

VILNIAUS UNIVERSITETAS

TATJANA LIOGIENĖ

**HIERARCHINIS ŠNEKOS EMOCIJŲ KLASIFIKAVIMAS**

Daktaro disertacija  
Fiziniai mokslai, informatika (09P)

Vilnius, 2017

Disertacija rengta 2012–2016 metais, Vilniaus universitete.

Mokslinis vadovas:

doc. dr. Gintautas Tamulevičius (Vilniaus universitetas, fiziniai mokslai,  
informatika – 09P).

## **Padėka**

*Nuoširdžiai dėkoju moksliniam vadovui doc. dr. Gintautui Tamulevičiui už nuoseklų vadovavimą, vertingas mokslines konsultacijas, visokeriopą pagalbą, kantrybę bei moralinį palaikymą.*

*Dėkoju disertacijos recenzentams prof. habil. dr. Romualdui Baušiui ir prof. habil. dr. Kaziui Kazlauskui už naudingas pastabas bei vertingus patarimus, kurie leido pagerinti šio darbo kokybę.*

*Taip pat dėkoju Vilniaus universiteto Matematikos ir informatikos instituto direktoriui prof. habil. dr. Gintautui Dzemydai už suteiktą doktorantūros studijų galimybę bei visapusišką paramą studijuojant.*

*Nuoširdžiai dėkoju savo vyrui, dukrytei ir tėveliams už kantrybę, supratingumą ir besąlyginį palaikymą.*

*Tatjana Liogienė*

## Santrauka

Šioje disertacijoje yra nagrinėjamas emocionalios šnekos atpažinimo uždavinys, kuris yra labai aktualus ir leidžia pagerinti žmogaus ir kompiuterio sąveiką.

Disertacijoje yra išanalizuota žmogaus emocinės būklės įtaka šnekos generavimo procesui, aprašytas požymių atrankos procesas, apžvelgtos požymių rinkinių formavimo metodikos, atlikta hierarchinio klasifikavimo analizė, apžvelgtos emocijų atpažinimo šnekoje siūlomos hierarchinės klasifikavimo schemas bei jų naudojimo rezultatai.

Atliktos analizės pagrindu pasirinkti požymių rinkinio formavimo metodai (maksimalaus efektyvumo kriterijus, minimalios koreliacijos kriterijus, nuoseklus aibės didinimas) bei pasiūlyta Sprendimų medžių teorija paremta hierarchinė emocijų šnekoje klasifikavimo schema, kuri leidžia naudoti skirtingus požymių rinkinius kiekviename klasifikavimo etape.

Atliktas požymių atrankos kriterijų palyginimas parodė, kad tiek maksimalaus efektyvumo, tiek minimalios koreliacijos, tiek nuoseklus aibės didinimo atrankos metodai lėmė panašius klasifikavimo tikslumo rezultatus. Požymių rinkinių apimties požiūriu, efektyviausias pasirodė nuoseklus aibės didinimo metodas, kuris leido gauti iki 4 kartų mažesnės apimties požymių rinkinius, palyginti su kitais dviem atrankos metodais.

Pasiūlytos hierarchinės klasifikavimo schemas eksperimentinis tyrimas buvo atliktas nagrinėjant 3 emocijų (pykčio, džiaugsmo ir neutralios būsenos), 4 emocijų (pykčio, džiaugsmo, neutralios būsenos, liūdesio) ir 5 emocijų (pykčio, džiaugsmo, neutralios būsenos, liūdesio ir baimės) klasifikavimo uždavinius. Eksperimentinio tyrimo rezultatai įrodė hierarchinės klasifikavimo schemas pranašumą prieš tiesioginį klasifikavimą – pasiektas iki 40 proc. didesnis vidutinis klasifikavimo tikslumas.

Disertacijos tyrimų rezultatai buvo pristatyti 2 tarptautinėse konferencijose ir publikuoti 2 periodiniuose recenzuojamuose leidiniuose bei 3 recenzuojamuose konferencijų leidiniuose.

## **Abstract**

Speech emotion recognition problem which is considered relevant problem in human computer interaction area is tackled in this dissertation.

Influence of human emotional condition to speech generation process is analysed, described feature selection process, reviewed feature selection methods, hierarchical classification is analysed and reviewed classification schemes, its experimental results in this dissertation.

Based on research on feature selection methods the maximal efficiency criterion, minimal cross-correlation criterion, sequential forward selection technique are chosen and new hierarchical speech emotion classification scheme is proposed in this dissertation, which allows to use different feature sets in each classification stage.

Experimental research on comparison of feature selection criteria shows that maximal efficiency, minimal cross-correlation, sequential forward selection provides similar results. No obvious advantage of either method can be stated after the research at the viewpoint of classification accuracy. Nevertheless, the Sequential forward selection technique enabled us to obtain up to 4 times smaller feature sets in comparison with other two techniques.

The experimental research on newly proposed hierarchical classification scheme proves that the proposed scheme is more efficient than flat classification. For experimental purposes the classification problem of three (anger, joy, neutral), four (anger, joy, neutral, sadness) and five (anger, joy, neutral, sadness, fear) emotion classes were solved. The experimental results have proved hierarchical classification scheme advantage over flat classification – an improvement of 40 % in classification accuracy was achieved.

Research results are presented in two scientific international conferences, published in two reviewed scientific periodicals and in three scientific conference publications.

## Santrumpos

Lietuvių kalba:

KNN –  $k$ -artimiausių kaimynų algoritmas.

KT – vidutinis emocijų klasifikavimo tikslumas, naudojant skirtingus požymių atrankos kriterijus.

ME – maksimalaus efektyvumo kriterijus.

