

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
INFORMATIKOS KATEDRA

Tiesinis neuroninių tinklų kombinavimas: Duomenų balinimo operacijos poveikis tiesiniam neuroninio tinklo kombinavimui su įvairaus skirtingumo ekspertais

Linear Neural Network Combination: The Effect of Data Whitening in Combination of Neural Networks Based on Different Classifiers

Magistro baigiamasis darbas

Atliko:

Šarūnas Gruodis

(parašas)

Darbo vadovas:

lekt. Arūnas Janeliūnas

(parašas)

Recenzentas:

a. Rimantas Kybartas

(parašas)

Vilnius, 2012

Santrauka

Darbe apžvelgtos bazinės neuroninių tinklų savybės, architektūra ir kombinavimo metodikos. Atlikta neuroninių tinklų tiesinių kombinavimo metodų analizė. Palyginti skirtingi neuroninių tinklų diversiškumo matavimai ir pritaikyti eksperimente. Išnagrinėta duomenų balinimo transformacija ir jos taikymo galimybės neuroninių tinklų kontekste.

Šio darbo tikslas – ištirti duomenų balinimo operacijos poveikį tiesinių kombinavimo metodų ir skirtingų ekspertų parinkimui, formuojant kokybiškesnį neuroninių tinklų kolektyvą.

Pritaikytas naujas metodas blogos kokybės skirtingų ir geros kokybės panašių ekspertų klasifikatoriams sudaryti.

Svarbiausios eksperimento *išvados*:

- Balinimo transformacija taikoma pradiniam duomenims klasifikatorius padaro panašesnius.
- Duomenų balinimo transformacijos taikymas ekspertų aibės atsakymams blogos kokybės skirtingus klasifikatorius padaro panašesnius už geros kokybės vienodus klasifikatorius net iki 40%.

Raktiniai žodžiai: Neuroninių tinklų kombinavimas, diversiškumas, skirtingumo matavimai, tiesinis kombinavimas, balinimo transformacija, klasifikatorius, kolektyvas.

Summary

The purpose of this study was to research the effect of data whitening for different linear combination and diversity measure methods selection in context of building better classifier ensemble.

In work main characteristics, architecture and combination methods of neural network were identified. Different linear combination methods and diversity measures were analyzed in context of data whitening.

In experiment classifier quality effect of data whitening transformation was researched with different classifiers.

The new method for designing bad quality different and good quality similar classifiers was constructed.

Most important experiment *conclusions*:

- Data whitening transformation applied for initial data vectors makes classifiers more similar.
- Applying data whitening transformation for single classifier set answers makes bad quality different classifiers better than good quality similar classifiers in lower generalization error range up to 40%.

Keywords: Neural Network Ensembles, Combination of Classifiers, Classification, Diversity Measures, Data Whitening, Linear Ensembles.

Turinys

Įvadas	1
Darbo aktualumas ir tikslas	2
Darbo uždaviniai	3
Siekiami tyrimo rezultatai	3
Naudoti tyrimo ir analizės metodai	3
1. Neuroniniai tinklai ir tiesinis jų kombinavimas.....	4
1.1 Neurono modelis.....	4
1.2 Neuroninių tinklų architektūros	5
1.3 Kombinavimo etapai.....	6
1.4 Klasifikatorių tipai	7
1.5 Kombinavimo metodikos.....	9
2. Skirtingumo (diversiškumo) matavimai	12
2.1 Diversiškumas.....	13
2.2 Skirtingumo matavimai.....	13
2.3 Diversiškų tinklų kūrimas	15
3. Kolektyvo topologijos.....	17
4. Duomenų balinimas	20
5. Eksperimento duomenų atranka, eiga ir rezultatai.....	22
5.1. Eksperimento tikslas	22
5.2. Naudojami įrankiai ir bibliotekos	23
5.3. Eksperimento planas ir eiga	23
5.4. Tyrime naudojamos duomenų imtys.....	24
5.5. Praktiniai tyrimai	25
5.5.1 Balinimo transformacijos taikymas kombinuojant neuroninius tinklus	25
5.5.2 Balinimo transformacijos poveikis klasifikatorių diversiškumui	33
5.5.3 Balinimo transformacijos taikymo ekspertų atsakymams poveikis klasifikatoriaus kokybei	40
Darbo rezultatai	44
Išvados.....	44
Šaltinių sąrašas	45
Priedai	48
A priedas. Praktinio eksperimento bandymų programos ir gauti rezultatai.....	48

Ivadas

Tiesinis neuroninių tinklų kombinavimas yra ekspertų kolektyvo sudarymo procesas, kuriame baigtinis skaičius neuroninių tinklų yra apmokomi tai pačiai klasifikavimo ar prognozavimo užduočiai spręsti. Bendrai tai apima du etapus: skirtingų neuroninių tinklų apmokymą ir tiesinį šių ekspertų atsakymų kombinavimą. Į procesą įtraukiama duomenų balinimo operacija, kuri transformuoja duomenis ir jais apmokant neuroninius tinklus, leidžia sukurti kokybiškesnius klasifikatorius, pasinaudojant ekspertų diversiškumo matavimais.

Neuroninis tinklas yra lygiagreti sistema, galinti spręsti įvairius uždavinius, kurių nesugebėtų išspręsti paprasti tiesiniai, linijiniai skaičiavimai. Dirbtinis neuroninis tinklas yra paremtas biologinio neuroninio tinklo pagrindu. Kitais žodžiais tariant, tai biologinės neuroninės sistemos emuliacija.

Paprasčiausias tiesinio neuroninio tinklo tipas yra perceptronas. Jį sugalvojo amerikiečių psichologas ir informatikas Frankas Rosenblatas (Frank Rosenblatt), dar 1958 metais [Ros58]. Paprasčiausias perceptronas susideda iš n įvesties parametru, perdavimo funkcijos, bei vieno išvesties parametro. Perceptronas atlieka klasifikatoriaus vaidmenį (gali būti taikomas ir prognozavimo uždaviniuose), kuris sugeba spręsti tiesiškai padalinamas užduotis. Šiandien egzistuoja nemažai įvairių metodikų, kurios leidžia kombinuoti skirtingus klasifikatorius. Jų kombinavimas gali būti skirstomas į kelias kategorijas, kurios paremtos skirtingų klasifikatorių rezultatais. Paprastai tokių klasifikatorių kombinavimas leidžia gauti geresnius rezultatus [KS92]. Taip pat pastebėta, jog netgi blogi klasifikatoriai gali turėti vertingą informaciją, gerinant galutinį rezultatą. Galutinis neuroninio tinklo rezultatas priklauso ir nuo to, kokie klasifikatorių skirtingumo matavimai bus parinkti, kaip bus apdorojami skirtingų klasifikatorių rezultatai. Jie gali būti klasifikuojami pagal daugumos taisyklę, vidurkinami (regresija) ar priskiriami skirtingiems svoriams ir pan. Krogh ir Vedelsby pagrindė, kad geriausią klasifikatorių kombinaciją, duodančią mažiausią klasifikavimo klaidą, galima sudaryti tada, kai tie klasifikatoriai yra gana aukštos kokybės, tačiau jų rezultatai maksimaliai prieštarauja vienas kitam [Ved95].

Kombinuojant vienodus klasifikatorius galutinis rezultatas nesikeičia, o tik išauga skaičiavimo sąnaudos. L. Hansen ir P. Salamon įrodė, jei kiekvieno klasifikatoriaus paklaida yra mažesnė nei 50%, tai su sąlyga, kad kiekvienas klasifikatorius neįtakoja kitų klasifikatorių rezultatų, augant tokių klasifikatorių kombinacijoms į begalybę, galutinio rezultato paklaida artėja prie nulio [HS90].

Bendrai, klasifikatorių skirtingumo matavimas, nėra lengva užduotis. Pagrindinė to priežastis yra ta, kad nėra bendro visų klasifikatorių formalaus matavimo. L. Kuncheva ir J. Whitaker

išstudijavo 10 skirtingų klasifikatorių rezultatų skirtingumo matavimo metodų [KW02]. Rezultate buvo nustatytos sąajos tarp klasifikatorių panašumo ir kokybės, tačiau praktikoje, kaip teigia autoriai, ne visada jos pasitvirtina. Todėl dar ir dabar nėra visiškai apibrėžto skirtingumo ar priklausomybės matavimo tarp klasifikatorių.

Darbo aktualumas ir tikslas

Bėgant metams, kartu tobulėja ir technologijos. Dirbtinių neuroninių tinklų (DNT) atsiradimas, programų sistemoms davė naujus, geresnius vystymosi sprendimus. Pažangių sistemų projektavimas ir įgyvendinimas šiandien yra lemiamas veiksnys kuriant inovatyvius ir geresnius produktus visuomenei. Taigi neuroninių tinklų taikymas šiuolaikinėse sistemose, neretai tampa būtinybe. Įvairus jų kombinavimas leidžia pasiekti geresnius rezultatus.

Norint gauti geresnius neuroninio tinklo rezultatus, galima atlikti įvairius skaičiavimus su pačiais mokymo duomenimis, įvairiuose neuroninio tinklo veikimo etapuose. Duomenis galima apdoroti juos prieš paduodant neuroniniam tinklui, taip pat juos gavus, kaip rezultatus, iš aibės klasifikatorių, ar kombinuojant duomenų apdorojimą. Vienas iš duomenų apdorojimo būdų yra vadinamas – duomenų balinimu (angl. *Data Whitening*) [EC02]. Duomenų balinimo operacija transformuoja duomenų erdvę taip, kad duomenų išdėstymas erdvėje taptų rutulio formos, taip sumažinant vidutinę kvadratinę paklaidą tarp originalių ir išbalintų duomenų. T.y. transformuotų duomenų rezultatas tampa kiek įmanoma artimesnis įvestiems duomenims, kvadratinės paklaidos atžvilgiu. C. Eldar ir V. Oppenheim nagrinėja tiesines duomenų balinimo transformacijas tiek baigtiniams vektoriams, tiek begalinio ilgio signalams. Yra įrodyta, kad duomenų balinimo operacijų taikymas, pagerina įvairių realaus pasaulio elementų atpažinimo kokybę [Eld01]. C. Eldar ir M. Chan nagrinėja bevielų komunikavimo sistemų duomenų balinimo įtaką, signalais perduodamiems duomenims [EC02].

Š. Raudys ir M. Iwamura savo tyrimuose taiko statistinius ir neuroninius tinklus, kaip pagrindinius klasifikatorius. Juos kombinuoja kartu su duomenų balinimo operacijomis bei vienasluoksniais perceptronais, taip siekiant geresnių rezultatų, atpažįstant Japonų rašytinius simbolius. Duomenų balinimo procese taip pat naudojamas duomenų pasukimas, siekiant pagreitinti perceptrono mokymosi procesą. Autoriai savo tyrime pabrėžtinai nagrinėja išskirtinių, panašių simbolių atpažinimo subtilybes ir kylančias problemas, kombinuojant duomenų balinimo operaciją ir tiesinius neuroninius tinklus [ŠI04].

Šiandien skirtingų tyrimų rezultatai neatsako, kaip duomenų balinimo operacija įtakoja klasifikatorių skirtingumą, bei kokią įtaką ji turi klasifikatorių kolektyvo parinkimui.

Darbo tikslas – išnagrinėti neuroninių tinklų tiesinius kombinavimo metodus, ekspertų skirtingumo matavimus ir duomenų balinimo operaciją. Ištirti duomenų balinimo operacijos poveikį tiesinių kombinavimo metodų ir skirtingų ekspertų parinkimui, formuojant neuroninių tinklų kolektyvą, kai duomenų balinimo operacija taikoma tiek pradiniais duomenimis, tiek ekspertų aibės atsakymams.

Darbo uždaviniai

Darbe nagrinėjamos darbo tikslui pasiekti skirtos užduotys:

1. Neuroninių tinklų taikymo ir tiesinių kombinavimo metodų literatūros analizė.
2. Diversiškumo matavimų ir skirtingų neuroninių tinklų konstravimo literatūros analizė.
3. Duomenų balinimo transformacijos literatūros analizė.
4. Klasifikatorių, apmokytų su išbalintais ir originaliais duomenimis panašumo tyrimas, tokių klasifikatorių kokybės vertinimas;
5. Duomenų balinimo taikymo įtaka balinant tiek pradinis duomenimis, tiek klasifikatorių aibės atsakymus.
6. Kombinuotų neuroninių tinklų su taikyta balinimo transformacija ir be jos rezultatų palyginimas.

Siekiami tyrimo rezultatai

Šis tyrimas turėtų pateikti rezultatus, kurie atsakytų į sekančius klausimus:

1. Ar išbalintų duomenų klasifikatoriai yra vienodesni ir pagal kokius matavimo būdus tai galima spręsti?
2. Kaip keičiasi klasifikatorių kolektyvo parinkimo, atsižvelgiant į jų skirtingumą, strategija kai į klasifikatorių mokymą įvedama duomenų balinimo operacija?
3. Kokią įtaką rezultatams daro duomenų balinimo operacija, ją taikant ir balinant klasifikatorių aibių atsakymus?

Naudoti tyrimo ir analizės metodai

Darbe naudojami tiesiniai kombinavimo metodai ir neuroninių tinklų teorija. Ekspertų panašumui įvertinti naudojami diversiškumo matavimo metodai.

Eksperimentiniai tyrimai atlikti su MATLAB programine įranga ir papildomomis neuroninių tinklų bibliotekomis.

1. Neuroniniai tinklai ir tiesinis jų kombinavimas

Šiame skyriuje apžvelgsiu neurono modelį ir pagrindines jo savybes. Sekančiame poskyryje kategorizuosiu neuroninių tinklų architektūras ir apžvelgsiu pagrindinius tokių tinklų parametrus. 1.3 Poskyryje nagrinėsiu du kombinavimo etapus. Jie yra skirstomi į požymių ir sprendimų etapus. 1.4 Poskyryje apžvelgsiu keturis klasifikatorių tipus ir jiems būdingas savybes. 1.5 poskyryje išanalizuosiu praktikoje dažniausiai pasitaikančias klasifikatorių kombinavimo metodikas.

1.1 Neurono modelis

Neuroninių tinklų vystymasis prasidėjo nuo paprasčiausio neuroninio tinklo vadinamo Perceptronu. Iki to laiko W. McCulloch ir W.Pitts pristatė užuomazgas į neuroninius tinklus, apibrėžę dirbtinio neurono modelį. Šio neurono esmė buvo slenksčio (angl. *Threshold*) įverčiu paremtas klasifikavimas. Kur įvesties požymių svorių suma yra palyginama su duotuoju slenksčiu ir pagal tai apsprendžiama išvestis. Kai suma yra mažesnė už slenkstį, išvestis yra 0. O kai suma didesnė ar lygi slenkščiui, išvestis lygi 1 [MP43].

Minėtasis klasifikatorius – tai paprasčiausias neuronas k (1), kurio aktyvacijos funkcija φ yra slenkstinė (angl. *Hard Limit*). Bendrai neuronas susideda iš trijų pagrindinių dalių:

- *Sinapsių* aibės, kitaip vadinamų jungiamųjų *ryšių*, kur kiekvienas ryšys j yra apibūdinamas tam tikru svoriu w_{kj} . Čia svoriai gali būti tiek teigiami, tiek ir neigiami.
- *Sumatoriaus*, kuris sumuoja įėjimų j signalus padaugintus iš atitinkamų svorių ir sudaro tiesinę kombinaciją.
- *Aktyvacijos funkcijos* φ , kuri riboja išėjimo amplitudę. Būtent todėl kartais ji yra vadinama slopinančiąja (angl. *squashing*) funkcija. Dažniausiai neuroniniuose tinkluose yra naudojamas vienetinio ilgio išėjimo intervalas $[0,1]$ arba $[-1,1]$.

$$u_k = \varphi \left(\sum_{j=1}^m w_{kj} x_j + b_k \right) \quad (1)$$

Be slenkstinės aktyvacijos funkcijos egzistuoja ir kitos, praktikoje dažnai pasitaikančios, kitaip dar vadinamos perdavimo funkcijos. Tai gali būti tiesinė, sigmoidinė ar tangento aktyvavimo funkcijos. Toks neuronas vadinamas vienasluoksniu perceptronu.

Pagal tai, kaip neuronai yra sujungti tarpusavyje į neuroninį tinklą, galime išskirti pagrindines neuroninių tinklų struktūras. Tokie tinklai pasižymi savitais bruožais, kuriuos apžvelgsime sekančiame poskyryje.

1.2 Neuroninių tinklų architektūros

Neuroninių tinklų architektūros skirstomos pagal tai, kaip neuroninis tinklas yra apmokomas. Visumoje yra išskiriamos trys pagrindinės neuroninių tinklų struktūros.

Vienasluoksniai tiesioginio sklidimo (angl. Feedforward) tinklai. Sluoksninėje neuroninių tinklų struktūroje, neuronai yra išdėstyti sluoksniais. Paprasčiausiame, tiesioginio sklidimo neuroniniame tinkle yra įvesties sluoksniu kuris sujungtas su išvesties sluoksniu – kuriame vyksta skaičiavimai. Paprastai toks tinklas vadinamas vienasluoksniu, kadangi yra tik vienas skaičiavimų sluoksniu. Čia įvesties sluoksniu neskaiciuojame, nes jame skaičiavimai nevyksta. Tokie tinklai vadinami tiesioginio sklidimo dėl to, kad įvesties sluoksniu negauna jokios informacijos, po to kai ją perduoda sekančiam sluoksniu.

Daugiasluoksniai tiesioginio sklidimo tinklai. Šis tinklas skiriasi nuo pirmojo sluoksniu skaičiumi. Čia tarp įvesties ir išvesties sluoksniu atsiranda dar viena sluoksniu aibė – vadinama paslėptais sluoksniu. Paslėptų sluoksniu funkcija atlieka tarpininko vaidmenį tarp įvesties ir išvesties neuronų, atlikdama tam tikrus skaičiavimus. Į neuroninį tinklą įtraukiant paslėptus sluoksnius, neuroninis tinklas geba atlikti aukštesnės eilės matavimus. Tai yra naudinga, kai neuroninis tinklas turi didelį įvesties sluoksniu.

Tokiame tinke įvesties sluoksniu neuronai atitinka požymių vektorių, kur požymiai per jungtis atiduodami sekančiam – paslėptam sluoksniu, kuriame vyksta skaičiavimai. Atitinkamai to sluoksniu išvestis atiduodama sekančiam sluoksniu, kol pasiekiamas išvesties sluoksniu.

Tokio tipo tinklai gali būti *pilnai sujungti*, tuomet kai kiekvienas neuronas kiekviename sluoksnyje yra sujungtas su kiekvienu sekančio sluoksniu neuronu. *Dalinai sujungti* – kai trūksta vieno sluoksniu neuronų jungčių su sekančio sluoksniu neuronais, pilnai sujungto tinklo atžvilgiu.

