4.66 .	-0.05 .
DIS	11.92 .	7.65 .	27.74 +	27.74 +	-1.40 .	-0.84 .	0.00 .
ECHO	10.39 .	17.56 .	7.48 .	10.17 .	-3.06 .	3.39 .	2.11 .
GERMAN	6.60 .	15.74 +	-3.23 .	1.16 .	3.15 .	1.99 .	0.08 .
GLASS	-18.49 .	-33.06 -	-48.65 -	-5.60 .	3.37 .	0.60 .	-4.60 .
HEART-C	12.34 .	15.53 .	4.99 .	3.79 .	-12.68 .	-12.45 .	-7.89 .
HEART-H	18.35 +	19.28 +	12.57 .	-1.02 .	-11.76 .	-10.02 .	-14.09 .
HEART	20.41 .	27.72 .	11.83 .	13.62 .	0.93 .	0.93 .	-1.67 .
HEPATITIS	10.91 .	14.34 .	2.78 .	11.87 .	0.41 .	3.92 .	2.00 .
HYP0	7.29 .	32.45 +	27.53 .	46.62 +	0.00 .	45.48 +	0.00 .
IMAGE	-12.05 .	-55.30 -	-50.34 .	2.65 .	-0.15 .	-0.61 .	-2.64 .
IONOSPHERE	6.18 .	-14.67 .	-14.67 .	-2.79 .	12.54 .	14.57 .	0.00 .
IRIS	17.44 .	18.39 .	12.35 .	-7.58 .	1.39 .	1.39 .	-2.90 .
SOLAR-FLARE-C	-1.33 .	2.77 +	13.03 +	1.56 .	-1.38 .	-1.38 .	-1.38 .
SOLAR-FLARE-M	-5.28 .	18.69 +	11.90 .	-0.84 .	-5.59 .	-4.36 .	-5.59 .
SOLAR-FLARE-X	-4.17 .	31.32 .	21.88 .	10.07 .	-3.31 .	-3.31 .	-4.17 .
SONAR	38.65 +	33.10 +	17.11 .	21.94 .	0.00 .	6.95 .	0.00 .
SOYA	2.03 .	1.63 .	-5.00 .	-5.46 .	2.03 .	0.82 .	-4.77 .
TIC-TAC-TOE	91.40 +	82.98 +	82.11 +	93.67 +	39.13 .	39.13 .	0.00 .
VOTE	9.94 .	25.60 .	-0.65 .	48.67 +	0.00 .	0.00 .	0.00 .
WAVEFORM	5.75 .	8.71 +	-12.11 .	7.87 +	11.83 +	3.80 .	0.00 .
WINE	43.96 +	29.17 +	-37.84 .	-64.52 .	1.92 .	1.92 .	-4.08 .
POVPREČJE	18.14	17.37	10.67	18.50	1.98	3.72	-0.48
ZMAGE/PORAZI	6:0	12:3	4:1	7:0	1:1	1:2	1:1

Tabela 7: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega ansambla uporabimo metodo skladanja klasifikatorjev z MLR s **sedmimi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	BAGG	BOOS	RFOR	VOTE	SELB	SMDT	SCMLR
AUSTRALIAN	-2.23 .	10.33 .	-6.40 .	0.10 .	5.86 .	7.22 .	0.41 .
BALANCE	39.28 +	51.08 +	42.62 +	-3.63 .	-23.96 -	-24.43 -	-6.83 -
BREAST-CANCER	-3.90 .	15.98 .	7.95 .	-8.41 .	0.50 .	1.24 .	0.12 .
BREAST-W	45.40 +	26.92 .	22.76 .	25.49 +	-1.06 .	-1.06 .	-0.00 .
BRIDGES-TD	-4.61 .	19.70 .	-1.92 .	-1.27 .	4.79 .	9.66 .	-0.63 .
CAR	37.49 +	-1.95 .	21.21 .	37.06 +	25.53 +	-13.84 -	15.28 +
CHESS	1.55 .	-56.55 .	23.60 .	62.48 +	0.00 .	0.00 .	0.00 .
CONTRACEPTIVE	4.18 .	8.89 +	7.90 .	4.90 +	2.20 .	5.66 .	-0.29 .
DIABETES	3.75 .	16.93 +	1.46 .	1.67 +	-2.54 .	2.36 .	-0.39 .
DIS	8.52 .	4.08 .	24.95 .	24.95 +	-3.30 .	-2.73 .	-0.54 .
ECHO	6.52 .	14.00 .	3.49 .	4.44 .	-3.20 .	3.01 .	-5.74 .
GERMAN	12.02 +	20.63 +	2.77 .	3.65 .	5.96 .	4.49 .	-0.30 .
GLASS	5.39 .	-6.25 .	-18.69 -	4.53 .	2.23 .	3.66 .	-8.44 .
HEART-C	9.05 .	12.36 .	1.43 .	-0.18 .	-14.49 .	-14.02 .	-13.79 .
HEART-H	19.34 +	20.26 +	13.63 .	0.00 .	-15.37 .	-16.47 .	-9.17 .
HEART	18.91 .	26.36 .	10.17 .	7.08 .	0.69 .	3.56 .	-0.46 .
HEPATITIS	5.09 .	8.74 .	-3.57 .	-2.76 .	-2.76 .	-0.38 .	2.61 .
HYPO	1.62 .	28.32 +	23.10 .	50.71 +	-1.67 .	43.09 +	0.00 .
IMAGE	20.71 +	-9.88 .	-6.38 .	-5.18 .	33.19 +	17.78 .	-4.24 .
IONOSPHERE	9.09 .	-11.11 .	-11.11 .	16.39 .	16.39 .	18.83 .	-0.40 .
IRIS	13.95 .	14.94 .	8.64 .	1.33 .	-2.78 .	-2.78 .	-8.82 .
SOLAR-FLARE-C	-4.08 .	0.13 .	10.67 +	-1.52 .	-3.75 .	-3.70 .	-3.89 .
SOLAR-FLARE-M	-5.87 .	18.23 +	11.41 .	1.63 .	-5.40 .	-4.19 .	-5.71 .
SOLAR-FLARE-X	-3.33 .	31.87 .	22.50 .	10.79 .	-1.64 .	-1.64 .	-3.33 .
SONAR	34.06 +	28.10 .	10.91 .	5.03 +	1.31 .	2.27 .	-1.68 .
SOYA	-2.03 .	-2.44 .	-9.35 .	-9.83 .	-18.35 -	-16.17 -	-10.55 -
TIC-TAC-TOE	90.63 +	81.46 +	80.51 +	82.22 +	33.70 .	33.70 .	-29.79 .
VOTE	4.68 .	21.26 .	-6.54 .	40.07 +	4.68 .	4.68 .	-0.62 .
WAVEFORM	13.09 +	15.81 +	-3.38 .	5.98 +	-11.41 -	-12.98 -	2.71 .
WINE	59.34 +	48.61 .	0.00 .	-42.31 .	9.76 .	5.13 .	-2.78 .
POVPREČJE	20.00	19.24	12.69	15.37	2.09	2.82	-3.01
ZMAGE/PORAZI	10:0	9:0	3:1	11:0	2:3	1:4	1:2

Tabela 8: Relativno izboljšanje klasifikacijske natančnosti (v %), če v ansamblih uporabimo sedem osnovnih klasifikatorjev namesto treh.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR
AUSTRALIAN	1.33 .	-0.69 .	-5.27 .	2.13 .	1.73 .
BALANCE	24.34 +	0.00 .	0.38 .	-3.46 .	-4.06 -
BREAST-CANCER	-1.80 .	0.25 .	-0.87 .	-0.76 .	-1.91 .
BREAST-W	-5.37 -	0.00 .	0.00 .	0.52 .	0.52 .
BRIDGES-TD	-0.64 .	-3.73 .	-7.32 .	-10.42 .	-9.72 .
CAR	-3.65 .	2.48 .	25.92 +	24.54 +	22.02 +
CHESS	-11.38 .	0.00 .	0.00 .	0.00 .	0.00 .