Grįžtamojo ryšio (angl. recurrent) tinklai. Šių neuroninių tinklų architektūra išsiskiria nuo prieš tai minėtų tuo, kad šiuose tinkluose egzistuoja bent vienas grįžtamasis ryšys, dar vadinamas atvirkštiniu sklidimu (angl. *back-propagation*). Pavyzdžiui grįžtamojo ryšio neuroninis tinklas gali

turėti vieną sluoksnį neuronų, kur kiekvienas neuronas yra sujungtas su įvesties sluoksniu ir perduoda išvesties rezultatus [Hay98].

Praktikoje vis dažniau imama kooperuoti tokius neuroninius tinklus, siekiant gauti geresnius rezultatus. Neuroninių tinklų kombinavimas gali būti padalintas į kelis etapus, kuriuos apžvelgsiu sekančiame poskyryje.

1.3 Kombinavimo etapai

Tyrinėjant neuroninius tinklus buvo pastebėta, kad labai dažnai kombinuojant vienas kitą papildančius klasifikatorius ar pradinius duomenis, sumoje gauname kokybiškesnius ir greitesnius klasifikatorius arba kitaip vadinamus ekspertus [WMT99]. Atpažinimo uždaviniuose kombinavimas atveria naują tyrimų kryptį, kurios tikslas yra sukurti patikimas atpažinimo sistemas.

Ekspertų sprendimų kombinavimas gali būti atliekamas skirtinguose etapuose. Vienas iš jų vadinamas požymių (angl. *features*) etapas, kitas – klasifikatorių sprendimų (angl. *decisions*) etapas.

Požymių etapas. Kiekvienas klasifikatorius, spręsdamas specifinės užduoties klasifikavimo problemas, gali gauti skirtingus rezultatus. T.y. klasifikatoriaus kokybės procentas klasifikuojant duomenis gali skirtis nuo kitų ekspertų kokybės. Net ir tokiu atveju, visi klasifikatoriai gali būti prasti, paėmus juos atskirai. O išrinkus vieną geriausią ekspertą, jis nebūtinai gerai elgsis realiose sąlygose (t.y. ne su mokymo ar testavimo duomenimis, o realiose situacijose – praktiniame taikyme). Šiame etape yra kombinuojami vektorių požymiai, kurių rezultatas vėliau klasifikuojamas kitais metodais. Sekančiame, aukštesnio lygio etape, vėlgi gali būti atliekamas kiekvieno eksperto sprendimų kombinavimas. Požymių etapo rezultatas yra klasifikatorių sprendimai, kitaip tariant – išvestys.

Kombinuojant požymius pirmame etape apsaugome nuo informacijos praradimo. Šio etapo kombinavimas labiau tinka nedaug dimensijų turintiems vektoriams, tačiau esant dideliems požymių vektoriams, įvairios klasifikavimo technikos čia tampa ne visada praktiškos. Taip pat klasifikavimą apsunkina vienareikšmio heterogeninių požymių tipo atvaizdavimo trūkumas. Be to, nesudėtingas kelių požymių aibių išskyrimas iš vektoriaus, gali pareikalauti sudėtingos klasifikavimo schemos, jei kitų požymių išskyrimas nėra paprastas. Todėl šiame etape tinkamų klasifikatorių ir jų mokymo algoritmų pasirinkimas tampa gana sudėtingas.

Sprendimų etapas. Šis etapas yra jau aukštesniame lygyje ir čia klasifikuojami ne vektorių požymiai, o ekspertų sprendimai. Jų kombinavimas leidžia išvelgti vertingos informacijos praradimo riziką. Čia mes jau neturime pradinių duomenų, o tik klasifikatorių atsakymus. Kadangi

pirmo etapo klasifikavimo procesas negali būti atgamintas, kombinavimo savybės natūraliai yra prastesnės nei kad požymių etape. Sprendimų etape taip pat nelieta originalių požymių heterogeniškumo. Čia mes turime tik ekspertų iškirtus požymius, kurie šiame etape yra negrįžtamai heterogeniniai lyginant su pirmuoju etapu. Vieni klasifikatoriai gali būti koreliuoti, kiti tiesiškai atskirti. Kiekvieno klasifikatoriaus nuopelnai turi būti pasverti ir įvertinti.

Be šių trūkumų galime įžvelgti ir privalumus. Informacijos filtravimas sumažina neuroninio tinklo kompleksumą ir padaro kombinavimą lengviau valdomu procesu. Siekiant efektyvių kombinacijų, gali būti pasirinktas tinkamas klasifikatorius, atsižvelgiant į pirmo lygio klasifikavimo sprendimus. Be to, klasifikatorių atsakymai turi fiksuotą dimensijų skaičių ir savas interpretacijas [Lee95].

Kombinavimo metodika yra ypač naudinga sudėtingiems uždaviniams spręsti. Sudėtinga juos spręsti pasidaro tada, kai duomenyse yra daug triukšmo (pvz. balso atpažinimo uždavinys, esant triukšmingam fonui). Taip pat uždaviniai su daug dimensijų turinčiais vektoriais (pvz. radaro atpažinimo uždavinys), ar situacijos kai mokymo duomenų nėra daug (pvz. naujų ligų identifikavimas, kuriomis yra sirgę tik nedaugelis žmonių).

1.4 Klasifikatorių tipai

Kooperuojant neuroninius tinklus yra svarbu suvokti klasifikatorių tipus. Skirtingų tipų klasifikatoriams yra taikomos skirtingos kombinavimo metodikos ir algoritmai. L. Xu klasifikatorius skirsto į tris tipus pagal jų išvesties rezultatus [KSX92].

I tipo (abstraktus lygio) klasifikatoriai laikomi tokie, kurie pateikia tik klasę kaip rezultatą. Jei turime objektų (šablonų) aibę P , kuri susideda iš M skirtingų elementų $P = C_1 \cup \dots \cup C_M$, kur kiekvienas $C_i: \forall_i \in \Lambda = \{1, 2, \dots, M\}$ atstoja vieną unikalią klasę (pvz. $M = 10$, kai klasifikuojame skaitmenis, skaitmenų atpažinimo uždavinyje, kur skaitmenys gali būti 0, 1 ... 9). Tada jei imame bet kurį požymių vektorių x , kaip įvesties parametą iš P , klasifikatoriaus θ užduotis yra vektoriui x priskirti indeksą $j \in \Lambda \cup \{M + 1\}$, kuris identifikuoja vieną iš galimų klasių C_j , kai $j \neq M + 1$. Kai $j = M + 1$, klasifikatorius θ nežino kuriai klasei priskirti vektorių x . Kitaip tariant x nėra atpažintas arba yra pripažintas netinkamu. Nepaisant klasifikatoriaus vidinės struktūros ir metodų kuriais jis remiasi, mes šiuo atveju į klasifikatorių žiūrime kaip į funkciją $\theta(x) = j$. Čia paduodame požymių vektorių x ir gauname klasės indeksą j .

II tipo (reitingavimo lygio) klasifikatoriai pateikia išvesties rezultatą, kuris yra surūšiuotas sąrašas klasių pagal reitingą. Klasifikatoriaus išvesties informacija j mus labiausiai domina kaip klasifikavimo rezultatas (sužinome kuriai klasei priklauso duotasis požymių vektorius). Tačiau praktikoje dauguma klasifikavimo algoritmų paprastai pateikia, ar gali pateikti ir kitos susijusios informacijos. Pavyzdžiui Bajeso (angl. *Bayes*) klasifikatorius gali išvesti klasių aibės M tikimybių reikšmes $P(i/x): i = 1, \dots, M$ kiekvienai klasei ar klasių poaibiui. Finale, rezultatas būtų C_j klasė iš M klasių aibės, kur indeksas j yra rezultatų sąrašo viršuje (pagal reitingą).

III tipo (matavimų lygio) klasifikatoriai priskiria balus kiekvienai klasei, kurie vėliau gali būti naudojami klasių pasiklovimo tikslais. Čia klasifikatorius θ kiekvienai klasei iš Λ (ar klasių poaibiui $J \subset \Lambda$) priskiria tikimybinę reikšmę, kuri pasako su kokia tikimybe vektorius x yra priskirtas tai klasei. III tipo klasifikatoriai dar vadinami tikimybiniai klasifikatoriais.

Tarp šių trijų klasifikatorių tipų, matavimų lygio klasifikatoriai gali pateikti daugiausiai informacijos, o abstraktaus lygio – mažiausiai. Remiantis III ir II tipo klasifikatorių rezultatais, galima surūšiuoti visas klases iš Λ pagal rūšiavimo taisykles – didėjant arba mažėjant. Dauguma klasifikavimo algoritmų matavimo lygyje gali pateikti papildomą informaciją. Pvz. Bajeso klasifikatorius pateikia klasių tikimybes, arba įvairūs atstumų klasifikatoriai (Pvz. Euklidinis klasifikatorius) pateikia atstumus iki artimiausių klasių vidurkių. Kitais žodžiais tariant, III tipo klasifikatorių procesas yra tarpinis skaičiavimas, kurio rezultatai dažniausiai vėl apdorojami kitų neuroninio tinklo ekspertų pagalba.

Be šių klasifikatorių tipų, L. Kuncheva išskiria dar vieną tipą [Kun04]. Tai *0 tipo* klasifikatoriai, dar vadinami Orakulo tipo (angl. *Oracle level*). Jie pasižymi tuo, kad jų atsakymai yra *true* arba *false*. Tokie klasifikatoriai atmeta visą informaciją, tame tarpe ir klasės numerį. Toks klasifikatorius yra dirbtinis, nes jį galima pritaikyti tik testiniams duomenis. T.y visi požymių vektoriai turi būti papildyti ir klase. Jei toks klasifikatorius teisingai suklasifikuoja požymių vektorius, jis grąžina tik *true* reikšmę. Priešingai – nesėkmės atveju.

Klasifikatorių tipai vaidina svarbią rolę renkantis kombinavimo algoritmą. I tipo ekspertų sprendimai, kurie yra pateikiami kaip vienas atsakymas iš daugelio klasių, yra lengvai apdorojami įvairių algoritmų. II tipo atsakymai yra panašūs į I tipo atsakymus tuo, kad sprendimai yra baigtiniai ir diskretūs. Skiriasi tik galimybių skaičiumi. III tipo sprendimai dažnai būna tolydūs ir beribiai erdvėje, kuriuos yra sudėtinga kombinuoti.

1.5 Kombinavimo metodikos

Įvairiais tyrimais yra pagrįsta, jog kelių ar daugiau skirtingų algoritmų kombinavimas leidžia ženkliai pagerinti įvairių problemų atpažinimo kokybę [JNS94]. J. Hull su kitais tyrimo nariais nagrinėjo rašytinius simbolius specializuodamiesi į pašto adresų atpažinimą [CHH90]. Kombinuojant tris skirtingus ekspertus (holistinį, kontūrinį ir struktūrinį), jiems pavyko gauti geresnius rezultatus. L. Xu ir A. Krzyzak tyrė skirtingus kombinavimo metodų taikymus, panaudojant skirtingų tipų klasifikatorius skaitmenų atpažinimo srityje [KSX92]. Visumoje egzistuoja be galo daug įvairių kombinavimo metodikų. Tačiau apžvelgsime tik pagrindines, praktikoje dažniausiai pasitaikančias metodikas, kurias vėliau panaudosime praktinėje šio darbo dalyje, kombinuojant neuroninius tinklus. Verta pabrėžti, kiekvienas metodas turi savų privalumų bei apribojimų. Tiesiniame neuroninių tinklų kombinavime dažnai yra naudojamas balsavimo principas. Be jo taikomi ir kiti metodai: Bajeso taisyklė, Dempster-Shafter teorija, regresiniai modeliai ir t.t. Kai kurie iš jų reikalauja netiesinių skaičiavimų, todėl šiame darbe jų nenagrinėsime. Apžvelgsime esminius balsavimo metodus, kuriais remiantis praktinėje tyrimo dalyje kooperuosime neuroninius tinklus.

Balsavimo metodai. Balsavimo principas yra vienas iš seniausių ir paprasčiausių sprendimų kombinavimo metodų. Panašiai kaip ir rinkimuose, kiekvieno klasifikatoriaus atsakymai yra laikomi lygiasvoriais. Klasifikavime, vadovaujantis daugumos principu, klasė surinkusi daugiau nei pusę ekspertų balsų – laikoma galutiniu sprendimu. S. Srihari savo tyrime parodė, jog dviejų klasių uždavinyje, balsavimo metodo kokybė kyla monotoniškai, kai konkretaus klasifikatoriaus atsakymai yra geresni nei atsitiktiniai [Sri82].

Tačiau gali atsitikti ir toks įvykis, kai nei viena klasė nesurenka daugumos balsų. Tuomet sprendimas yra atmetamas arba naudojami papildomi metodai. Vienas iš daugumos principo variantų yra pasirinkti klasę, kuri surenka daugiausiai balsų, nepaisant ar už tą klasę balsavo daugiau nei pusė kolektyvo narių. Kai klasifikatoriai yra abstraktaus lygio, jie pateikia rezultatą klasių vektorių. Todėl tokio neuroninio tinklo atsakymas yra rezultatų vektorių suma. Tačiau šiuo atveju gali būti lygiosios tarp tam tikrų klasių.

Tokiu atveju lygiųjų problemą galima išspręsti priimant geresnių klasifikatorių sprendimus. Toks balsavimo variantas realizuojamas priskiriant atitinkamus svorius $w_i = 1 + P_i$, kiekvienam klasifikatoriui. P_i - tai klasifikatoriaus tikslumas (kokybė). Kadangi P_i reikšmė gali būti tarp 0 ir 1, tai neįvykus lygiosioms, toks patobulinimas galutinio tinklo sprendimo neįtakoja. Toks balsavimo metodas gali būti užrašytas funkcija VT:

$$VT_j(T_1(S_1), \dots, T_1(S_M)) = \sum_{i=1}^M (1 + P_i) s_{ij} \quad (2)$$

Formulėje s_{ij} yra klasifikatoriaus rezultatų vektoriaus S_i elementai. Kai klasifikatoriai yra I tipo (abstraktus lygio), elementų s_{ij} reikšmės yra 0 arba 1, kur i nurodo rezultatų vektorius, j – klasės, tame vektoriuje, balsavimo rezultatą. Norint užtikrinti, kad s_{ij} įgytų reikšmes tik 0 arba 1, taikome filtrą T_1 kiekvienam rezultatų vektoriumi. Tuomet kiekvienas rezultatų vektorius įgyja tik vieną 1. Visi kiti vektoriaus elementai lygūs 0. Didžiausią įvertinimą gavusi klasė iš kiekvieno klasifikatoriaus, įgyja reikšmę tarp 1 ir 2, kai $s_{ij} = 1$. Visos kitos klasės iš to klasifikatoriaus vertinamos 0. Taip išsprendžiama lygiųjų tarp klasių vertinimo problema. Be abejo kiekvienas klasifikatorius neturi įgyti vienodų svorių w_i , nes tokiu atveju vėl atsirastų tikimybė, kad klasės surinks vienodai balsų.

Yra dar viena dažnai naudojama balsavimo schema, atspindinti klasifikatoriaus patikimumą sprendimų etape. Pavyzdžiui, jei klasifikatoriaus θ_A tikslumas yra 90%, o klasifikatoriaus θ_B – tik 60%, tuomet θ_A atsakymams galime priskirti svorį 0,9. Atitinkamai θ_B – 0,6. Šis algoritmas paremtas pasvertu balsavimu, gali būti užrašytas funkcija WV:

$$VW_j(T_1(S_1), \dots, T_1(S_M)) = \sum_{i=1}^M P_i s_{ij} \quad (3)$$

Nuo (2) formulės, jis skiriasi vektorių komponentų reikšmėmis. Didžiausią įvertinimą gavusi klasė iš kiekvieno klasifikatoriaus, įgyja reikšmę tarp 0 ir 1, kai $s_{ij} = 1$. Šiuo atveju klasifikatoriai nėra laikomi lygiaverčiais. Todėl didesnę svorį turintis klasifikatorius turės didesnę įtaką galutiniams rezultatams tinkle. Pavyzdžiui, turint du klasifikatorius su tikslumu 0,9 ir 0,6 – balsų sumavimo metu, klasifikavimo rezultatai bus padauginami iš šių svorių. Todėl atitinkamai antrasis klasifikatorius (su 0,6 svoriu), galutinio sprendimo atžvilgiu, turės pateikti daugiau tos pačios klasės spėjimų, nei pirmasis – kitos klasės spėjimų.

Šis metodas gali būti pakeistas, į balsavimą įtraukiant pirmumo tvarką. Tokiu atveju balsavime gali dalyvauti ir II tipo klasifikatoriai, kurie pateikia rezultatų vektorius – klases, surūšiuotas pagal reitingą. Tokia metodo modifikacija vadinama Borda skaičiavimu (angl. *Borda Count*). Kadangi neatliekamos jokios modifikacijos po to kai pateikiami rezultatų vektoriai, formulė BC yra pati paprasčiausia:

$$BC_j(S_1, \dots, S_M) = \sum_{i=1}^M s_{ij} \quad (4)$$

Ji susumuoja kiekvienos klasės reikšmes, kurias pateikia klasifikatoriai. Nors tai atrodo paprastas sprendimas, tačiau jam pritaikius skirtingo tipo klasifikatorius, galima gauti vertus dėmesio rezultatus. I tipo klasifikatoriai pateiks tik vieną $s_{ij} = 1$. Kiti vektoriaus elementai bus lygūs 0. Šis metodas panašus į (2), tačiau neišsprendžia problemos, kai atsiranda vienodai balsų už kelias klases.

Sumuojant II tipo klasifikatorių vektorių balsus, s_{ij} perskaičiuojamas taip:

$$s_{ij} = |@ \leq (S_i, s_{ij})|/n \quad (4.1)$$

Siekiant suprasti šią išraišką, apirašykime struktūrą: $@op(S, k)$. Čia $@op$ - laikysime funkcija, kuri grąžina aibę elementų iš S vektoriaus ir tenkina sąlygą „ $s_j op k$ “. Pavyzdžiui, $@ \leq (S, 3)$ grąžins aibę elementų iš vektoriaus S , kurie yra mažesni arba lygus 3. Kaip ir įprasta, žymėjimas $|A|$, nusako aibės A kardinalumą (jos galią). Todėl (4.1) nurodo klasių, esančių žemiau jų rango skaičiaus, ir bendrą klasių skaičiaus santykį. Borda skaičiavimą gausime (4.1) pritaikius (4) formulėje. Šis būdas yra ekvivalentus daugumos taisyklei, kai klasifikavime dalyvauja lygiai dvi klasės [HHS92].