CONTRACEPTIVE	1.09 .	2.90 +	1.16 .	2.11 .	2.42 +
DIABETES	0.48 .	7.89 .	2.41 .	0.05 .	0.38 .
DIS	-0.00 .	-1.96 .	-1.95 .	-3.87 .	-3.31 .
ECHO	1.94 .	-4.17 .	-3.91 .	-4.31 .	3.43 .
GERMAN	3.37 .	2.99 .	3.34 .	5.81 .	6.17 .
GLASS	11.68 .	21.08 .	17.62 .	20.15 .	22.98 .
HEART-C	0.36 .	-2.11 .	-2.32 .	-3.75 .	1.62 .
HEART-H	0.20 .	4.30 +	6.68 .	1.21 .	-3.23 .
HEART	5.28 .	-1.63 .	-4.66 .	-1.88 .	-3.11 .
HEPATITIS	8.63 .	-3.25 .	-1.96 .	-6.53 -	-7.20 .
HYPO	-14.92 .	-4.37 .	-1.67 .	-6.11 .	-6.11 .
IMAGE	34.51 +	-6.07 .	13.41 .	29.24 +	30.33 +
IONOSPHERE	-19.12 .	-1.36 .	-1.99 .	3.10 .	3.49 .
IRIS	-13.64 .	-0.00 .	-0.00 .	-4.23 .	1.45 .
SOLAR-FLARE-C	0.40 .	-0.37 .	-0.41 .	-2.72 -	-0.23 .
SOLAR-FLARE-M	-3.09 .	-0.74 .	-0.73 .	-0.56 .	-0.44 .
SOLAR-FLARE-X	0.00 .	-0.83 .	-0.83 .	0.80 .	0.00 .
SONAR	11.67 .	-8.90 .	-2.32 .	-7.47 .	-5.69 .
SOYA	0.00 .	13.79 .	0.75 .	-3.18 .	1.30 .
TIC-TAC-TOE	61.24 +	0.00 .	0.00 .	-8.93 .	16.07 .
VOTE	9.33 .	-11.04 .	-11.04 .	-5.84 .	-5.19 .
WAVEFORM	9.63 +	27.01 +	21.48 +	7.78 +	5.21 +
WINE	16.13 .	21.15 .	25.00 .	27.45 .	26.53 .
POVPREČJE	5.88	2.16	2.82	2.29	3.76
ZMAGE/PORAZI	4:1	3:0	2:0	3:2	4:1

B. Rezultati poskusov z izboljšanimi metodami

Tabela 9: Klasifikacijske napake (v %) ansamblov klasifikatorjev. Metode za kombiniranje heterogenih klasifikatorjev uporabljajo **tri različne osnovne klasifikatorje**.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE	SMM5
AUSTRALIAN	14.17	14.74	14.30	14.28	14.28	15.71	14.45
BALANCE	13.41	8.48	8.48	10.16	9.46	6.61	4.59
BREAST-CANCER	25.31	28.15	28.04	27.73	27.45	27.52	27.62
BREAST-W	3.46	2.69	2.69	2.73	2.73	2.58	2.79
BRIDGES-TD	15.29	15.78	16.08	14.12	14.12	14.41	14.31
CAR	6.49	5.83	5.02	5.61	6.41	4.11	1.53
CHESS	1.43	0.60	0.60	0.60	0.60	0.60	0.61
CONTRACEPTIVE	47.78	47.33	48.19	45.91	45.93	46.23	46.07
DIABETES	24.22	25.09	24.87	23.71	23.70	24.22	23.78
DIS	1.33	0.95	0.95	0.96	0.96	0.95	1.04
ECHO	31.53	27.48	29.31	28.32	28.93	27.71	28.17
GERMAN	24.92	25.43	25.13	24.63	24.65	24.83	25.02
GLASS	29.21	31.92	31.03	30.84	29.49	34.30	31.12
HEART-C	18.28	15.61	15.64	17.59	16.30	19.97	16.34
HEART-H	16.63	15.03	15.27	16.80	14.73	18.64	15.10
HEART	18.22	15.89	15.89	15.74	15.48	18.63	16.00
HEPATITIS	17.94	15.87	16.45	15.81	16.13	16.39	15.29
HYP0	1.36	0.72	1.33	0.72	0.72	0.72	0.76
IMAGE	2.94	2.85	2.84	2.86	2.78	2.80	2.80
IONOSPHERE	7.15	8.40	8.60	7.35	7.35	7.44	7.83
IRIS	4.40	4.80	4.80	4.73	4.60	4.93	4.40
SOLAR-FLARE-C	16.16	15.69	15.69	15.91	15.69	16.25	15.99
SOLAR-FLARE-M	5.13	4.90	4.95	5.17	4.90	5.36	5.08
SOLAR-FLARE-X	1.00	0.87	0.87	0.90	0.86	0.95	0.87
SONAR	17.31	13.51	14.52	13.51	13.51	13.89	13.65
SOYA	6.71	7.22	7.13	7.07	6.75	7.20	6.37
TIC-TAC-TOE	9.24	0.96	0.96	0.58	0.58	1.39	0.26
VOTE	6.90	3.54	3.54	3.54	3.54	3.56	3.43
WAVEFORM	18.42	19.24	17.64	16.97	16.97	15.80	14.78
WINE	1.74	2.92	2.92	2.87	2.75	3.31	2.08
POVPREČJE	13.60	12.75	12.79	12.59	12.41	12.90	12.07

Tabela 10: Klasifikacijske napake (v %) ansamblov klasifikatorjev. Metode za kombiniranje heterogenih klasifikatorjev uporabljajo **sedem različnih osnovnih klasifikatorjev**.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE	SMM5
AUSTRALIAN	13.99	14.84	15.06	13.97	14.03	17.10	14.32
BALANCE	10.14	8.48	8.45	10.51	9.84	5.78	5.09
BREAST-CANCER	25.77	28.08	28.29	27.94	27.97	28.57	28.01
BREAST-W	3.65	2.69	2.69	2.72	2.72	2.59	2.72
BRIDGES-TD	15.39	16.37	17.25	15.59	15.49	15.69	15.78
CAR	6.73	5.69	3.72	4.24	5.00	3.52	1.24
CHESS	1.59	0.60	0.60	0.60	0.60	0.60	0.62
CONTRACEPTIVE	47.26	45.95	47.64	44.94	44.81	45.69	45.57
DIABETES	24.10	23.11	24.27	23.70	23.61	24.01	24.02
DIS	1.33	0.96	0.97	1.00	0.99	1.03	1.07
ECHO	30.92	28.63	30.46	29.54	27.94	30.23	30.38
GERMAN	24.08	24.67	24.29	23.20	23.13	23.68	23.24
GLASS	25.79	25.19	25.56	24.63	22.71	30.56	25.19
HEART-C	18.22	15.94	16.01	18.25	16.04	19.64	16.20
HEART-H	16.60	14.39	14.25	16.60	15.20	17.07	16.05
HEART	17.26	16.15	16.63	16.04	15.96	17.70	16.04
HEPATITIS	16.39	16.39	16.77	16.84	17.29	17.94	16.45
HYP0	1.56	0.76	1.35	0.77	0.77	0.76	0.79
IMAGE	1.92	3.03	2.46	2.02	1.94	1.95	2.06
IONOSPHERE	8.52	8.52	8.77	7.12	7.09	7.66	7.52
IRIS	5.00	4.80	4.80	4.93	4.53	6.20	4.80
SOLAR-FLARE-C	16.10	15.75	15.76	16.34	15.73	16.82	16.17
SOLAR-FLARE-M	5.28	4.93	4.99	5.20	4.92	5.54	5.13
SOLAR-FLARE-X	1.00	0.88	0.88	0.89	0.86	0.98	0.90
SONAR	15.29	14.71	14.86	14.52	14.28	17.55	15.29
SOYA	6.71	6.22	6.34	7.36	6.66	7.44	6.16
TIC-TAC-TOE	3.58	0.96	0.96	0.64	0.49	1.30	0.25
VOTE	6.25	3.93	3.93	3.75	3.72	3.98	3.63
WAVEFORM	16.64	14.04	13.85	15.65	16.08	14.45	13.56
WINE	1.46	2.30	2.19	2.08	2.02	2.25	1.97
POVPREČJE	12.95	12.30	12.47	12.39	12.08	12.94	12.01

Tabela 11: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega ansambla uporabimo metodo skladanja klasifikatorjev z večličnimi modelskimi drevesi s **tremi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE
AUSTRALIAN	-1.94 .	1.97 .	-1.01 .	-1.22 .	-1.22 .	8.03 .