Kombinuojant III tipo klasifikatorių rezultatus naudoti (4) sumavimą, nėra tikslo. Taip yra dėl to, kad rezultatų vektoriai iš skirtingų klasifikatorių dažnai pateikia skirtingus intervalus. Todėl kombinavimo kokybė paprastai priklauso nuo gerai parinktų normalizavimo metodų.

Apibendrinimas. Nagrinėti balsavimo metodai leidžia sudaryti reikšmingas prielaidas. Iš jų išplaukia, kad skaičiavimai padaryti su klase C_j neturi jokios įtakos kitoms klasėms. Taip pat,

kiekvienas klasifikatorius visas klases atpažįsta vienodai gerai. *VT* algoritmas sudaro prielaidą – jog visi klasifikatoriai yra vienodo tipo, bei yra vienas nuo kito nepriklausomi [Lee95]. Akivaizdu, jei mes padubliuotume prasčiausios kokybės klasifikatorių keletą kartų, tai tų klasifikatorių atsakymų vektoriai dominuotų balsavimo metu. Didėjant tokių pačių klasifikatorių skaičiui tinkle, viso tinklo kokybė tampa priklausoma nuo tų klasifikatorių. Todėl šiuo atžvilgiu vadovautis tik balsavimo metodais nėra tikslu. Svarbų vaidmenį kooperuojant neuroninį tinklą atlieka klasifikatorių skirtingumo matavimai. Sekančiame poskyryje tai ir apžvelgsime.

2. Skirtingumo (diversiškumo) matavimai

Klasifikatorių skirtingumo sąvoka neuroninių tinklų kombinavime atsirado su pirmaisiais bandymais kombinuoti įvairių klasifikatorių atsakymus. Net ir dabar matuoti skirtingumą nėra lengva ir aiški užduotis, kadangi vis dar nėra aiškaus priimto apibrėžimo, kuris tiktų bendroms situacijoms. Didelį įnašą klasifikatorių skirtingumo matavimų srityje atliko L. Kuncheva ir J. Whitaker [KW02]. Jie išstudijavo 10 įvairių klasifikatorių skirtingumo matavimo metodų bei suskirstė juos į dvi grupes – porinius ir neporinius. Atlikę tyrimus su skirtingais metodais ir binariniais klasifikatorių atsakymais, tyrėjai bandė nustatyti sąsajas tarp klasifikatorių skirtingumo ir kolektyvo kokybės. Šiame poskyryje apžvelgsime kelis, gerus rezultatus davusius metodus, kuriuos vėliau panaudosime praktinėje šio darbo dalyje, į kombinavimą įtraukiant ir duomenų balinimo operaciją.

Intuityviai galima manyti, kad klasifikatoriai neuroniniame tinkle turi būti kiek įmanoma tikslesni ir neturi daryti atsitiktinių klaidų. Nors iš pirmo žvilgsnio tai atrodo paprasta, tačiau skirtingi matavimo metodai ir jų santykis su viso tinklo kokybe nėra aiškiai apibrėžtas. Tai viena iš problemų kuri dar nėra atsakyta. Įvairūs kombinavimo metodai kurie intuityviai skatina skirtingumą yra gana sėkmingi. Netgi tikslingai bloginant klasifikatorių kokybę, toks jų kombinavimas gali sukurti gana kokybišką kolektyvą. Tokiu principu yra paremtas AdaSkatinimo (angl. *AdaBoost*) algoritmas. Tačiau ne visada klasifikatorių diversiškumas duoda geresnius viso kolektyvo rezultatus. Nepaisant to diversiškumo kriterijus neuroninių tinklų kombinavime yra vienas iš kertinių parametrų. Todėl sėkmingai patinktas tinklo skirtingumo matavimas gali padėti sukurti kokybišką klasifikatorių.

2.1 Diversiškumas

Jei problemai spręsti turėtume kokybišką klasifikatorių kuris neklysta, tuomet praktiškai nereikėtų net kolektyvo. Tačiau, jei klasifikatorius daro klaidas, stengiamės jį papildyti su kitu klasifikatoriumi, kuris taip pat klysta. Tačiau jis turi klysti visai kitose vietose (su kitais požymiais). Todėl tokių klasifikatorių atsakymų diversiškumas yra svarbus reikalavimas bendrame kolektyve. Intuityviai galvojant, neuroninį tinklą siekiame kooperuoti iš kuo mažiau klystančių klasifikatorių. Ir tuo atveju kai jie klysta norime, kad jų klaidos būtų ne vienodos, o skirtingose srityse. Visa tai atlikti praktiškai yra sudėtinga užduotis ir jai spręsti nėra vieno universalaus metodo. Tuo labiau susieti klasifikatorių skirtingumą ir viso kolektyvo sėkmingą klasifikavimą – šiandien yra aktuali problema.

2.2 Skirtingumo matavimai

Yra sukurta nemažai metodų su tikslu matuoti klasifikatorių skirtingumą. Šie metodai yra skirstomi į dvi grupes pagal norimą gauti skirtingumo matavimų kiekį ar tikslumą. Viena metodų grupė – tai poriniai (angl. *Pairwise*) arba cikliniai (angl. *Round Robin*) matavimai. Porinis klasifikavimas yra klasių binarizavimo procedūra, kuri daugelio klasių klasifikavimo užduotį paverčia į seriją dviejų klasių užduočių. Matematiškai tai užsirašytų taip: jei turime kolektyvą iš L klasifikatorių, tai tas kolektyvas gali pateikti $L(L-1)/2$ porinių diversiškumo matavimo tarp klasifikatorių reikšmių. Norėdami gauti vieną bendrą viso tinklo matavimą, paprasčiausiai galime visus gautus matavimus suvidurkinti. Neporiniai (angl. *nonpairwise*) matavimai – sekanti metodų grupė, kur matavimai remiasi entropija arba kiekvieno kolektyvo nario koreliacija su suvidurkintais viso tinklo matavimais.

Klasifikavimo rezultatus, gautus iš bet kurių dviejų klasifikatorių tikimybiškai galima padalinti į keturias dalis (Lentelė 1):

- a) Požymių vektoriai suklasifikuoti teisingai abiejų klasifikatorių. Lentelėje pažymėta a .
- b) Požymių vektoriai suklasifikuoti teisingai pirmojo θ_i , klaidingai – antrojo θ_j klasifikatoriaus. Pažymėta raide b .
- c) Požymių vektoriai suklasifikuoti klaidingai pirmojo θ_i , teisingai – antrojo θ_j klasifikatoriaus. Pažymėta raide c .
- d) Požymių vektoriai suklasifikuoti klaidingai abiejų klasifikatorių. Lentelėje pažymėta d .

Lentelė 1. Dviejų klasifikatorių θ_i, θ_j tikimybiniai rezultatų pasiskirstymai [DSZ09]

	θ_j teisingas (1)	θ_j klaidingas (0)
θ_i teisingas (1)	a	b
θ_i klaidingas (0)	c	d

$$a + b + c + d = 1$$

Poriniai matavimai. Intuityviausias skirtingumo tarp klasifikatorių rezultatų matavimo metodas yra vadinamas prieštaravimo (angl. *Disagreement*) matavimu. 0 tipo klasifikatorių atsakymams šis matavimas yra lygus tikimybei dviem klasifikatoriams pateikti prieštarigus atsakymus. Tai užsirašo taip:

$$\theta_{i,j} = b + c \quad (5)$$

Dviejų klasifikatorių atveju, jei vienas iš jų atsakymą pateikė teisingai, kitas – klaidingai, tai tokie klasifikatoriai yra skirtingi. Nepriklausomai nuo to kuris klasifikatorius melavo, matavimo rezultatas bus lygus 1. Šiuo matavimu T. K. Ho darbe buvo nagrinėjamas sprendimų medžių diversiškumas bei jo įtaka viso kolektyvo klasifikavimo kokybei [Ho98]. Norint gauti viso kolektyvo klasifikatorių skirtingumo matavimą *dis*, reikia suvidurkinti visų porinių klasifikatorių matavimus. Tai užsirašo formule (5.1). Kolektyvas yra tuo skirtingesnis kuo matavimo rezultatas aukštesnis. Čia L – klasifikatorių skaičius. Sumuojamai dviejų klasifikatorių θ_i, θ_j tikimybiniai rezultatų pasiskirstymai pagal 1 lentelės pavyzdį:

$$dis = \frac{2}{L(L-1)} \sum_{i=1}^{L-1} \sum_{j=i+1}^L \theta_{i,j} \quad (5.1)$$

Kitas matavimas vadinamas dvigubos klaidos (angl. *Double-fault*). Intuityviai galima suvokti jog šis matavimas gražina dviejų klasifikatorių tikimybę, kai jie abu klysta. Lent. 1 žymima raide d . Matematiškai tai užsirašo taip:

$$\theta_{i,j} = d \quad (6)$$

Šis matavimas intuityviai kilo iš idėjos - du klasifikatoriai turi elgtis skirtingai, kad jie būtų diversiški. G. Giacinto ir F. Roli tvirtina – kuo skirtingesni yra klasifikatoriai, tuo mažiau atsitiktinių klaidų jie daro [GR01]. Viso kolektyvo skirtingumas yra skaičiuojamas taip pat kaip ir prieštaravimo atveju – formulė (5.1). Šiuo atveju kuo matavimo rezultatas yra mažesnis – tuo didesnis yra diversiškumas.

Neporiniai matavimai. Panagrinėsime vieną neporinį kolektyvo skirtingumo matavimą – entropiją. Ji užsirašo formule:

$$E = \frac{2}{N(L-1)} \sum_{j=1}^N \min \left\{ \left(\sum_{i=1}^L y_{j,i} \right), \left(L - \sum_{i=1}^L y_{j,i} \right) \right\} \quad (7)$$

Kolektyvas yra laikomas labiausiai diversišku, kai konkretų požymių vektorių pusė kolektyvo narių suklasifikavo teisingai ir kita pusė narių – klaidingai. Jei visi nariai konkretų požymių vektorių suklasifikavo klaidingai, arba teisingai – toks kolektyvas nėra diversiškas. (7) formulėje E kinta tarp 0 ir 1. Kuo yra skirtingesnis kolektyvas – tuo matavimas bus arčiau 1.

Be Entropijos matavimo, egzistuoja ir KW-variacinis, trapusavio sutarimo, sudėtingumo, generalizuoto skirtingumo, atsitiktinės klaidos ir kiti skirtingumo matavimo metodai [KW02][TSY06].

2.3 Diversiškų tinklų kūrimas

A. Sharkey savo darbe nagrinėja 4 esmines parametrų kategorijas, kurios įtakoja kolektyvo diversiškumą [Sha98]. Tai pradiniai svoriai, mokymo duomenys, neuroninių tinklų architektūra ir parinktas mokymo algoritmas. Pritaikius skirtingus parametrus kiekvienam klasifikatoriui atskirai ir juos apjungus į bendrą kolektyvą – gausime skirtingus matavimų ir klasifikavimo rezultatus. Tuo tarpu G. Brown išskiria tris platesnes hipotetines kategorijas, kurios apima daugumą neuroninių tinklų kombinavimo būdų [Bro04]. Jis kategorijas laiko tam tikromis hipotezių erdvėmis: pradžios erdvė (ang. *Starting-point*), aibė pasiekiamų hipotezių ir hipotezių erdvės perėjimas (angl. *traversal*).

Pradinė erdvė. Vienas iš kolektyvo skirtingumo įtakojančių pradinės erdvės parametrų – yra pradiniai neuronų svoriai. Jų keitimas, tai bene paprasčiausias būdas kombinuoti neuroninius tinklus. Svorius galima parinkti tiek atsitiktinai, tiek naudojant įvairius metodus. G. Brown pastebi, kad pradinių svorių parinkimo būdas yra mažiausiai efektyvus stengiantis pasiekti gerą diversiškumą. Su pradiniais svoriais ir jų ivairiomis kombinacijomis, buvo atlikta keletas bandymų

ties su sintetiniais, ties su realaus pasaulio duomenimis. Buvo prieita išvados: lyginant su kitais parametrais (paslėptų sluoksnių skaičiumi, mokymo duomenų struktūra, tinklo architektūra ir kt.), atsitiktinis pradinių svorių parinkimas turi mažiausiai įtakos generuojant diversiškumą.

Hipotezių erdvė. Šią erdvę galima padalinti į kelias sritis, kurios skirtingai įtakos kolektyvo diversiškumą, o taip pat ir kokybę. Vienoje srityje galima keisti mokymo duomenų pateikimą kolektyvui, kitoje – keisti pačią kolektyvo architektūrą.

Labiausiai paplitę ir išnagrinėti metodai, siekiantys sukonstruoti skirtingesnius arba vienas kitą papildančius neuroninius tinklus, dažnai yra paremti skirtingomis mokymosi duomenų aibėmis. T.y. skirtingus kolektyvo narius mokome su duomenimis iš tos pačios požymių vektorių erdvės su tikslu, kad skirtingi klasifikatoriai išmokyti skirtingų dalykų, tačiau iš tos pačios srities. Vieni metodai tai leidžia padaryti mokant kolektyvo narius su skirtingais požymių vektoriais. Kiti – padalina požymių erdvę dalimis ir kiekvieno vektoriaus dalį atiduoda skirtingam kolektyvo nariui. Šie du būdai dar vadinami – pakartotine atranka (angl. *resampling*). Pakartotinė atranka taip pat gali būti atliekama įvairiai. Tiek požymius ar vektorius imant nuosekliai, tiek juos perdengiant. Dar vienas būdas įtakoti klasifikatorius – tai požymių perdirbimas. Metodai kurie tai atlieka vadinami deformuojančiais (angl. *distortion*). Vėliau nagrinėsime duomenų balinimo metodą, kuris požymius erdvėje padaro rutulio formos. Šio metodo įtaką kolektyvo atsakymų kokybei su skirtingais ekspertais ir įvairiomis neuroninio tinklo kombinacijomis nagrinėsiu praktinėje šio darbo dalyje.

R. Duin ir D. Tax atliko tyrimus su ranka rašytinių skaitmenų klasifikavimu ir priėjo prie išvadų [DT00]:

- Kokybiškesnį klasifikavimą galima pasiekti kombinuojant kolektyvo narius apmokytus su skirtingais požymių vektoriais.

- Strategija skaldyk ir valdyk – nepriklausomų požymių išskyrimas, duoda geresnius rezultatus. Sudėtingi duomenys neturi būti išmesti – jie kartais pateikia svarbią informaciją, tik juos reikia tinkamai apdoroti. Atsitiktinai parinkti požymiai, konkretaus bandymo metu davė labai gerus rezultatus. Ypatingai taikant Bajeso klasifikatorių.

- Kombinuojant atsakymus vieno tipo klasifikatoriaus, apmokyto su skirtingais požymiais yra kur kas efektyviau, nei kombinuojant atsakymus skirtingų klasifikatorių, apmokytų su tais pačiais požymiais.

Tyrimų atliktų su skirtingomis neuroninių tinklų architektūromis ir jų kombinacijomis nėra daug. D. Partridge teigia, jog paslėptų neuronų sluoksnyje skaičiaus variacijos nedarė didelės įtakos kolektyvo diversiškumui. Tačiau toks būdas yra geresnis už atsitiktinių pradinių svorių parinkimą. Kadangi tyrimas buvo atliekamas tik su 8-12 paslėptų mazgų variacijomis ir su vieno tipo mokymo

duomenimis, tai nėra taikytina išvada kitiems tyrimams. Partidže taip pat taikė daugiasluoksnius perceptronus ir radialines bazines funkcijas kolektyvo sudarymui. Šis bandymas buvo kur kas sėkmingesnis už paslėptų neuronų kiekio varijavimą.

Taip pat dažnai yra kombinuojami ne tik neurono paslėpti sluoksniai ar jo struktūra, bet ir skaičius neuroninių tinklų pačiame kolektyve. Toks kolektyvo kombinavimas paprastai yra atliekamas projektuotojo, o tai reikalauja gana geros žinių bazės ir patirties neuroninių tinklų srityje. Į tai atsižvelgdami D. Optiz ir J. Shavaliik panaudojo genetinį Addemup algoritimą, kuris optimizuoja tinklo topologiją sudarydamas kolektyvą [OS96]. Šis algoritmas mokantis naudoja atvirkštinį sklidimą, tuomet pasirenka diversiškas neuroninių tinklų grupes remdamasis skirtingumo matavimais. Z. Wang su kolegomis patobulino CNNE (angl. *Constructive Neural Network Ensemble*) algoritimą, kuris konstruoja neuroninių tinklų kolektyvą, eigoje stebėdamas diversiškumą. Šis algoritmas lygiagrečiai projektuoja kolektyvo architektūrą ir moko atskirus neuroninius tinklus. Kolektyvo narių mokymui yra naudojamas neigiamos koreliacijos metodas. Wang neuroninio tinklo kokybei testuoti naudojo medicininius duomenys. Rezultate gavosi gana tikslus klasifikavimas.

Be šių minėtų kombinavimo metodų, egzistuoja ir begalė kitų, tarp kurių galima rasti nemažai bandymų kombinuoti hibridinius neuroninius tinklus. Tai atliekama į kolektyvą įtraukiant skirtingus klasifikatorius: sprendimų medžius, k-arčiausio kaimyno ir kitus klasifikatorius. Tyrimai rodo jog toks kombinavimas leidžia geriau atskirti požymius erdvėje ir sukurti daugelio ekspertų kolektyvą su mažesne klasifikavimo klaida, lyginant su pavieniais klasifikatoriais.

Hipotezių erdvės perėjimas. Turint hipotezių erdvę, kuri yra sudaryta iš neuroninių tinklų architektūros ir pateiktų mokymo duomenų, galima imti vieną tašką toje erdvėje, kuris atspindės tam tikrą hipotezę (klasifikatoriaus parinkimą). Tai kaip eisime ta erdve tarp skirtingų hipotezių, leis sudaryti bendrą galutinį kolektyvą. Toks G. Brown pasiūlytas neuroninių tinklų kombinavimo būdas erdvėje, leidžia išskirti kombinavime naudojamus etapus bei lengviau suvokti patį kombinavimo procesą.