BALANCE	65.75 +	45.85 +	45.85 +	54.80 +	51.44 +	30.51 +
BREAST-CANCER	-9.12 .	1.86 .	1.50 .	0.38 .	-0.64 .	-0.38 .
BREAST-W	19.42 .	-3.72 .	-3.72 .	-2.09 .	-2.09 .	-8.33 .
BRIDGES-TD	6.41 .	9.32 .	10.98 .	-1.39 .	-1.39 .	0.68 .
CAR	76.38 +	73.71 +	69.47 +	72.68 +	76.08 +	62.73 +
CHESS	57.55 +	-1.57 .	-1.57 .	-1.57 .	-1.57 .	-1.57 .
CONTRACEPTIVE	3.58 .	2.65 .	4.41 .	-0.34 .	-0.31 .	0.35 .
DIABETES	1.83 .	5.24 .	4.40 .	-0.27 .	-0.33 .	1.83 .
DIS	21.96 +	-9.52 .	-8.91 .	-8.01 .	-8.01 .	-8.61 .
ECHO	10.65 .	-2.50 .	3.91 .	0.54 .	2.64 .	-1.65 .
GERMAN	-0.40 .	1.61 .	0.44 .	-1.58 .	-1.50 .	-0.77 .
GLASS	-6.56 .	2.49 .	-0.30 .	-0.91 .	-5.55 .	9.26 .
HEART-C	10.65 .	-4.65 .	-4.43 .	7.13 .	-0.20 .	18.18 .
HEART-H	9.20 .	-0.45 .	1.11 .	10.12 .	-2.54 .	18.98 .
HEART	12.20 .	-0.70 .	-0.70 .	-1.65 .	-3.35 .	14.12 .
HEPATITIS	14.75 .	3.66 .	7.06 .	3.27 .	5.20 .	6.69 .
HYPO	44.29 +	-4.37 .	43.09 +	-4.37 .	-4.37 .	-4.37 .
IMAGE	4.72 +	1.97 .	1.52 .	2.12 .	-0.47 .	-0.00 .
IONOSPHERE	-9.56 .	6.78 .	8.94 +	-6.59 .	-6.59 .	-5.36 .
IRIS	0.00 .	8.33 .	8.33 .	7.04 .	4.35 .	10.81 .
SOLAR-FLARE-C	1.07 .	-1.88 .	-1.88 .	-0.50 .	-1.88 .	1.60 .
SOLAR-FLARE-M	0.98 .	-3.68 .	-2.47 .	1.81 .	-3.68 .	5.37 .
SOLAR-FLARE-X	12.95 .	0.00 .	0.00 .	3.20 .	-0.83 .	8.33 .
SONAR	21.11 .	-1.07 .	5.96 .	-1.07 .	-1.07 .	1.73 .
SOYA	5.02 .	11.76 .	10.68 .	9.94 .	5.64 .	11.59 +
TIC-TAC-TOE	97.18 +	72.83 .	72.83 .	55.35 .	55.35 .	81.20 +
VOTE	50.33 +	3.25 .	3.25 .	3.25 .	3.25 .	3.87 .
WAVEFORM	19.72 +	23.17 +	16.17 +	12.86 +	12.86 +	6.44 +
WINE	-19.35 .	28.85 .	28.85 .	27.45 .	24.49 .	37.29 .
POVPREČJE	27.61	12.93	14.48	11.17	9.90	14.07
ZMAGE/PORAZI	9:0	3:0	5:0	3:0	3:0	5:0

Tabela 12: Relativno izboljšanje klasifikacijske natančnosti (v %), če namesto posameznega ansambla uporabimo metodo skladanja klasifikatorjev z večličnimi modelskimi drevesi s **sedmimi osnovnimi klasifikatorji** ter signifikantnost morebitnih razlik. »+« pomeni signifikantno izboljšanje, »-« signifikantno poslabšanje, ».« pa pomeni, da razlika ni signifikantna.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	ScMLR	SMLRE
AUSTRALIAN	-2.38 .	3.52 .	4.91 .	-2.49 .	-2.07 .	16.27 .
BALANCE	49.84 +	40.00 +	39.77 +	51.60 +	48.29 +	11.91 .
BREAST-CANCER	-8.68 .	0.25 .	0.99 .	-0.25 .	-0.12 .	1.96 .
BREAST-W	25.49 +	-1.06 .	-1.06 .	0.00 .	0.00 .	-4.97 .
BRIDGES-TD	-2.55 .	3.59 .	8.52 .	-1.26 .	-1.90 .	-0.63 .
CAR	81.51 +	78.13 +	66.56 +	70.63 +	75.12 +	64.64 +
CHESS	60.90 +	-4.19 .	-4.19 .	-4.19 .	-4.19 .	-4.19 .
CONTRACEPTIVE	3.56 .	0.83 .	4.33 .	-1.40 .	-1.70 .	0.25 .
DIABETES	0.32 .	-3.94 .	1.02 .	-1.37 .	-1.77 .	-0.05 .
DIS	19.16 +	-11.26 .	-10.66 .	-7.71 .	-8.29 .	-4.65 .
ECHO	1.73 .	-6.13 .	0.25 .	-2.84 .	-8.74 .	-0.51 .
GERMAN	3.49 .	5.80 .	4.32 .	-0.17 .	-0.48 .	1.86 .
GLASS	2.36 .	-0.00 .	1.46 .	-2.28 .	-10.91 .	17.58 .
HEART-C	11.05 .	-1.66 .	-1.24 .	11.21 .	-1.03 .	17.48 .
HEART-H	3.28 .	-11.58 .	-12.65 .	3.28 .	-5.59 .	5.98 .
HEART	7.08 .	0.69 .	3.56 .	0.00 -	-0.46 -	9.41 .
HEPATITIS	-0.39 .	-0.39 .	1.92 .	2.30 .	4.85 .	8.27 .
HYP0	49.29 +	-4.60 .	41.45 +	-2.88 .	-2.88 .	-3.73 .
IMAGE	-7.43 .	31.76 +	16.02 .	-2.14 .	-6.47 .	-5.77 .
IONOSPHERE	11.71 .	11.71 .	14.29 +	-5.60 .	-6.02 .	1.86 .
IRIS	4.00 .	0.00 .	0.00 .	2.70 .	-5.88 .	22.58 .
SOLAR-FLARE-C	-0.45 .	-2.65 -	-2.60 -	1.06 .	-2.79 -	3.85 .
SOLAR-FLARE-M	2.86 .	-4.09 .	-2.89 .	1.25 .	-4.39 .	7.28 .
SOLAR-FLARE-X	10.07 .	-2.46 .	-2.46 .	-0.81 .	-4.17 .	8.09 .
SONAR	0.00 .	-3.92 .	-2.91 .	-5.30 .	-7.07 .	12.88 .
SOYA	8.08 .	0.94 .	2.77 .	16.30 +	7.47 .	17.13 .
TIC-TAC-TOE	93.00 +	73.91 .	73.91 .	60.65 .	48.93 .	80.80 .
VOTE	41.91 +	7.60 +	7.60 +	3.07 .	2.47 .	8.67 .
WAVEFORM	18.50 +	3.42 .	2.05 .	13.31 +	15.66 +	6.16 .
WINE	-34.62 .	14.63 +	10.26 .	5.41 .	2.78 .	12.50 .
POVPREČJE	23.73	11.76	12.41	9.87	7.16	13.97
ZMAGE/PORAZI	9:0	5:1	5:1	4:1	3:2	1:0

Tabela 13: Relativno izboljšanje klasifikacijske natančnosti (v %), če v ansamblih uporabimo sedem osnovnih klasifikatorjev namesto treh.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE	SMM5
AUSTRALIAN	1.33 .	-0.69 .	-5.27 .	2.13 .	1.73 .	-8.86 .	0.90 .