3. Kolektyvo topologijos

A. Sharkey tvirtina jog neuroninių tinklų kolektyvo topologija yra antras pagal svarbą kriterijus kombinuojant neuroninius tinklus [Sha98]. Pirmoje vietoje yra kolektyvo narių diversiškumas, apie kurį plačiau galima paskaityti 2 darbo skyriuje. L. Kuncheva teigia, kad taip pat svarbu yra parinkti gerą kombinavimo metodiką (angl. *fusion*), kuri turi didelę įtaką kolektyvo kokybei. Kombinavimo

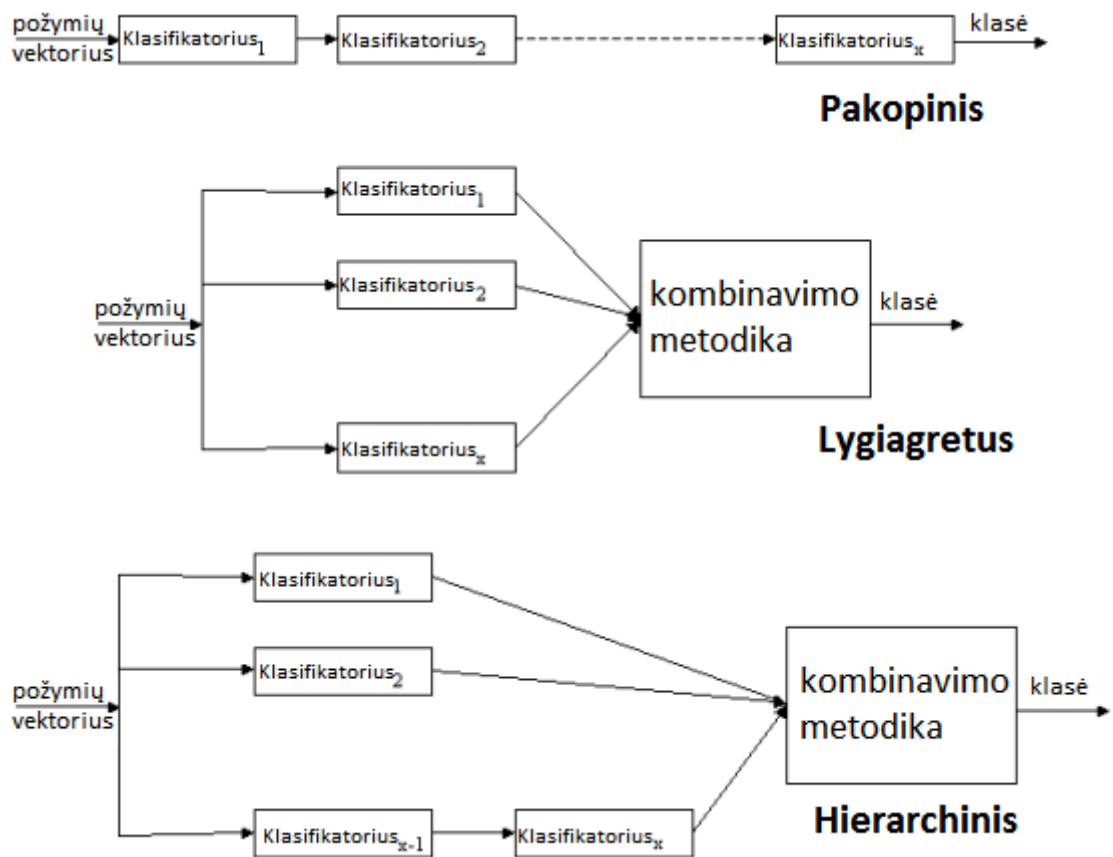
metodikos buvo nagrinėtos 1.5 skyrelyje. Tačiau prieš parenkant tinkamą kombinavimo metodiką yra svarbu suprojektuoti gerą viso tinklo išdėstymą. Gera strategija projektuojant kolektyvą yra padalinti jį į kompetencijos sritis. T.y. kiekvienas klasifikatorius ar jų kolektyvo pogrupis turi būti savo srities ekspertas. Tada klasifikuojant naujas užduotis yra atsižvelgiama į kiekvienos srities ekspertų nuomones. Pavyzdžiui turint dvi požymių erdves temperatūros ir drėgnumo matavimus, galime sudaryti kolektyvą iš dviejų ekspertų, kurie gebėtų prognozuoti sekančios dienos oro sąlygas vertinamas kažkoku tai matavimu. Vienas ekspertas atliktų prognozavimą remiantis temperatūra, kitas – drėgnumu. Tuomet apjungus abu ekspertus kombinavimo taisykle turėtume gauti tikslesnę prognozę. Tačiau tokiu atveju atsiranda problemos, kai ekspertas pradeda meluoti.

Tokias problemas galima sumažinti teisingai įvertinus kiekvieną ekspertą. Tai gali būti atliekama mokant kolektyvą ir analizuojant pateiktus rezultatus. Visa tai atliekama neuroninio tinklo mokymo fazėje. Kitas tokių problemų sprendimas yra teisingai parinkti duomenų srauto kelią prieš sprendimų pateikimą. Tokiu atveju pirmiausiai gali būti atliekami kelių kolektyvo narių sprendimai, kurie perduodami kitam pogrupiui kolektyvo narių ir t.t. Tinkama tokios struktūros konstrukcija gali pagerinti viso kolektyvo kokybę.

Y. Lu topologijas išskiria į tris grupes (pav. 1): pakopines, lygiagrečias ir hierarchines [Lu96]. Pakopinėje kolektyvo topologijoje, pirmo nario klasifikavimo rezultatas yra duodamas sekančiam klasifikatoriui. Taip kiekvieno klasifikatoriaus rezultatai yra perduodami sekantiems klasifikatoriams, kol gaunamas atsakymas iš paskutinio grandyje esančio nario. Vienintelis didesnis šio tinklo trūkumas – tai negalėjimas sekantiems klasifikatoriams grandyje pataisyti klaidas prieš tai buvusių kolektyvo narių.

Lygiagretus klasifikatorių išdėstymas leidžia integruoti jų atsakymus į vieną vietą. Projektuojant tinklą šia topologija, svarbiausias dalykas yra parinkti tinkamą kombinavimo metodiką. Jei toks tinklas yra gerai suprojektuotas, galima pasiekti labai gerus klasifikavimo rezultatus. Kadangi tokia tinkle svarbų vaidmenį atlieka kombinavimo metodika, netinkamai ją parinkus, gali pasireikšti tinkle esantys blogesnės kokybės klasifikatoriai ir finale suprastinti viso kolektyvo kokybę.

Hierarchinis tinklas apima pakopines ir lygiagrečias topologijas ir įvairias jų konfigūracijas. Toks išdėstymas gali padėti išvengti pakopinio tinklo nuostolių. Taip pat tokie kolektyvai gali lengviau susitvarkyti su klaidų tikrinimu ir rezultate sumažinti blogai klasifikuojančių narių įtaką kolektyvui.



Pav. 1. Trys klasifikatorių kabinavimo variantai (topologijos) [Lu96].

L. Lam pateikia bendresnę ir literatūroje dažniau randamą topologijų klasifikavimo modelį [Lam00]. Topologijos čia skirstomos į keturias kategorijas: sąlygines, hierarchines, hibridines ir sudėtines.

Sąlyginė topologija. Ši strategija pirmiausia pasirenka vieną klasifikatorių ir atlieka klasifikavimą. Jei šis klasifikatorius neidentifikuoja teisingos klasės – imamas kitas klasifikatorius. Praktikoje dažnai yra pasirenkama kuris klasifikatorius yra pirminis. Sekančio klasifikatoriaus pasirinkimas paprastai būna arba statinis arba dinaminis – jo pasirinkimas priklausomas nuo pirmojo klasifikatoriaus pateiktų atsakymų. Klasifikatorių pasirinkimą gali apspręsti įvairūs metodai (pvz. Sprendimų medžiai). Toks pasirinkimo procesas gali tęstis, kol yra klasifikatorių arba kol požymių vektorius yra teisingai suklasifikuotas. Jei pirmasis klasifikatorius yra kokybiškas, tai tuomet ir visų požymių klasifikavimas tampa efektyvus skaičiavimų atžvilgiu. Todėl yra naudinga sudaryti klasifikatorių eilę nuo kokybiškiausio iki mažiausiai kokybiško klasifikatoriaus. Tokio tinklo didžiausia problema yra apspręsti, kaip bus teisiama arba skatinami klasifikatoriai, kurie

atitinkamai blogai arba gerai klasifikavo. Taip pat augant klasifikatorių skaičius kolektyve, visas procesas gali tapti labai komplikuoatas.

Hierarchinė (nuosekli) topologija. Šioje struktūroje yra pasirenkami kolektyvo nariai kurie klasifikavo sėkmingai. Kiekvienas klasifikatorius naudojamas sumažinti galimų klasių skaičių, kurioms priklauso požymių vektoriai. Su kiekviena naujų požymių iteracija, klasių kiekis tampa labiau fokusuotas, siauresnis. Formuojant tokio kolektyvo narių eilę, klasifikatoriai yra rikiuojami mažėjančia klasifikavimo klaidos tvarka. T.y eilėje pirmieji klasifikatoriai yra tie, kurių klaida didžiausia. Taip pat turi būti saugikliai, kurie užtikrins, kad tarp bet kurio klasifikatoriaus atmestų klasių nebuvo tikrosios, teisingos klasės. Priešingu atveju, sekantis klasifikatorius negalėtų pateikti teisingos klasės, nes jos nebūtų tarp pirmesniojo klasifikatoriaus klasių pasirinkimo.

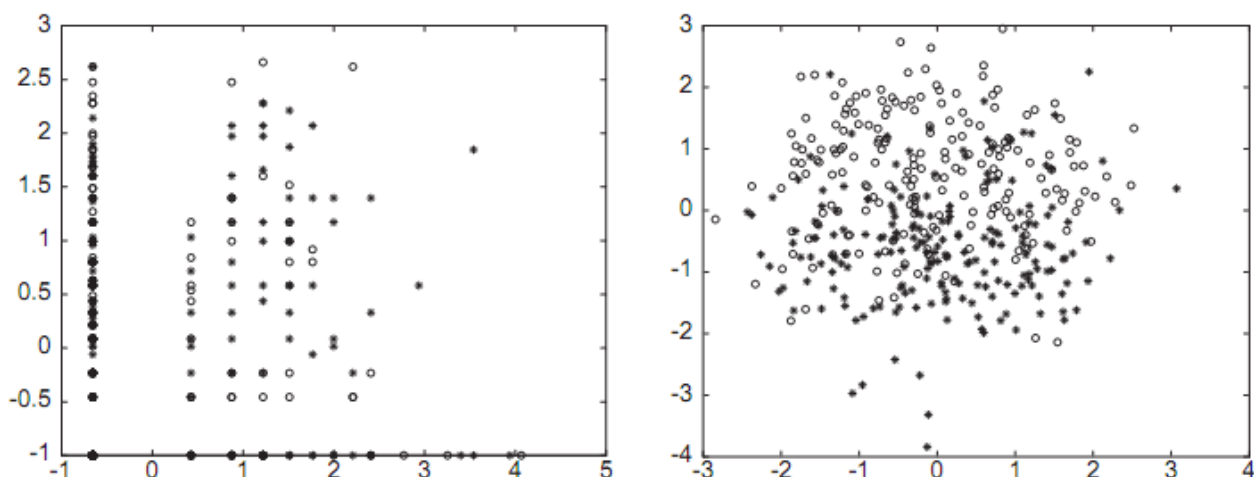
Hibridinė topologija. Hibridinis kolektyvas parenka geriausią klasifikatorių pagal pateiktą požymių vektorių. Šio tinklo idėja yra ta, kad ekspertas geriausiai gali klasifikuoti savo srityje, todėl kiti klasifikatoriai konkrečiam požymių vektoriui nereikalingi. Toks kolektyvas, nepriklausomai nuo ekspertų parinkimo, supaprastina klasifikavimo procesą.

Sudėtinė (lygiagreti) topologija. Tai praktikoje dažniausiai naudojama topologija. Klasifikatoriai šioje struktūroje veikia lygiagrečiai ir jų atsakymai vėliau apdorojami įvairiais metodais. Tai praktiškai tas pačias savybes įgyjanti topologija kaip ir Y. Lu atveju nagrinėtas lygiagretus modelis.

Hierarchinės ir sudėtinės topologijos dar yra vadinamos atitinkamai paremtos pasirinkimu (angl. *selection-based*) ir kombinavimu (angl. *fusion-based*).

4. Duomenų balinimas

Duomenų balinimo (angl. *whitening*) sąvoka atėjo iš signalų apdorojimo srities. Duomenų balinimo transformacija duomenis erdvėje paverčia rutulio formos (pav. 2). Rezultate gauname kiek įmanoma labiau nekoreliuotus požymius erdvėje [DHS97].



Pav. 2. Pirmas iš kairės – originali požymių erdvė, dešinėje – požymių erdvė atlikus duomenų balinimo transformaciją. „*“ ir „o“ žymi dviejų klasių požymius, kurie erdvėje buvo pastumti ir pasukti. Abscisių ir ordinačių ašyse vaizduojami du skirtingi požymiai [RI04].

Duomenų balinimo operaciją Š. Raudys ir M. Iwamura taikė savo tyrime siekiant atpažinti ranka rašytinius japoniškus simbolius, kurių poros yra panašios vienas į kitą [RI04]. Darbe buvo naudojami trys klasifikatoriai. Siekiant juos padaryti robastiškus, buvo panaudota statistinė ir neuroninių tinklų kombinacija (duomenų balinimo transformacija ir vienasluoksnio perceptrono mokymas). Atlikus įvairias neuroninių tinklų kombinacijas, buvo gautas gana kokybiškas klasifikatorius. Tyrimo išvadose teigiama jog visi trys klasifikatoriai buvo svarbūs ir jų eliminavimas būtų tik neigiamai įtakojęs rezultatus.

Duomenų balinimo operaciją galima įvardinti kaip koordinačių transformaciją, kuri konvertuoja daugiamatį normalųjį skirstinį į sferinį. T.y tokį, kurio kovariacijos matrica yra proporcinga tapatumo matricai I . Apsibrėšime matricą Φ , kurios stulpeliai yra ortonormuoti (skirtingų stulpelių skaliarinė sandauga yra lygi 0, o kiekvieno stulpelio skaliarinė sandauga su juo pačiu yra lygi 1) kovariacijos matricos Σ tikriniai vektoriai. Λ – atitinkamų tikrinių verčių diagonalinė matrica. Tuomet atliekama transformacija originalių požymių erdvėje:

$$A_w = \Phi^T \Lambda^{-1/2} \quad (8)$$

Transformuotos erdvės pasiskirstymas turi kovariacijos matricą lygią tapatumo matricai. Transformacija (8) yra vadinama duomenų balinimu, kadangi ji tikrinių vektorių spektrą paverčia

vienodo dažnio pasiskirstymu. T.y požymių erdvę transformuoja arti balto triukšmo. Š. Raudys teigia, jog dažnai vienasluoksniu perceptrono mokymas ant išbalintų duomenų ženkliai paspartėja.

Toliau praktinėje šio darbo dalyje stengsiuosi atsakyti į klausimus kaip duomenų balinimo operacija įtakoja neuroninių tinklų kokybę. Balinimo transformaciją taikysiu ne tik pradinei požymiu erdvei transformuoti, tačiau atliksiu bandymus ir balinant klasifikatorių aibių atsakymus. Visą tai apjungus į bendrą kolektyvą su skirtingais ekspertais, remiantis pseudo sintetinių duomenų ir realaus pasaulio problemomis ieškosiu kombinacijų, duodančių gerus klasifikavimo rezultatus.

5. Eksperimento duomenų atranka, eiga ir rezultatai

Šiame skyriuje aprašoma praktinė darbo eiga ir gauti rezultatai. 1.1 Poskyryje apibrėžiamas šio eksperimento tikslas. 1.2 Poskyryje aprašomi eksperimente naudojami įrankiai ir bibliotekos. Sekančiame poskyryje nurodomos esminės eksperimento dalys ir jų eiga. 1.4 Poskyryje nurodomos ir paaiškinamos eksperimente naudojamos duomenų imtys. 1.5 Poskyryje atliekami eksperimentai, tiriantys duomenų balinimo transformacijos įtaką neuroninių tinklų kombinavime. 1.5.1 Tiriamas balinimo transformacijos klasifikatorių kokybei, juos kombinuojant skirtingais metodais. 1.5.2 Poskyryje atliekami bandymai su skirtingais diversiškumo matavimo metodais ir aprašomas balinimo transformacijos efektas klasifikatorių diversiškumui. 1.5.3 Poskyryje tiriamas ekspertų atsakymų aibės balinimo poveikis klasifikatoriaus kokybei. Tyrime nagrinėjamos geros kokybės skirtingų ir blogos kokybės panašių ekspertų grupės. Į tyrimą įtraukiamas tiesinio perceptrono ekspertas, kaip kombinavimo metodas. Skyriaus pabaigoje pateikiamos bendros eksperimento išvados ir esminiai tyrimo pastebėjimai.

5.1. Eksperimento tikslas

Eksperimento tikslas atspindi šio darbo nagrinėjamą sritį ir pateikia praktiškai pasiektus rezultatus. Literatūroje didelis dėmesys yra skiriamas į nagrinėjamos srities matematinius modelius, tačiau retai sutinkama išsamesnė nagrinėjamos srities eksperimentinė dalis, kuri atspindėtų konkrečius pavyzdžius. Todėl šiame skyriuje pagrindinis dėmesys yra skiriamas eksperimento eigai.

Nagrinėjami esminiai eksperimentą įtakojantys faktoriai yra šie: klasifikatorių diversiškumas, balinimo transformacijos (BT) taikymas skirtinguose neuroninio tinklo kūrimo etapuose, neuroninio tinklo kombinavimo metodikos su skirtingomis duomenų imtimis. Siekiant įvertinti šių faktorių įtaką klasifikavimo rezultatams, eksperimento eiga yra suskirstyta į mažesnius eksperimentus, atspindinčius konkrečias neuroninio tinklo kombinacijas ir bandymus su jomis. Šios kombinacijos savo ruožtu gali pateikti gana konkrečius rezultatus. Todėl siekiant gauti kuo tikslesnius rezultatus,

kurie atspindėtų bendresnes situacijas, eksperimentai bus atlikti su kuo įvairesniais duomenų rinkiniais (pseudo sintetiniais ir realaus pasaulio). Duomenų įvairovė leis vertinti platesniu aspektu eksperimentuose taikomas metodikas ir veikiančius faktorius.

5.2.Naudojami įrankiai ir bibliotekos

Eksperimentui atlikti buvo pasirinkta matricų laboratorija – MATLAB [LM84]. Šis paketas yra patogus savo galimybėmis greitai atlikti įvairias matematinės operacijas ir atvaizduoti duomenis grafikų pagalba. MATLAB savyje turi pagrindines neuroniniuose tinkluose naudojamas funkcijas kartu su standartiniu neuroninių tinklų įrankių paketu. Tačiau eksperimentiniam darbui pagreitinti bus naudojamas PRTools MATLAB instrumentinių priemonių komplektas [DUI93], kuris yra skirtas atpažinimo (klasifikavimo) uždaviniams spręsti. Savyje jis turi komandų rinkinį kuris leidžia greitai ir patogiai operuoti duomenų imtimis, jomis kurti ir apmokyti pagrindinius neuroninius tinklus, juos kombinuoti, stebėti klasifikavimą ir paklaidas grafikuose ar kitose matematinėse struktūrose.

Šis įrankių ir bibliotekų komplektas leidžia susikoncentruoti ties eksperimento tikslu ir negaišti laiko įvairių funkcijų rašymui. Tyrime naudojamos MATLAB ir PRTools pagrindinės komandos bus paaiškintos siekiant pristatyti jų esmę ir panaudojimą eksperimente.

5.3.Eksperimento planas ir eiga

Tyrimo dalis yra suskirstyta į eksperimentus. Kiekvienas eksperimentas turi savo eigos dalis: siekiamą tikslą (hipotezę), praktinę dalį ir išvadas.

Siekiamas tikslas. Tai įvadinė eksperimento dalis, kurioje iškeliamą hipotezę. Šios dalies tikslas yra prognozuoti galimus rezultatus remiantis jau atliktas tyrimo bandymais ir pateikti tolimesnius eksperimento principinius veiksmus rezultatams pasiekti.

Praktinė dalis. Šioje dalyje atliekamas pats eksperimentas. Aprašoma eksperimento eiga, naudojami algoritmai ir matematiniai modeliai. Ši eksperimento dalis gali būti skaidoma į bandymus, siekiant pabrėžti konkrečių bandymų esminius modelius, rezultatus ir jų įtaką eksperimento eigai. Eksperimentai atlikti su skirtingais duomenų rinkiniais apibendrinami lentelėse. Grafiškai pateikiami esminiai tarpiniai ir galutiniai eksperimento rezultatai.

Išvados. Tai dalis, kurioje pateikiamos eksperimento išvados. Trumpai apibendrinamos esminės eksperimento vietos ir pateikiami gauti rezultatai. Patvirtinamos arba paneigiamos prielaidos iškeltos pradinėje eksperimento dalyje.

5.4. Tyrime naudojamos duomenų imtys

Bandymams atlikti buvo pasirinkta 10 duomenų imčių tipų. 4 imtys – tai atsitiktinai sugeneruotos imtys, turinčios tam tikrą duomenų formą erdvėje. Kitos 6 – grįstos realaus pasaulio problemomis. Jos paimtos iš UCI duomenų talpyklos [BKM99]. Tai populiariausia Kalifornijos universiteto mašinų mokymosi tikslams skirta duomenų saugykla, kurioje sukaupta virš 200 įvairių duomenų rinkinių.

Pirmuose bandymuose naudojamos atsitiktinai sugeneruotos imtys, pasinaudojant PRTools bibliotekos funkcijomis:

- *Gendatb*. Ši duomenų imtis turi banano formos išsidėstymą erdvėje. Ją sudaro 2 klasės. Jų tikimybės vienodos – 0,5. Ši imtis susideda iš dviejų požymių.
- *Gendatc* generuoja n požymių, 2 klasių duomenų imtį, kurios erdvėje yra paskirstytos pagal Gauso skirstinį. Klasių tikimybės yra taip pat po 0,5. Pagal nutylėjimą šių klasių duomenys persidengia 16%.
- *Gendatd* komanda sukuria n požymių, 2 klasių duomenų rinkinį, kur dviejų požymių dispersija yra labai skirtinga. Pirmieji du požymiai yra pasukti 45 laipsnių kampu, todėl atsiranda stipri koreliacija. Klasių tikimybės vienodos – 0,5.
- *Gendats* generuoja n požymių, 2 klasių duomenų imtį. Generuojant duomenis galima nurodyti atstumą, kuris skiria 2 klasių vidurkius. Pagal nutylėjimą atstumas lygus 1. Požymių vektorių pasiskirstymas lygus.

Sekančios duomenų imtys paimtos iš realaus pasaulio problemų:

- *Iris* – tai turbūt labiausiai paplitusi duomenų imtis atpažinimo srityje. Ją pristatė Fišeris – vienas iš pirmųjų statistinio klasifikavimo pradininkų. Duomenų imtį sudaro 3 klasės po 50 vektorių. Kiekviena klasė – tai tam tikras Iriso augalų tipas. Viena iš trijų klasių yra tiesiškai atskiriama nuo sekančių dviejų, kurios persidengia. Vektorius sudaro 4 požymiai.
- *Breast* duomenų imtį sudaro 699 vektoriai su 9 požymiais. Imtis klasifikuojama į dvi klases. Klasėse vektoriai pasiskirstę į 458 ir 241 atvejus. Duomenys imčiai surinkti iš Viskonsino ligoninių daktaro William H. Wolberg. Užduoties tikslas yra nustatyti kokio tipo krūties vėžys (piktybinis ar ne).
- *Wine* duomenų imtį sudaro 13 cheminių vyno požymių, pagal kuriuos imtis dalijama į 3 klases. Užduotis yra nustatyti kuriam iš trijų kultivarui Italijos vynuogynuose priklauso vinas. Duomenų imtį sudaro viso 178 vektoriai.

- *Satellite* – tai didžiausia pasirinkta duomenų imtis kurioje yra net 6435 požymių vektoriai. Juos sudaro 36 požymiai. Pagal juos imtis klasifikuojama į 6 klases. Imtį sudaro supaprastinti skaitmeninių nuotraukų vektoriai, kurie remiantis spektrine analize, identifikuoja vieną iš šešių žemės dirvožemių tipų.
- *Diabetes* duomenų imtį sudaro 786 vektoriai su 8 požymiais. Jie klasifikuojami į 2 klases. Tai Indijos gyventojų biologiniai parametrai, kuriais remiantis siekiama nustatyti ar žmogus yra linkęs į diabeto ligą.
- *Ecoli* sudaro 336 požymių vektoriai su 7 požymiais. Ši turi 8 klases, kurių dydis stipriai varijuoja požymių vektorių atžvilgių. Mažiausiai vektorių sudaro 2 klases - po 2 požymių vektorius. Didžiausia klasė turi 143 vektorius.

5.5. Praktiniai tyrimai

Siekiant ištirti balinimo transformacijos (BT) poveikį tiesiniam neuroninių tinklų kombinavimui, svarbu atlikti ne vieną eksperimentą su skirtingomis duomenų imtimis ir ekspertų kombinacijomis. Svarbu pabrėžti, jog kartojant eksperimentą su skirtingais parametrais ar duomenimis daug kartų, jo rezultatų apibendrinimas tampa tikslesnis. Nepaisant to konkrečios situacijos gali pateikti skirtingus rezultatus.

Šiame poskyryje stengsiuosi atsakyti į klausimą, kaip BT įtakoja neuroninių tinklų mokymąsi konkrečiose situacijose. Nėra apibrėžtai aišku kaip elgiasi ekspertų kombinacijos, apmokytos su išbalintais duomenimis ir su originaliais duomenimis. Ar ekspertai skiriasi, ar yra vienodi? Tyrime bus vertinamas ekspertų diversiškumas, kuris matuojamas skirtingais metodais. Juos apžvelgėme literatūros analizės dalyje. Kaip ekspertų diversiškumas įtakoja kolektyvo parinkimo metodiką ir galutinius viso kolektyvo rezultatus? Ar taikant balinimo transformaciją klasifikatorių atsakymams gaunami geresni klasifikatoriai?

5.5.1 Balinimo transformacijos taikymas kombinuojant neuroninius tinklus

Yra atlikta nemažai tyrimų, nagrinėjančių klasifikatorių skirtingumą ir jų įtaką kolektyvo atsakymams [KW02], [Ho98], [GR01]. Tyrimo rezultatai yra skirtingi konkrečiose situacijose, tačiau pastebėtas dėsningumas – skirtingesni ekspertai daro mažiau atsitiktinių klaidų. Taip pat pastebėta, jog duomenų balinimo transformacija, atlikta mokymosi duomenų imčiai – paspartina neuroninio tinklo mokymosi laiką ir neretai duoda geresnius rezultatus [RI04]. Iš čia kyla *pirmasis eksperimentas*:

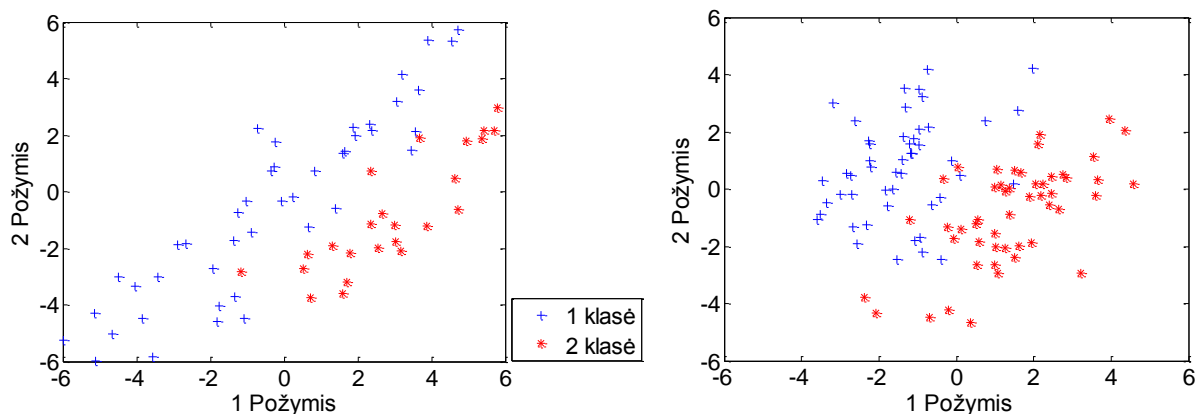
Siekiamas tikslas. Atsižvelgiant į minėtus tyrimus, galime daryti prielaidą, jog skirtingų ekspertų klasifikatorius, apmokytas su išbalintais duomenimis, duos geresnius rezultatus – klasifikatorius bus geresnės kokybės, nei klasifikatorius, išmokytas su originaliais (netransformuotais) duomenimis.

Šis eksperimentas padalintas į tris bandymus. Pirmame bandyme surenkami skirtingi ekspertai iš trijų tipų neuroninių tinklų. Atliekami bandymai su pseudo sintetiniais originaliais ir balintais duomenimis. Antrame bandyme sudaromas klasifikatorius ir su tais pačiais duomenimis atliekami bandymai su skirtingais kombinavimo metodais. Trečiame bandyme įtraukiama 10 skirtingų duomenų imčių ir pakartojamas antrasis bandymas.

Praktinė dalis. 1 bandymas. Norint sudaryti klasifikatorių, pirmiausiai surenku jo narius. Tam tikslui parenku 15 ekspertų. Toks ekspertų skaičius parinktas siekiant varijuoti skirtingais parametrais. Pirmoje darbo dalyje, neatsižvelgiant į duomenų kilmę, parenku skirtingus ekspertų parametrus. Jie yra vienas iš kriterijų, stipriai įtakojančių diversiškumą (jį nagrinėsiu sekančiame poskyryje). Ekspertai sudaromi iš trijų tipų neuroninių tinklų: 5 nariai iš daugiasluoksnių tiesioginio sklidimo, 5 – iš daugiasluoksnių grįžtamojo ryšio ir 5 – vienasluoksnių perceptronų. Kiekvienam ekspertui parenkami skirtingi pradiniai parametrai.

Sekančius šio eksperimento bandymus atlieku su atsitiktinai sugeneruotais duomenimis. Duomenims sugeneruoti pasinaudoju PRTools bibliotekos funkcija *gendatd*. Duomenys erdvėje sugeneruojami pasukti 45 laipsnių kampu (pav. 3 kairėje). Gaunu dviejų klasių duomenų imtį su dviem požymiais. Imtyje sugeneruota 100 vektorių (50 vektorių priklauso vienai klasei, 50 – kitai).

Pasidarę šių duomenų kopiją – jiems pritaikome balinimo transformaciją (pav. 3 dešinėje). Tokiu būdu duomenys yra dekokoreliuojami. T.y jų kovariacijos matricos diagonalios reikšmės tampa lygios. Duomenų forma erdvėje pasidaro rutulio formos, kuri centruojama koordinačių ašių pradžioje.



Pav. 3. Atsitiktinai sugeneruoti dviejų požymių ir dviejų klasių 100 vektorių duomenys kairėje. Jiems pritaikyta balinimo transformacija – dešinėje. + ir * vaizduoja skirtingas klases, absčių ir ordinačių ašys – skirtingus požymius.

Sekančiame žingsnyje apmokomi parametrizuoti neuroniniai tinklai. Bet prieš tai išskaidoma turima duomenų imtis į testinių ir mokymo duomenų imtis. Atsižvelgiant į praktikoje taikomą duomenų padalijimą, duomenys padalijami santykiu $\frac{1}{2}$. T.y. 50% požymių vektorių skiriami neuroninių tinklų apmokymui, 50% požymių vektorių – testavimui. Duomenų atsitiktiniam išskaidymui naudojama *gendat* funkcija. Ji atsitiktinai išskaido duomenis taip, kad parinktas procentas duomenų imties, apimtų visas klases esančias pradinėje imtyje.

Viena klasifikatorių aibė apmokoma su originaliais duomenimis, kita – su transformuotais balinimo operacijos. Abiejų aibių ekspertų parametrai paliekami tokie patys. Atlikus testavimą gaunami generalizavimo klaidų įverčiai (A priedas). Testavimas atliktas 2×10 kartų. Kiekvieną kartą (taip 10 kartų) duomenys buvo atsitiktinai generuojami. Tačiau mokymosi ir testavimo imčių dydžiai palikti tie patys. Vis tai buvo kartojama du kartus: vieną kartą klasifikatorius dirbo su originaliais imties duomenimis, kitą – su balintais duomenimis.

Kiekvieno eksperto klaidos įverčius nesudėtingai galima gauti PRTools bibliotekos apjungimo mechanizmo pagalba (angl. *mapping*). Komanda užsirašo taip: $D * W * testc$. Čia D – testinė duomenų imtis, kuri pateikiama ekspertui, ar ekspertų aibei – W , o klasifikavimo rezultatas pateikiamas *testc* funkcijai, kuri pagal nutylėjimą grąžina klasifikavimo klaidą (ar jų aibę, jei klasifikavime dalyvavo keli ekspertai).

2 lentelėje atvaizduojami gauti geriausi (viršutinėje eilutės dalyje) ir blogiausi (vidurinėje eilutės dalyje) kiekvieno eksperto atsakymai (generalizavimo klaidos). Apatinėje eilutės dalyje vaizduojamas klasifikavimo klaidų skirtumas tarp geriausio ir blogiausio rezultato. Geriausias ir blogiausias eilutės (bandymo) klasifikavimo rezultatas, bei didžiausias skirtumas tarp klasifikavimo klaidų išskirti pajuodintu šriftu.

Ekspertai		Vienasluksniai perceptronai					Daugiasluksniai tiesioginio sklaidimo NT					Daugiasluksniai grįžtamojo ryšio NT				
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Originalūs duomenys	geriausias	0,0170	0,0170	0,0170	0,0170	0,0170	0,0500	0,0500	0,0500	0,0500	0,0500	0,0167	0,0500	0,0500	0,0667	0,0500
	blogiausias	0,1670	0,1670	0,2000	0,1670	0,2500	0,5000	0,1667	0,1667	0,1833	0,1833	0,5167	0,2000	0,5000	0,1833	0,4833
	skirtumas	0,1500	0,1500	0,1830	0,1500	0,2330	0,4500	0,1167	0,1167	0,1333	0,1333	0,5000	0,1500	0,4500	0,1167	0,4333
Transformuoti i balinimo operacijos ritinys	geriausias	0,0330	0,0330	0,0000	0,0500	0,0330	0,0333	0,0500	0,0333	0,0667	0,0167	0,0500	0,0333	0,0500	0,0167	0,0500
	blogiausias	0,2500	0,2000	0,1830	0,1670	0,1830	0,1667	0,2500	0,1667	0,2500	0,2500	0,5000	0,1667	0,5333	0,2000	0,1833
	skirtumas	0,2170	0,1670	0,1830	0,1170	0,1500	0,1333	0,2000	0,1333	0,1833	0,2333	0,4500	0,1333	0,4833	0,1833	0,1333

Lentelė 2. 3x5 ekspertų geriausi ir prasčiausi klasifikavimo rezultatai tarp balintų ir nebalintų duomenų. Geriausias rezultatas – 0, blogiausias – 1.

Iš gautų rezultatų pastebime, kad geriausią klasifikavimo rezultatą pavyko gauti 3-iam ekspertui. Jo klasifikavimo klaida – 0. Tai reiškia, kad jis iš 50% testinių duomenų, teisingai suklasifikavo juos visus. Toks atvejis pasitaikė tik vienas per visą bandymų seką. Reikia atkreipti dėmesį, kad šis rezultatas buvo pasiektas su ekspertu, kuris buvo apmokytas ant išbalintų duomenų (svarbu paminėti, kad balinimo transformacija analogiškai buvo taikoma ne tik mokymo, bet ir testiniams duomenims, siekiant duomenis turėti toje pačioje transformacijų erdvėje).

Pažvelkime į skirtumus tarp geriausių ir blogiausių klasifikavimo rezultatų. Mažesnis skirtumas šiame bandyme yra antroje lentelės pusėje – su balintais duomenimis. Iš to sprendžiame, kad balintų duomenų ekspertai klysta mažesniame diapazone. Tai nereiškia, kad šie ekspertai yra geresni klasifikatoriai. Šis bandymas neįrodo, kad duomenų transformacija padeda klasifikuoti geriau.

Jeigu panagrinėtume rezultatų lentelę A priede, pastebėtume, kad 3 atlikti bandymai (1, 2 ir 8) geriausius klasifikavimo rezultatus pateikė ekspertai apmokyti su originaliais duomenimis. Dar 3 geriausi rezultatai atsispindi 3,4 ir 7 bandymuose su transformuotais erdvėje duomenimis. Bet tai tik pavienių ekspertų atsakymai. Norint įsitikinti, ar rezultatas bus geresnis apjungus šiuos klasifikatorius ir atlikus analogiškus bandymus, atliekame tiesinę šių ekspertų kombinaciją su skirtingais kombinavimo metodais.

2 bandymas. Į klasifikatorių įtraukiame visus (15) eksperimente dalyvaujančius ekspertus. Kolektyvą tiesiškai kombinuojame 6 skirtingais metodais (gauname 6 klasifikatorius): sandaugos, vidurkio, medianos, maksimumo, minimumo ir balsavimo. Kiekvienas metodas skirtinguose klasifikatoriuose pateikia skirtingus atsakymus.