BALANCE	24.34 +	0.00 .	0.38 .	-3.46 .	-4.06 -	12.59 .	-10.80 .
BREAST-CANCER	-1.80 .	0.25 .	-0.87 .	-0.76 .	-1.91 .	-3.81 .	-1.39 .
BREAST-W	-5.37 -	0.00 .	0.00 .	0.52 .	0.52 .	-0.56 .	2.56 .
BRIDGES-TD	-0.64 .	-3.73 .	-7.32 .	-10.42 .	-9.72 .	-8.84 .	-10.27 .
CAR	-3.65 .	2.48 .	25.92 +	24.54 +	22.02 +	14.49 +	18.87 +
CHESS	-11.38 .	0.00 .	0.00 .	0.00 .	0.00 .	0.00 .	-2.58 .
CONTRACEPTIVE	1.09 .	2.90 +	1.16 .	2.11 .	2.42 +	1.17 .	1.08 .
DIABETES	0.48 .	7.89 .	2.41 .	0.05 .	0.38 .	0.86 .	-1.04 .
DIS	-0.00 .	-1.96 .	-1.95 .	-3.87 .	-3.31 .	-7.50 .	-3.58 .
ECHO	1.94 .	-4.17 .	-3.91 .	-4.31 .	3.43 .	-9.09 .	-7.86 .
GERMAN	3.37 .	2.99 .	3.34 .	5.81 .	6.17 .	4.63 .	7.11 .
GLASS	11.68 .	21.08 .	17.62 .	20.15 .	22.98 .	10.90 .	19.07 .
HEART-C	0.36 .	-2.11 .	-2.32 .	-3.75 .	1.62 .	1.65 .	0.81 .
HEART-H	0.20 .	4.30 +	6.68 .	1.21 .	-3.23 .	8.39 .	-6.31 .
HEART	5.28 .	-1.63 .	-4.66 .	-1.88 .	-3.11 .	4.97 .	-0.23 .
HEPATITIS	8.63 .	-3.25 .	-1.96 .	-6.53 -	-7.20 .	-9.45 .	-7.60 .
HYPO	-14.92 .	-4.37 .	-1.67 .	-6.11 .	-6.11 .	-5.24 .	-4.60 .
IMAGE	34.51 +	-6.07 .	13.41 .	29.24 +	30.33 +	30.19 +	26.16 .
IONOSPHERE	-19.12 .	-1.36 .	-1.99 .	3.10 .	3.49 .	-3.07 .	4.00 .
IRIS	-13.64 .	-0.00 .	-0.00 .	-4.23 .	1.45 .	-25.68 .	-9.09 .
SOLAR-FLARE-C	0.40 .	-0.37 .	-0.41 .	-2.72 -	-0.23 .	-3.50 .	-1.13 .
SOLAR-FLARE-M	-3.09 .	-0.74 .	-0.73 .	-0.56 .	-0.44 .	-3.22 .	-1.13 .
SOLAR-FLARE-X	0.00 .	-0.83 .	-0.83 .	0.80 .	0.00 .	-3.03 .	-3.31 .
SONAR	11.67 .	-8.90 .	-2.32 .	-7.47 .	-5.69 .	-26.30 .	-11.97 .
SOYA	0.00 .	13.79 .	0.75 .	-3.18 .	1.30 .	-3.42 .	3.22 .
TIC-TAC-TOE	61.24 +	0.00 .	0.00 .	-8.93 .	16.07 .	6.01 .	4.00 .
VOTE	9.33 .	-11.04 .	-11.04 .	-5.84 .	-5.19 .	-11.61 .	-6.04 .
WAVEFORM	9.63 +	27.01 +	21.48 +	7.78 +	5.21 +	8.53 +	8.25 .
WINE	16.13 .	21.15 .	25.00 .	27.45 .	26.53 .	32.20 .	5.41 .
POVPREČJE	5.88	2.16	2.82	2.29	3.76	0.94	0.83
ZMAGE/PORAZI	4:1	3:0	2:0	3:2	4:1	3:0	1:0

C. Rezultati poskusov na praktičnih problemih

Tabela 14: Klasifikacijske napake (v %) osnovnih klasifikatorjev na praktičnih problemih.

MNOŽICA POD.	J4.8	IBK	NB	K*	KDE	DT	MLR
ACUTE_ABDOMINAL_PAIN	15.58	12.27	15.16	12.69	12.18	15.73	12.96
WATER	32.19	39.62	28.32	37.84	45.34	38.01	34.79
WATER+DIVERSITY	27.84	39.38	27.98	36.10	41.88	25.79	35.27
SEA_CUCUMBER	18.44	19.77	15.78	16.41	20.31	16.48	13.05
REUMA+BKGRD_KNOW	50.54	59.44	49.22	57.75	58.14	52.29	51.41
REUMA	50.69	57.79	52.77	59.61	57.55	51.84	50.82
BIODEG-FUNCTION_GROUP	45.70	44.51	48.72	44.09	43.08	50.64	51.13
BIODEG-FRAGMENTS	42.96	43.51	49.97	43.87	42.47	43.90	47.07
DIATOMA_VULGARE	35.07	38.02	34.99	34.98	33.87	33.44	33.75
TRAFFIC	12.85	11.48	19.10	10.51	10.82	13.40	18.63
POVPREČJE	33.19	36.58	34.20	35.39	36.56	34.15	34.89

Tabela 15: Klasifikacijske napake (v %) ansamblov klasifikatorjev na praktičnih problemih. Metode uporabljajo **tri različne osnovne klasifikatorje**.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE	SMM5
ACUTE_ABDOMINAL_PAIN	12.18	12.45	12.51	11.88	11.55	12.09	12.12
WATER	28.87	29.32	28.63	27.33	27.57	26.23	27.02
WATER+DIVERSITY	26.51	29.28	26.95	25.99	26.44	28.87	25.75
SEA_CUCUMBER	17.11	18.05	18.36	16.02	16.41	16.25	15.31
REUMA+BKGRD_KNOW	48.96	49.22	49.74	48.98	47.84	55.26	49.61
REUMA	49.33	52.64	51.71	51.23	49.94	54.61	50.95
BIODEG-FUNCTION_GROUP	42.56	45.00	45.61	42.65	41.62	43.63	43.17
BIODEG-FRAGMENTS	42.59	43.78	43.14	41.04	40.12	42.20	42.01
DIATOMA_VULGARE	33.61	35.22	34.77	34.09	34.09	33.70	34.28
TRAFFIC	10.86	11.99	12.03	11.09	10.98	11.41	11.41
POVPREČJE	31.26	32.70	32.35	31.03	30.66	32.43	31.16

Tabela 16: Klasifikacijske napake (v %) ansamblorov klasifikatorjev na praktičnih problemih. Ansamblorji uporabljajo **sedem različnih osnovnih klasifikatorjev**.

MNOŽICA POD.	VOTE	SELB	SMDT	SMLR	SCMLR	SMLRE	SMM5
ACUTE_ABDOMINAL_PAIN	10.12	13.49	13.19	10.75	10.69	11.79	11.31
WATER	29.69	29.32	29.62	26.44	26.64	27.40	26.03
WATER+DIVERSITY	27.05	28.90	28.46	24.93	24.01	28.01	24.83
SEA_CUCUMBER	14.84	13.83	14.06	15.70	15.47	15.63	15.16
REUMA+BKGRD_KNOW	49.24	49.55	50.69	50.22	47.73	58.46	50.50
REUMA	49.70	52.16	51.65	51.02	49.72	58.94	51.08
BIODEG-FUNCTION_GROUP	40.73	45.12	46.16	42.90	41.95	45.64	44.09
BIODEG-FRAGMENTS	39.94	43.78	43.38	42.44	40.61	44.54	42.96
DIATOMA_VULGARE	32.36	34.92	33.90	32.00	31.99	31.83	32.68
TRAFFIC	9.77	11.52	11.25	11.05	10.94	12.23	11.45
POVPREČJE	30.34	32.26	32.24	30.75	29.98	33.45	31.01