Abstraktaus lygio ekspertų išvestims t -tojo eksperto atsakymas yra apibrėžiamas formule:

$$d_{t,j} \in \{0,1\}, t = 1, \dots, T; j = 1, \dots, C \quad (9)$$

Kur T – ekspertų skaičius ir C – klasių skaičius. Jei t -tasis ekspertas pasirenka klasę ω_j , tuomet atsakymas $d_{t,j} = 1$, priešingu atveju – $d_{t,j} = 0$. Matavimų lygio ekspertams išvestys apibrėžiamos intervale $d_{t,j} \in [0,1]$. Galutinis kolektyvo atsakymas h_{galut} yra klasė j kuri surenka daugiausiai balsų $\mu_j(x)$, po to kai kombinavimo metodai yra pritaikomi visoms imtyje dalyvaujančioms klasėms:

$$h_{galut}(x) = \max(\mu_1(x), \dots, \mu_1(x)) \quad (10)$$

Tuomet esminiai skirtumai tarp šių metodų yra šie galutiniai surinkti klasių balsų skaičiavimai:

- *Sandaugos* kombinavimo metodui užsirašo formule

$$\mu_j(x) = \prod_{t=1}^T d_{t,j}(x), \quad (11)$$

- *Vidurkio*:

$$\mu_j(x) = \frac{1}{T} \sum_{t=1}^T d_{t,j}(x), \quad (12)$$

- *Medianos*:

$$\mu_j(x) = \underset{t=1, \dots, T}{\text{med}} \{d_{t,j}(x)\}, \quad (13)$$

- *Maksimumo*:

$$\mu_j(x) = \underset{t=1, \dots, T}{\text{max}} \{d_{t,j}(x)\}, \quad (14)$$

- *Minimumo*:

$$\mu_j(x) = \underset{t=1, \dots, T}{\text{min}} \{d_{t,j}(x)\}, \quad (15)$$

Balsavimo metodas veikia abstrakčiame lygyje daugumos balsų principu, o ekspertų atsakymai yra binariniai. Todėl klasifikatorius pateikia klasę J , kaip atsakymą tokiu būdu:

$$\sum_{t=1}^T d_{t,J}(x) = \underset{t=1, \dots, C}{\text{max}} \sum_{t=1}^T d_{t,j}, \quad (16)$$

Atliekame bandymą (apmokome ekspertus ir testuojame) su tomis pačiomis duomenų imtimis, 2x10 kartų. Gauti rezultatai pateikti A priede.

Apibendriname gautus rezultatus (lentelė 3). Klasifikatoriai kuriems buvo taikoma BT, požymių vektorius klasifikavo neblogiau nei klasifikatoriai, kurie dirbo tik su originaliais požymių vektoriais, geriausių bandymų atžvilgiu.

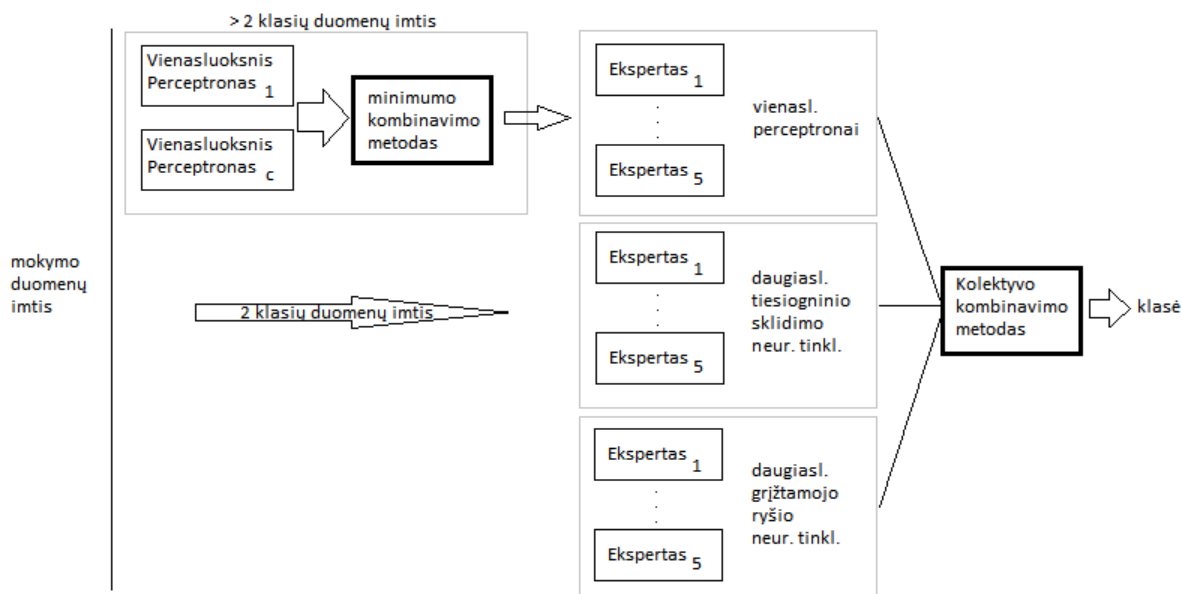
Bandymai		Kombinavimo metodai					
		Sandaugos	Vidurkio	Medianos	Maksimumo	Minimumo	Balsavimo
Originalūs duomenys	geriausias	0,050	0,033	0,033	0,067	0,067	0,033
	blogiausias	0,167	0,167	0,167	0,167	0,167	0,167
	skirtumas	0,117	0,134	0,134	0,100	0,100	0,134
Transformuoti balinimo operacijos duomenys	geriausias	0,000	0,017	0,033	0,000	0,000	0,033
	blogiausias	0,200	0,167	0,167	0,283	0,283	0,167
	skirtumas	0,200	0,150	0,134	0,283	0,283	0,134

Lentelė 3. 2x6 Klasifikatorių geriausi ir prasčiausi klasifikavimo rezultatai tarp balintų ir nebalintų duomenų. Geriausias rezultatas – 0, blogiausias – 1.

Lyginant pavienių ekspertų atsakymus ir bendrus klasifikatorių rezultatus – ekspertų kombinavimas pateikia kur kas geresnius rezultatus. Tačiau žvelgiant iš balinimo transformacijos perspektyvos, akivaizdus skirtumo nepateikia nei kombinuoti, nei pavieniai ekspertai. Atliktuose bandymuose klasifikavimo rezultatus įtakojo mokymosi ir testavimo vektorių parinkimas. Tokių bandymų baigtį galėjo įtakoti ir maža duomenų imtis. Todėl sekantis šio eksperimento bandymas su skirtingomis duomenų imtimis leis įvertinti BT įtaką kolektyvo kokybei platesniu aspektu.

3 bandymas. Šiam bandymui atlikti naudojamos skirtingos duomenų imtys. Vienos iš jų pseudo sintetinės, kitos – imtos iš realaus pasaulio problemų. Naudojamos duomenų imtys konkrečiau aprašytos 1.4 poskyryje. Šiame bandyme visi parametrai paliekami tokie patys, kaip ir buvo naudoti 2 bandyme. Taip pat taikome tuos pačius kombinavimo metodus visiems 15 ekspertų. Kombinuojame ekspertus tik vieną kartą su kiekviena skirtinga duomenų imtimi po du kartus. Pirmą kartą su originaliais duomenimis, antrą – jiems pritaikius BT.

Kadangi šiame bandyme pasitaiko duomenų imčių su daugiau nei dvejomis klasėmis, todėl kombinuojant ekspertus, reikėtų atkreipti dėmesį į vienasluoksnių perceptronų panaudojimą. Jie apsiraso *perlc()* komanda. Esant dviejų klasių problemai, ši komanda grąžins tiesinį ekspertą. Tačiau kai duomenų imtyje dominuoja n klasių, atliekama vidinė ekspertų, apmokytų su kiekviena iš n klasių, kombinacija. Kombinavimo metodas PRTools bibliotekoje pagal nutylėjimą yra minimumo – *minc*. Šis metodas grąžina klasę, kuri tarp vidinių ekspertų atsakymų pasitaikė mažiausiai kartų. Todėl daugelio klasių situacijoje, konstruojant klasifikatorių šiame bandyme, gauname mišrią tinklo topologiją (Pav. 4). Vienasluoksnių perceptrono ekspertas gaunamas kaip klasifikatoriaus, kombinuoto su minimumo kombinavimo taisykle, atsakymas. Vėliau visi ekspertai, kuriems pritaikius vieną iš naudojamų kombinavimo metodų, vėl apjungiami į klasifikatorių, kuris pateikia klasę kaip atsakymą.



Pav. 4. 3 bandyje naudojama klasifikatoriaus topologija su vidine kombinavimo strategija. Iš kairės į dešinę vaizduojama mokymosi ir klasifikavimo seka.

Bandymas atliekamas su 6 skirtingais kombinavimo metodais, kaip ir 2 bandyje. Atliekama 2x10x10 bandymų. T.y kiekvienas klasifikatorius apmokomas ir testuojamas 10 kartų, su 10 skirtingų duomenų imčių po du kartus (su balintais ir originaliais duomenimis). Gauti rezultatai pateikiami A priede. Gauti klasifikatorių, su skirtingais kombinavimo metodais ir duomenų imtimis, klasifikavimo klaidų vidurkiai atvaizduoti 4 lentelėje.

Lentelėje pateikiama 10 skilčių su duomenų imčių pavadinimais. Kiekvienoje skilties viršutinėje eilutės dalyje pateikiami skirtingų kombinavimo metodų klasifikavimo klaidų vidurkiai – dirbant su originaliais duomenimis, apatinėje eilutės dalyje – duomenims atlikus balinimo transformaciją. Pajuodintu šriftu pabrėžiami gauti geriausi rezultatai.

Duomenų imtys	Kombinavimo metodai		Sandaugos	Vidurkio	Medianos	Maksimumo	Minimumo	Balsavimo
Gendatb			0,139	0,134	0,148	0,147	0,147	0,148
			0,129	0,136	0,153	0,149	0,150	0,153
Gendatc			0,199	0,200	0,251	0,214	0,214	0,251
			0,209	0,209	0,237	0,206	0,206	0,237
Gendatd			0,073	0,071	0,043	0,097	0,097	0,043
			0,061	0,056	0,043	0,074	0,073	0,043
Gendats			0,159	0,162	0,163	0,168	0,168	0,163
			0,165	0,164	0,163	0,161	0,161	0,163
Iris			0,315	0,024	0,027	0,571	0,586	0,026
			0,279	0,024	0,026	0,380	0,509	0,047
Breast			0,048	0,049	0,049	0,048	0,048	0,049
			0,045	0,046	0,048	0,063	0,063	0,048
Wine			0,213	0,007	0,010	0,404	0,353	0,015
			0,109	0,027	0,022	0,115	0,152	0,027
Satellite			0,637	0,251	0,230	0,811	0,770	0,179
			0,561	0,186	0,191	0,793	0,811	0,200
Diabetes			0,282	0,268	0,269	0,290	0,290	0,269
			0,283	0,267	0,268	0,287	0,287	0,268
Ecoli			0,589	0,267	0,179	0,577	0,766	0,158
			0,451	0,131	0,146	0,682	0,798	0,144

Lentelė 4. Gauti generalizavimo klaidų vidurkiai atlikus 2x10x10 bandymų su skirtingomis duomenų imtimis. Geriausias rezultatas – 0, blogiausias – 1.

Analizuojant bandymo rezultatus pastebime, kad BT taikymas ne visada duoda geresnius klasifikavimo rezultatus. Klasifikuojant pseudo sintetinius ir realaus pasaulio duomenis, BT taikymas geresnius rezultatus davė tik 4 duomenų imtims – Gendatb, Breast, Diabetes ir Ecoli. Kitos duomenų imtys buvo geriau suklasifikuotos su originalių duomenų klasifikatoriais.

Žvelgiant iš skirtingų kombinavimo metodų perspektyvos – taip pat nėra tokio metodo, kuris būtų geriausias. Nors klasifikatorių topologija išliko tokia pati visiems kombinavimo metodams, rezultatai gavosi skirtingai. Tai įtakojo skirtingos duomenų imtys, skirtingi klasių, požymių, mokymosi ir testavimo duomenų kiekiai.

Žvelgiant bendrai, geriausiai pasirodė vidurkio kombinavimo metodas. Jo klasifikatoriai tiksliausiai suklasifikavo 7 duomenų imtis. Tačiau reikia atkreipti dėmesį, čia parinktas geriausias rezultatas tarp balintų ir nebalintų duomenų.

Išvados. Eksperimente buvo nagrinėjama duomenų balinimo transformacijos įtaka klasifikatoriaus kokybei pavieniams neuroniniams tinklams ir jų kolektyvams su skirtingais tiesiniais kombinavimo metodais. Atlikus eksperimentą paaiškėjo, kad atskiri ekspertai apmokyti su balintais duomenimis ne visada yra geresni klasifikatoriai. Dvejopai elgiasi ir ekspertų kolektyvai – vieni tampa kokybiškesni ant išbalintų duomenų, kiti – ant originalių. Klasifikatorių atsakymams nemažą įtaką turi kombinavimo metodo parinkimas. Šiame eksperimente duomenų balinimo transformacija nebuvo vienareikšmiškai geras sprendimas visoms duomenų imtims. Pastebėjau, jog nėra tiesioginio santykio tarp duomenų balinimo operacijos ir klasifikatoriaus kokybės.

Tai galėjo lemti klasifikatorių panašumas. Iš to kyla *antrasis eksperimentas*, nagrinėsiantis BT poveikį klasifikatorių diversiškumui.

5.5.2 Balinimo transformacijos poveikis klasifikatorių diversiškumui

Siekiamas tikslas. Diversiški klasifikatoriai neretai duoda geresnius rezultatus [KW02]. Šiuo eksperimentu siekiama ištirti ar klasifikatoriai išmokyti ant balintų duomenų yra vienodesni ar skirtingesni. Matavimams atlikti pasitelkiami skirtingi metodai. Ar skirtingi diversiškumo matavimo metodai pateikia skirtingus atsakymus? Kokią įtaką rezultatams duoda sekančios ekspertų kombinacijos:

- Geresnės kokybės, tačiau vienodesnių ekspertų kombinacijos;
- Prastesnės kokybės, tačiau skirtingesnių ekspertų kombinacijos.

Siekiant atsakyti į šiuos klausimus, pirmuoju bandymu skirtingais metodais įvertinu kaip BT įtakoja ekspertų diversiškumą, bei šių metodų panašumą. Sekančiame bandyme palyginu geros kokybės vienodų ir blogos kokybės skirtingų klasifikatorių atsakymus. Pateikiu pastebėjimus ir išvadas.

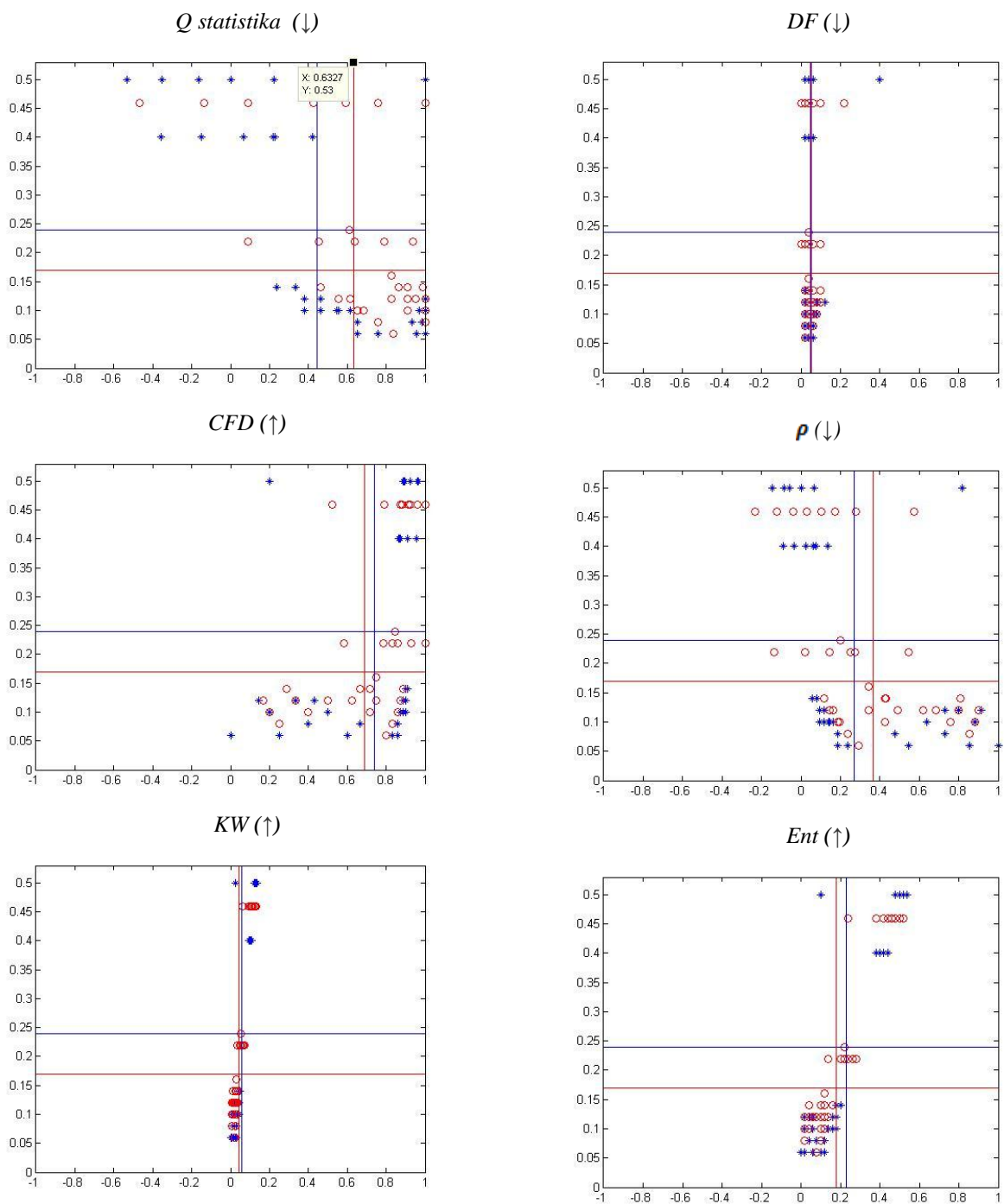
Praktinė dalis. 1 bandymas. Pirmajam bandymui atlikti pasinaudoju *gendatd* funkcija. T.y atsitiktinai sugeneruoju 100 požymių vektorių su dvejomis klasėmis (su tikimybe 1/2). Padalinu duomenų imtį į 50 testinių ir 50 mokymo vektorių. Šiam bandymui atlikti panaudoju 10 ekspertų. Visi parinkti ekspertai yra tiesiniai perceptronai. Kiekvienam iš jų parenku skirtingas parametrų kombinacijas, siekiant ekspertus padaryti diversiškus. Parenkami skirtingi:

- pradiniai svoriai;
- mokymosi iteracijų skaičiai (intervale [1;10000]);
- mokymosi žingsniai (intervale [0,01;0,9]);
- mokymosi duomenų imties dydžiai (nuo 10 iki 50% iš mokymosi duomenų imties).

Siekiant įvertinti ekspertų diversiškumą, į bandymą įtraukiu 6 diversiškumo matavimo metodus: Q statistikos (*Q*), dvigubos klaidos (*DF*), atsitiktinės klaidos (*CFD*), koreliacijos koeficiento (ρ), Kohavi-Wolpert dispersijos (*kw*) ir entropijos (*Ent*). *Q*, *DF* ir ρ yra poriniai skirtingumo matavimai, likę – neporiniai.

Siekiant įvertinti BT poveikį klasifikatorių skirtingumui, lygiagrečiai originaliems duomenims tuos pačius žingsnius atlieku ir balintiems imties duomenims. T.y bandymo pradžioje pasigaminu dvi duomenų imtis: *A* – originalią, *A_w* – balintą (*A* imčiai atlikus balinimo transformaciją). Vienodai *A* ir *A_w* padalinu į mokymo ir testines imtis. Sudaryti originalių ir balintų duomenų ekspertų

poroms naudoju tuos pačius parametrus. Sekančiame žingsnyje visus ekspertus apmokau. Iš viso gaunu 2×10 ekspertų. Diversiškumo matavimus atlieku ekspertų poroms. Kadangi turime 2×10 ekspertų, gauname $2 \times 6 \times 45$ diversiškumo matavimų atsakymus. T.y iš 10 ekspertų, gauname $10 \times 9/2 = 45$ matavimus. Šiems matavimams atlikti naudoju 6 skirtingus matavimo metodus. Tuos pačius matavimus atlieku antrą kartą su A_w imtimi. Gautų ekspertų porų diversiškumo atsakymų ir klasifikavimo klaidų išsibarstymą atvaizduoju pav. 5. Ekspertų porų klasifikavimo kokybę ir jų panašumą, matuojamą skirtingais metodais, atspindi 6 grafikai. Mėlyna spalva vaizduojami rezultatai gauti klasifikuojant originalius duomenis, raudona spalva – rezultatai gauti klasifikuojant balintus duomenis. Vertikalios raudonos ir mėlynos tiesės žymi individualaus eksperto klasifikavimo klaidos vidurkį, horizontalios tiesės – diversiškumo matavimų tarp ekspertų porų vidurkį, atitinkamai ant išbalintų ir originalių duomenų.



Pav. 5. Ekspertų porų diversiškumo ir klasifikavimo klaidos santykis su originaliais duomenimis (mėlyna spalva) ir atlikus BT (raudona spalva). Skirtingumo matavimo metodai *Q, DF, CFD, ρ, kw, Ent*. Vertikalios tiesės – vidutinis individualaus eksperto tikslumas, horizontalios – diversiškumo matavimų vidurkis (ant originaliųjų ir balintų duomenų). X ašyje vaizduojamas diversiškumo matavimas (kintanti reikšmė pagal skirtingų metodų matavimus), Y ašyje – klasifikavimo klaida (1 – blogai, 0 – gerai).

Virš kiekvieno grafiko nurodytas metodas, kuriuo matuojamas skirtingumas. Šalia jo vaizduojamos rodyklės rodo, jog ekspertų poros tuo skirtingesnės, kuo matavimų rezultatai yra mažesni (↓) arba didesni (↑), priklausomai nuo diversiškumo matavimo metodo. *Q* statistikos ir

koreliacijos koeficiento ρ diversiškumo matavimo metodų atsakymai $\in [-1, 1]$, kitų $\in [0, 1]$. Jie vaizduojami abscisių ašyje. Klasifikatorių porų klasifikavimo klaidos $\in [0, 1]$, kurios vaizduojamos ordinačių ašyje.

Apžvelgsime tik vieną bandymą. Atlikus šį bandymą 20 kartų, beveik visais atvejais gauti panašūs rezultatai, o labiausiai varijavo klasifikavimo klaidos vidurkis intervale $[0,02; 0,4]$. Tačiau šiame bandyme svarbiausia atkreipti dėmesį į diversiškumo matavimus. Q statistikos matavimas dažnai pateikia didžiausią diversiškumo matavimų vidurkių skirtumą tarp išbalintų ir originalių duomenų. Šiame bandyme skirtumas yra 0,19. Mažiausią skirtumą pateikė DF metodas – 0,005.

Reikia pastebėti, kad visi matavimo metodai vienodai įvertino diversiškumą tarp balintų ir originalių duomenų. Nepaisant to balintų duomenų ekspertai yra panašesni už originalių duomenų ekspertus. Tai kartojosi per visus 18 iš 20 bandymų. Q, DF ir ρ metodų, balintų duomenų diversiškumo matavimų vidurkis yra arčiau 1, nei originalių duomenų. Kai šių metodų atsakymas yra 1 – laikoma, kad klasifikatoriai yra statistiškai vienodi. Skirtingumas tarp jų auga matavimams mažėjant. Statistiškai nepriklausomi klasifikatoriai būna tada, kai atsakymas yra 0. Q ir ρ metodai gali pateikti ir neigiamus atsakymus. Tai reikštų jog ekspertai daro klaidas su skirtingais požymių vektoriais. Kai atsakymas teigiamas – abu ekspertai vienodai klasifikuoja požymių vektorius. CFD, kw ir Ent metodai elgiasi priešingai – kuo atsakymas arčiau 1, tuo ekspertai yra skirtingesni.

Iš 5 pav. matome, jog skirtingi metodai diversiškumą įvertina nevienodai. Pavyzdžiui Q statistikos matavimai yra išsibarstę visame intervale tarp -1 ir 1. Toks vaizdas leidžia tiksliau vertinti ekspertų skirtingumą. Iš Q statistikos išsibarstymo grafiko matome, kad vienos klasifikatorių poros skirtingai pateikė atsakymus, kitos – vienodai. Tą patį galima pasakyti ir apie ρ metodą. Taip pat šie metodai, pateikė nemažai skirtingų ekspertų porų. T.y pasitaikė nemažai šių metodų atsakymų, kurie yra arti 0. Žvelgiant į kitus metodus (pvz. Ent, kw), diversiškumo matavimų išsibarstymas yra arti 0. Tai reiškia jog ekspertai laikomi panašiais. Pažvelgus bendrai Q, CFD ir ρ yra panašūs metodai. Taip pat ir DF, kw ir Ent . Tačiau tarp šių dviejų metodų grupių yra skirtumas išsibarstymo atžvilgiu.

Iš 5 pav. galime įvertinti ir diversiškumo įtaką klasifikavimo kokybei. Geriausiai šis santykis matosi su Q, CFD ir ρ metodais. Pažvelgus į šiuos grafikus matome, kad skirtingesni klasifikatoriai pateikė blogesnę klasifikavimo klaidos vidurkį už individualaus eksperto klasifikavimo klaidos vidurkį arba elgėsi dviprasmiškai (pvz. CFD grafike skirtingi klasifikatoriai išsibarstę apylygiai tiek

į viršų, tiek į apačią, nuo vidutinės klasifikavimo klaidos – horizontali linija). Ir priešingai, kuo ekspertai buvo panašesni, tuo geresnį klasifikavimo rezultatą pateikė.

Visuose matavimo metoduose vyrauja ta pati tendencija – lengviausia yra gauti geresnės kokybės vienodesnius ekspertus ir blogesnės kokybės skirtingesnius ekspertus. T.y žvelgiant į bet kurį pav. 5 grafiką, skirtingumo matavimai linkę išsibarstyti viena įstriža tiese. Padalijus grafiką į ketvirčius, dažniausiai užsipildo tik du įstriži ketvirčiai. Kurie ketvirčiai užsipildo, priklauso nuo matavimo metodo ir skirtingumą nurodančios krypties (↓).

Tokiam rezultatui galėjo įtakoti bandymo metu parinktas balsavimo kombinavimo metodas. naudojami tik tiesiniai perceptronai, bei vieno tipo pseudo sintetinė duomenų imtis. Siekiant tikslumo ir abstraktesnio rezultato atliksime antrą bandymą.

2 bandymas. Šiame bandyme į ekspertų aibę įtraukiu 10 grįžtamojo ryšio ir 10 tiesioginio sklidimo neuroninių tinklų. Šiems ekspertams sudaryti, parenku:

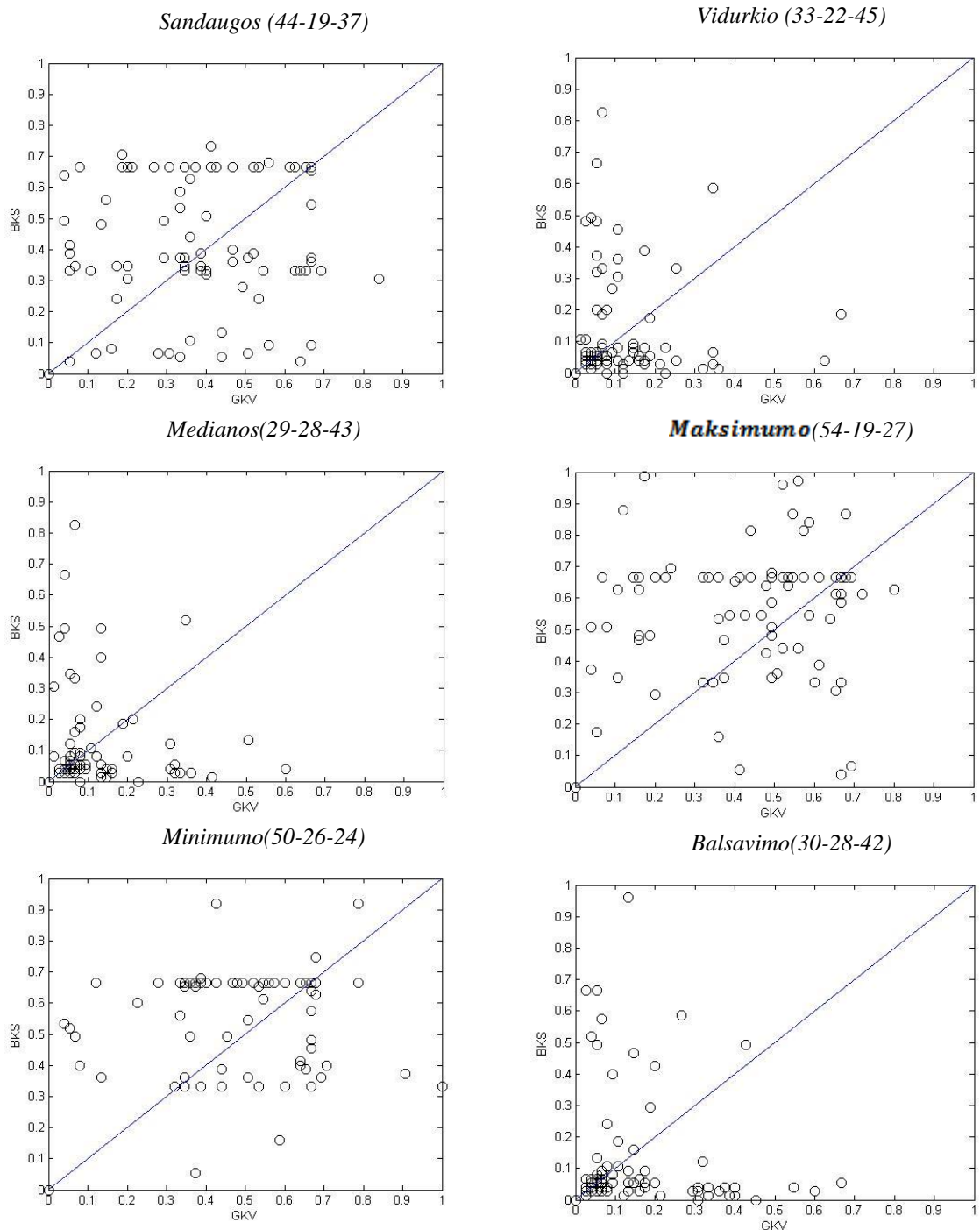
- skirtingus paslėptų sluoksnių skaičius (intervale [2, 10]);
- skirtingus pradinis svorius;
- mokymosi iteracijų skaičius (intervale [1,10000]);
- mokymosi duomenų imties dydžius (nuo 10 iki 50% iš mokymosi duomenų imties).

Sekančiu žingsniu apmokau neuroninius tinklus ir gaunu 20 ekspertų aibę. 1 bandymo pavyzdžiu, ekspertų poromis sukuriu 190 klasifikatorių ir kiekvienam iš jų įvertinu diversiškumą. Tam naudoju *CDF* metodą. Toliau atrenku:

- a) geresnės kokybės (kurių klasifikavimo klaidos vidurkis yra mažesnis už individualaus eksperto klaidos vidurkį) vienodesnius (ant balintų duomenų, kurių diversiškumo matavimai yra intervale [0; 0,5]) klasifikatorius.
- b) prastesnės kokybės (kurių klasifikavimo klaidos vidurkis yra didesnis už individualaus eksperto klaidos vidurkį) skirtingesnius (ant originalių duomenų kurių diversiškumo matavimai yra intervale (0,5; 1]) klasifikatorius.

Iš atrinktų ekspertų grupių suformuoju 2 klasifikatorius. Vieną iš *a* grupės ekspertų (jį vadinu GKV), kitą iš *b* grupės ekspertų (BKS). Bandymą atlieku su skirtingais ekspertų kombinavimo metodais (sandaugos, vidurkio, medianos, maksimumo, minimumo ir balsavimo). Šiam bandymui naudoju *Iris* duomenų imtį, kurią padalinu į mokymo ir testinę imtis santykiu ½. Bandymą pakartoju 100 kartų. Gauti rezultatai pateikti pav. 6. Skliausteliuose prie kombinavimo metodo pateikiami klasifikatorių atsakymų dažnumai iš 100 bandymų, eiliškumu nuo kairės į dešinę: BKS > GKV; BKS = GKV; BKS < GKV;

Iš pateiktų rezultatų matome kad abiejų grupių (GKV, BKS) klasifikatoriai elgiasi skirtingai, priklausomai nuo pasirinkto kombinavimo metodo. BKS klasifikatorius pranašiau elgėsi su sandaugos, maksimumo ir minimumo kombinavimo taisyklėmis. T.y jis buvo pranašesnis už GKV klasifikatorių. Su maksimumo kombinavimo metodu iš 100 bandymų, BKS kombinavo pranašiau daugiausiai kartų – 54.



Pav. 6. Blogos kokybės skirtingų (BKS) ekspertų ir geros kokybės vienu (GKV) ekspertų klasifikatorių klasifikavimo klaidų palyginimas, ekspertus kombinuojant skirtingais metodais. Skliausteliuose parašyta kiek kartų (iš 100 bandymų) BKS klasifikatorius (kairėje pusėje) geriau klasifikavo už GKV klasifikatorių (dešinėje pusėje). Vidurinė reikšmė – abu klasifikavo vienodai. Grafikų absčių ir ordinačių ašyse vaizduojamos klasifikavimo klaidos. 0 – geriausias rezultatas, 1 – blogiausias.

GKV klasifikatorius elgėsi priešingai. Jis buvo pranašesnis su vidurkio, medianos ir balsavimo kombinavimo metodais. Atkreipkime dėmesį į abiejų klasifikatorių atsakymų išsibarstymą erdvėje.

Pastebime, kad sandaugos, maksimumo ir minimumo metodais kombinuotų klasifikatorių atsakymai yra išsibarstę visoje erdvėje, intervale [0, 1].

Galime daryti išvadą, kad šie metodai nėra tikslūs sudarant klasifikatorius būtent šiai duomenų imčiai. Tačiau jei vis tik naudotume šiuos metodus, tiksliau čia elgtųsi BKS klasifikatorius. Pažvelgus į sekančių kombinavimo metodų grupę – atsakymų išsibarstymas yra kur kas siauresnis. Šiais metodais geriau kombinuojasi geresnės kokybės panašesni klasifikatoriai. Nepaisant kraštinių klasifikatorių atsakymų, atlikus bandymą 100 kartų su 6 skirtingais kombinavimo metodais, apie 20% BKS ir GKV klasifikatorių atsakymų sutapo.

Išvados. Atlikus pirmąjį bandymą, kurio tikslas – ištirti balinimo transformacijos poveikį klasifikatorių diversiškumui, priėjau išvados, kad BT klasifikatorius padaro panašesnius. Bandymas parodė, kad skirtingi diversiškumo matavimo metodai dažniausiai pateikia panašius rezultatus. Labiausiai bandymuose išsiskyrė atsitiktinės klaidos (*CFD*) metodas, kuris parodė didžiausią santykį tarp klasifikatorių skirtingumo ir klasifikavimo tikslumo.

Sekančiame bandyme matavimai buvo atlikti *CFD* metodu. Šio bandymo tikslas – sudarius blogos kokybės skirtingų ir geros kokybės panašių ekspertų klasifikatorius su skirtingais kombinavimo metodais, ištirti santykį tarp jų klasifikavimo klaidų. Bandyme geriausią rezultatą pateikė geros kokybės panašių ekspertų klasifikatorius su vidurkio kombinavimo metodu.

Buvo pastebėta, kad kombinavimo metodai išsiskiria į grupes. Bandymuose su *Iris* duomenų imtimi vidurkio, medianos ir balsavimo metodai buvo daug tikslesni už likusius tris. BKS klasifikatoriai buvo pranašesni su prastesniais kombinavimo metodais, GKV – su geresniais kombinavimo metodais.

Likusios ekspertų grupės – tai geros kokybės skirtingi ir blogos kokybės panašūs ekspertai. Šios ekspertų grupės bandymuose susidarydavo arba labai mažos, arba į jas ekspertai net nepakliūdavo. Taip pat kombinuojant tokias ekspertų grupes, klasifikatoriai geresnių rezultatų nedavė, nei kad BKS ar GVK grupių klasifikatoriai. Todėl siekiant supaprastinti skaičiavimus, BKV ir GKS ekspertų grupių į tolimesnius eksperimentus neįtrauksiu.

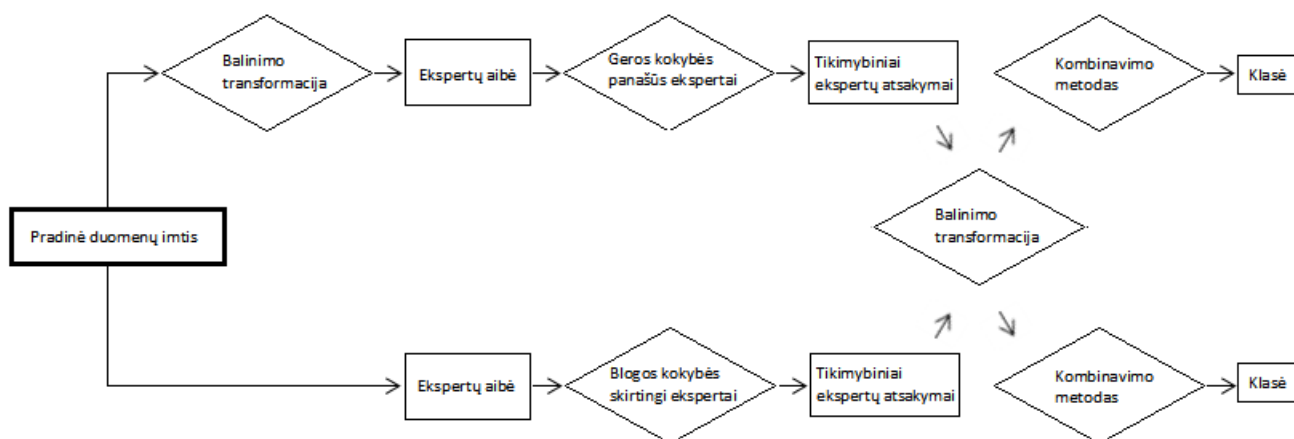
5.5.3 Balinimo transformacijos taikymo ekspertų atsakymams poveikis klasifikatoriaus kokybei

Balinimo transformacijos taikymas duomenims klasifikatorius padaro panašesnius. Jų kombinavimas atitinkamais metodais leidžia gauti geresnius rezultatus, nei originalių duomenų klasifikatorių kombinavimas (tai aprašyta 1.5.2 poskyryje). Tačiau jei BT taikytume ne tik duomenims, bet ir klasifikatorių atsakymams. Kaip tai įtakotų galutinį rezultatą? Jei įtrauktume

tiesinį perceptroną į pagrindinio ekspertų kombinuotojo vietą, kaip papildomą kombinavimo metodą, jis turėtų mokytis geriau su balintais ekspertų atsakymais. Ar tokio klasifikatoriaus atsakymai būtų pranašesni? Iš čia kyla *trečiasis eksperimentas*.

Siekiamas tikslas. Šiuo eksperimentu siekiama išsiaiškinti, kaip klasifikatorių atsakymų balinimas įtakoja galutinį klasifikavimo rezultatą. Šiame bandyme remiamės ir antrame eksperimente gautais rezultatais. T.y. atrenkame geros kokybės panašius ir prastesnės kokybės skirtingus ekspertus remiantis *CFD* metodu. Su skirtingais kombinavimo metodais tokie ekspertai kombinuojasi gana kokybiškai. Todėl siekiant geresnių klasifikavimo rezultatų, darome prielaidą, kad tokių ekspertų tolydinių atsakymų balinimas turėtų duoti geresnius rezultatus. Šiuo bandymu siekiama išsiaiškinti ar papildomas ekspertų atsakymų balinimas sumažina klasifikavimo klaidą. Ar skirtingų klasifikatorių atsakymų balinimas yra pranašesnis už panašių klasifikatorių atsakymų balinimą.

Praktinė dalis. 1 bandymas. Šis bandymas padalintas į 2 lygiagrečias dalis (pav. 7).



Pav. 7. Blogos kokybės skirtingų ir geros kokybės panašių ekspertų atsakymų balinimo poveikis klasifikatorių atsakymams. Schema vaizduoja neuroninio tinklo architektūrą ir jo mokymo seką.

Pirmoje dalyje konstruojamas neuroninis tinklas, kuris pirmiausiai apmokomas su balintais duomenimis, o vėliau gauti tikimybiniai ekspertų atsakymai dar kartą transformuojami balinimo operacijos. *CFD* metodu atrenkami geros kokybės panašūs ekspertai. Į kombinavimo metodų aibę įtraukiamas tiesinis perceptronas. Vėliau ekspertų atsakymai kombinuojami tiesinio perceptrono pagalba.

Dviejų žingsnių duomenų ir atsakymų balinimas reikalauja nemažai papildomo laiko. Siekiant įsitikinti, ar tokia neuroninio tinklo architektūra pasiteisina klasifikavimo klaidos atžvilgiu, lygiagrečiai atliekame antrą bandymo dalį. Antroje dalyje BT taikome tik ekspertų atsakymams, bet

ne pradiniais duomenimis. Taip *CFD* metodu atrenkami blogos kokybės skirtingi ekspertai. Vėliau kaip ir pirmajame bandyme taikome tiesinį perceptroną, kaip kombinavimo metodą.

Siekiant atrinkti blogos kokybės skirtingus ir geros kokybės panašius ekspertus iš mokymo duomenų imties pasiimame 50% vektorių. Šiuos vektorius toliau vadinsiu validavimo duomenimis. Validavimo duomenis paduodu ekspertams ir iš jų gautų binarinių atsakymų, *CFD* metodu atrenku BKS ir GKV ekspertų aibes. Abejose bandymo dalyse šių ekspertų atsakymų aibes sujungiu. Gaunu GKV ir BKS ekspertų atsakymų aibes. Atsakymai yra tikimybinės klasių reikšmės. Sekančiame žingsnyje šioms aibėms taikau BT. Prie anksčiau nagrinėtų kombinavimo metodų įtraukiu ir tiesinį perceptroną, kaip kombinatorių. Mokymo etape jį apmokau iš validavimo metu gautų ekspertų atsakymų. Tokiu būdu gaunu 2 galutinius klasifikatorius (GKV ir BKS).

Bandymą atlieku su Iris duomenų imtimi. Ją padalinu į mokymo ir testinę duomenų imtis santykiu $\frac{1}{2}$. Į ekspertų aibę įtraukiu 10 grįžtamojo ryšio ir 10 tiesioginio sklidimo neuroninių tinklų. Siekiant diversiškumo matavimų išsibarstymo, ekspertams parenku skirtingus parametrus:

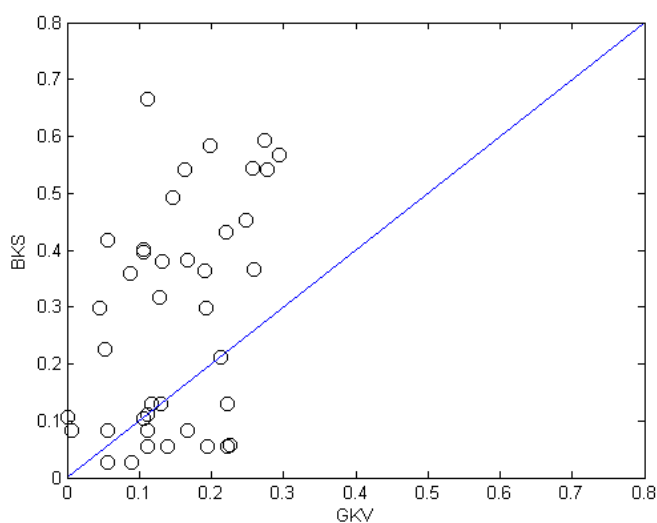
- paslėptų sluoksnių skaičius (intervale [1, 5]);
- skirtingus pradinis svorius;
- mokymosi iteracijų skaičius (intervale [1, 1000]);
- mokymosi duomenų imties dydžius (nuo 10 iki 50% iš mokymosi duomenų imties).

Sekančiame žingsnyje apmokau ekspertus su mokymo duomenimis. Tokią pat ekspertų aibę su tokiais pačiais parametrais apmokau su balintais duomenimis. BT padeda gauti panašesnius klasifikatorius (aprašyta 1.5.2 poskyryje).

Pasinaudodamas *CFD* diversiškumo matavimo metodu ir validavimui skirtais duomenimis iš mokymo duomenų aibės, atrenku GKV ir BKS ekspertus. BKS ekspertus atrenku iš ekspertų aibės, kuri buvo apmokyta su originaliais duomenimis, GKV – iš balintais duomenimis mokytojų ekspertų aibės.

Sekančiame žingsnyje apmokau du tiesinius perceptronus. Abu perceptronai mokomi ant BKS ir GKV ekspertų atsakymų aibių. Šios atsakymų aibės – tai tikimybiniai klasių įverčiai, gauti diversiškumo matavimo žingsnyje, klasifikuojant validavimo duomenis. Gauti tiesinio perceptrono ekspertai naudojami kaip pagrindiniai klasifikatorių kombinavimo metodai.

Bandymą kartoju 50 kartų su Iris duomenų imtimi, parenkant skirtingus diversiškumą įtakojančius parametrus. Gauti rezultatai pavaizduoti pav. 8.



Pav. 8. Blogos kokybės skirtingų (BKS) ir geros kokybės panašių (GKV) klasifikatorių palyginimas, kai abiemis klasifikatoriams taikomas ekspertų atsakymų balinimas. Geriau klasifikavo BKS – 28, blogiau – GKV – 13. Abscisių ašyje vaizduojami GKV, ordinačių ašyje - BKS klasifikatorių klasifikavimo klaidų rezultatai po 50 bandymų. Geriausias rezultatas – 0, blogiausias – 1;

Iš grafiko matome, kad BKS klasifikatorius klysta kur kas mažesniame klaidos intervale $[0,01; 0,3]$, nei GKV klasifikatorius $[0,02; 0,68]$. Taip pat BKS klasifikatorius, klasifikavo geriau 28 kartus. Tai beveik 2 kartus geriau už GKV klasifikatorių, kuris geriau klasifikavo 13 kartų. 9 iš 50 bandymų, abu klasifikatoriai pateikė vienodas klasifikavimo klaidas.

Šis bandymas parodė, kad ekspertų atsakymų balinimas BKS klasifikatorių padaro du kartus geresnį už GKV klasifikatorių. Papildomas ekspertų atsakymų balinimas ir jų klasifikavimas tiesinio perceptrono pagalba yra geresnis būdas už kitus kombinavimo metodus (sandaugos, vidurkio, medianos, maksimumo, minimumo ir balsavimo kombinavimo metodus) klasifikavimo klaidos diapazono atžvilgiu.

Šio eksperimento rezultatus galėjo įtakoti specifinė Iris duomenų imtis, todėl siekiant pažvelgti į kitų duomenų imčių kuriamus rezultatus bandymas buvo pakartotas su Ecoli ir gendatb duomenų imtimis. Gauti rezultatai labai neišsiskyrė pirmojo bandymo atžvilgiu ir BKS klasifikatorius čia taip pat buvo pranašesnis beveik kelis kartus už GKV klasifikatorių. Tačiau klaidos diapazonas buvo išaugęs iki $[0,02; 0,4]$, geriausio klasifikatoriaus atžvilgiu.

Išvados. Šio eksperimento metu, tiriant ekspertų tikimybinių atsakymų balinimo poveikį GKV ir BKS klasifikatorių kokybei pastebėjau, kad BKS klasifikatorius klysta mažesniame intervale. Toks neuroninis tinklas yra apmokomas su originaliais duomenimis, o BT taikoma tik ekspertų atsakymams. Tokiu būdu gaunamas gana mažas generalizavimo klaidos vidurkis.

Darbo rezultatai

1. Atliktas praktinis tyrimas vadovaujantis darbo planu, išskeltomis hipotezėmis ir nuosekliai gautais bandymų rezultatais.
2. Ištirta klasifikatorių apmokytų ant išbalintų duomenų ir klasifikatorių apmokytų ant originalių duomenų kokybė su 10 skirtingų duomenų imčių.
3. Palyginti 6 skirtingi klasifikatorių diversiškumo matavimo metodai.
4. Įvestas metodas skirtingos diversijos ir kokybės ekspertams grupuoti.
5. Ištirta duomenų balinimo transformacijos įtaka ekspertų diversiškumui ir jų klasifikatorių kokybei.
6. Ištirtas ekspertų atsakymų balinimo poveikis geros kokybės panašių ir blogos kokybės skirtingų klasifikatorių atsakymams.

Išvados

1. Klasifikatoriai sudaryti iš ekspertų apmokytų su balintais duomenimis ne visada yra geresni. Ekspertų kolektyvai elgiasi dvejopai – vieni tampa kokybiškesni ant išbalintų duomenų, kiti – ant originalių. Tokių klasifikatorių sudarymui nemažą įtaką turi tinkamo kombinavimo metodo parinkimas.
2. Skirtingi diversiškumo matavimo metodai pateikia panašius rezultatus. Tyrime labiau išsiskyrė atsitiktinės klaidos metodas (*CFD*), kuris daugelyje bandymų parodė didesnę santykį tarp klasifikatorių skirtingumo ir klasifikavimo tikslumo, nei kiti metodai. Jį verta naudoti siekiant diversiškų skirtingos kokybės ekspertų kolektyvams formuoti.
3. Balinimo transformacijos taikymas pradiniam duomenims, su jais apmokomus ekspertus padaro vienodesnius (panašesnius), remiantis atsitiktinės klaidos metodu.
4. Iš ekspertų padalintų į grupes, remiantis diversiškumo ir kokybės matavimais, dažnai konstruojasi geresni klasifikatoriai. Geresnį nei vidutinį rezultatą tyrimuose dažniau pateikė geros kokybės vienodų ir blogos kokybės skirtingų ekspertų klasifikatoriai. Kitos klasifikatorių poros dažniau klasifikavo blogiau.
5. Duomenų balinimo transformacijos taikymas ekspertų aibės atsakymams blogos kokybės skirtingus klasifikatorius padaro pranašesnius už geros kokybės vienodus klasifikatorius net iki 40%.

Šaltinių sąrašas

- [BKM99] C. Blake, E. Keogh, and C.J. Merz. UCI repository of Machine Learning databases, 1999. <http://archive.ics.uci.edu/ml/index.html>
- [Bro04] G. Brown. Diversity in Neural Network Ensembles. The university of Birmingham, 2004
- [CHH90] J. J. Hull, A. Commike, T. K. Ho, Multiple algorithms for handwritten character recognition. Frontier In Handwriting Recognition, p.p. 117-130, Canada, Montreal, 1990
- [DHS97] R. O. Duda, P. E. Hart, D. G. Stork. Pattern Classification, Ricoh California Research Center, USA, pp. 635, 1997.
- [DSZ09] P. Du, H. Sun, W. Zhang. Multiple Classifier Combination For Target Identification From High Resolution Remote Sensing Image. Departament of Remote Sensing and Geographical Information Science, China University of Mining and Technology, 2009.
- [DT00] R. P. W. Duin, D. M. J. Tax. Experiments with classifier combining rules. In Proc. Int. Workshop on Multiple Classier Systems, Calgiari, Italy, pp. 16, 2000.
- [DUI93] R. P. W. Duin. Delft Pattern Recognition Group, Faculty of Applied Physics of Delft University of Technology, 1993. <http://prtools.org/index.html>
- [GR01] G. Giacinto, F. Roli. An Approach to the Automatic Design of Multiple Classifier Systems, Pattern Recognition Letters, vol. 22, pp. 25-33, 2001.
- [Hay98] S. Haykin, Neural Networks: A Comprehensive Foundation. Personal Education, Singapore, 1998. pp 43-56.
- [HHS92] T. K. Ho, J. J. Hull, S. N. Srihari, On Multiple Classifier Systems for Pattern Recognition. Center for Document Analysis and Recognition, State University of New York, USA, 1992.
- [Ho98] T. K. Ho. The random space method for constructing decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 832–844, 1998.
- [JNS94] Y. N. Jima, H. Nishida T. Suzuki, A fast and high-performance system for handwritten character recognition by efficient combination of multiple classifiers, Association for Pattern Recognition Workshop on Document Analysis Systems p.p. 175-189, Germany, 1994.
- [KSX92] Lei Xu, A. Krzyzak, C. Y. Suen, Methods Of Combining Multiple Classifiers And Their Applications to Handwriting recognition., IEEE Transactions Systems, Man and

Cybernetics Vol. 22(3). Center for Pattern Recognition & Machine Intelligence, Concordia Univ., Montreal, Que., 1992

- [Kun02] L. I. Kuncheva. A theoretical study on six classifier fusion strategies. IEEE Transactions on PAMI, 24(2):281–286, 2002.
- [Kun04] L. I. Kuncheva, Combining Pattern Classifiers, Methods and Algorithms, A Wiley-Interscience publication, USA, 2004.
- [KW02] L. I. Kuncheva, C. J. Whitaker, Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy. School of Informatics, University of Wales, Bangor UK, 2002
- [Lam00] L. Lam. Classifier combinations: Implementations and theoretical issues. In MCS: Proceedings of the First International Workshop on Multiple Classifier Systems, London, UK, pp. 77–86, 2000.
- [LM84] J. Little, C. Moler. The MathWorks, Inc. 3 Apple Hill Drive Natick, Massachusetts, 1984. <http://www.mathworks.se/matlabcentral/index.html>
- [Lee95] Dar Shyang Lee, A Theory in Classifier Combination: The Neural Network Approach, Faculty of the Graduate School of State University of New York, 1995
- [Lu96] Y. Lu. Knowledge integration in a multiple classifier system. Appl. Intell., 6(2):75–86, 1996.
- [MP43] W. McCulloch, W. Pitts, A logical calculus of the ideas immanent in the nervous activity, Bulletin of Mathematical Biophysics, vol. 5, 1943, pp. 115-133
- [OS96] D. W. Opitz and J. W. Shavlik. Generating accurate and diverse members of a neural-network ensemble, 1996.
- [Par96] D. Partridge. Network generalization differences quantified. Neural Networks, 1996.
- [RI04] Š. Raudys, M. Iwamura. Multiple Classifiers System for Reducing Influences of Atypical Observations, Vilnius Gediminas Technical University, Vilnius, Lithuania. Tohoku University, Aoba 05, Aramaki, Aoba, Sendai, 980-8579 Japan, 2004
- [Sha98] A. Sharkey. Multi-Net Systems, Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems, pp. 1-30, 1998.
- [Sri82] S. N. Srihari, Reliability Analysis of Majority Vote Systems. Information Sciences Department of Computer Science, State University of New York, USA, 1982 pp. 243-256
- [TSY06] E. K. Tang, P. N. Suganthan, X. Yao. An analysis of diversity measures. Machine Learning, v.65(1), pp. 247-271, 2006.

- [WMT99] Y. Wang, S. Ma, T. N. Tan, Combination Of Multiple Classifiers with Neural Networks, National Laboratory of Pattern Recognition, Institute of Automation, 1999
- [WYX04] Z. Wang, X. Yao, Y. Xu. An Improved Constructive Neural Network Ensemble Approach to Medical Diagnoses, School of Computer Science, The University of Birmingham, 2004.

Priedai

A priedas. Praktinio eksperimento bandymų programos ir gauti rezultatai.

Programų aktualios dalies MATLAB išeities tekstai, reikalingos bibliotekos ir eksperimento bandymų rezultatai (Microsoft Excell skaičiuoklės formatu) pateikiami kompaktiniame diske.