MK – minimalios koreliacijos kriterijus.

SFS – nuoseklaus aibės didinimo metodas.

PR – pilnas požymių rinkinys.

Anglų kalba:

DT – Decision trees.

KNN –  $k$ -nearest neighbors classifier.

ME – Maximal efficiency criterion.

MC – Minimal cross-correlation criterion.

SFS – Sequential forward selection technique.

# Turinys

<b>1. Įvadas.....</b>	<b>1</b>
1.1 Tyrimo sritis ir problemos aktualumas.....	1
1.2 Tyrimų objektas.....	1
1.3 Darbo tikslas ir uždaviniai.....	2
1.4 Mokslinis darbo naujumas.....	2
1.5 Tyrimo metodika.....	3
1.6 Ginamieji teiginiai.....	3
1.7 Darbo rezultatų aprobavimas.....	3
1.8 Disertacijos struktūra.....	6
<b>2. Emocijų klasifikavimo uždavinys.....</b>	<b>7</b>
2.1 Emocijų poveikis šnekai.....	8
2.2 Emocijoms šnekoje atpažinti naudojami požymiai.....	9
2.2.1 Požymių rūšys.....	10
2.2.2 Požymių kombinacijos.....	13
2.3 Požymių atrankos procesas.....	14
2.4 Požymių rinkinio formavimo metodai.....	18
2.5 Hierarchinio klasifikavimo analizė.....	20
2.5.1 Klasifikavimo būdai.....	21
2.6 Emocijų klasifikavimo hierarchinių schemų apžvalga.....	26
2.7 Antro skyriaus apibendrinimas.....	36
<b>3. Hierarchinė emocijų klasifikavimo schema.....</b>	<b>38</b>
3.1 Sprendimų medžiai.....	38
3.1.1 Sprendimų medžio sudarymas.....	40
3.1.2 Sprendimų medžius realizuojantys algoritmai.....	41
3.1.3 Sprendimų medžių metodo privalumai ir trūkumai.....	42
3.2 Pasiūlytoji hierarchinė emocijų klasifikavimo schema.....	43
3.2.1 Apibendrinta hierarchinio klasifikavimo schema naudojant keletą lygių požymių.....	45



3.3	Klasifikavimo klaidos vertinimas.....	47
3.4	Požymių atrankos kriterijai.....	48
3.4.1	Maksimalaus efektyvumo požymių atranka.....	49
3.4.2	Minimalios koreliacijos požymių atranka.....	50
3.4.3	Nuoseklus aibės didinimo metodas.....	52
3.5	Trečio skyriaus apibendrinimas.....	53
<b>4.</b>	<b>Eksperimentinis tyrimas.....</b>	<b>54</b>
4.1	Tyrimo duomenys.....	54
4.2	Tyrimo priemonės.....	56
4.3	Tyrimo sąlygos.....	56
4.4	Tyrimo eiga ir rezultatai.....	58
4.5	Eksperimentų rezultatai.....	59
4.5.1	Atrankos kriterijų palyginimas.....	59
4.5.2	Hierarchinis 3 emocijų klasifikavimas.....	63
4.5.3	Hierarchinis 4 emocijų klasifikavimas.....	67
4.5.4	Hierarchinis 5 emocijų klasifikavimas.....	70
4.5.5	Alternatyvios hierarchinės schemos.....	77
4.6	Skyriaus apibendrinimas.....	80
	<b>Bendros išvados.....</b>	<b>81</b>
	<b>Literatūros sąrašas.....</b>	<b>83</b>

# **1. Įvadas**

## **1.1 Tyrimo sritis ir problemos aktualumas**

Istoriškai žmogaus ir kompiuterio sąveika prasidėjo nuo komandinio formato – visos instrukcijos kompiuteriui buvo užduodamos įvairiais formatais išreikštomis komandomis. Ilgainiui buvo pasiūlyta grafinė naudotojo sąsaja, kurios pagrindimas rėmėsi idėja, jog didžiąją dalį informacijos žmogus priima vizualiai. Šiame dešimtmetyje itin pradėjo populiarėti balso sąsaja, kurios esmė – verbaline forma grįsta žmogaus ir kompiuterio sąveika. Tokios sąsajos idėja remiasi teiginiu, jog verbalinis bendravimas yra pats natūraliausias žmogaus bendravimo būdas, kuris gali padidinti sąveikos su kompiuteriu efektyvumą.

Neatsiejama verbalinio bendravimo dalis yra emocijos. Emocija, kaip ir kita neverbalinė raiška (pavyzdžiui, veido išraiška, laikysena), perduoda dalį informacijos, kuria remdamiesi mes formuojame savo reakciją ir atsaką į gaunamą žinią. Taigi neverbalinės informacijos analizė tik padidins balso sąsajos efektyvumą. Tuo tikslu ir yra vykdomi šnekos emocijų atpažinimo tyrimai, tikintis sukurti patikimus, įvairiems veiksniams atsparius šnekos emocijų atpažinimo metodus, kurie leis žmogaus ir kompiuterio sąsajai suteikti daugiau natūralumo, informatyvumo. Kita vertus, šnekos emocijų analizė galėtų būti sėkmingai pritaikoma kriminalistikoje, skambučių centruose, kuriant robotus ir kitose srityse.

## **1.2 Tyrimų objektas**

Disertacijos tyrimų objektas – emocijų požymių šnekos signale atranka bei emocionalios šnekos klasifikavimas, siekiant atpažinti šnekančiojo emocinę būklę.

### **1.3 Darbo tikslas ir uždaviniai**

Pagrindinis darbo tikslas yra išnagrinėti šnekos emocijų klasifikavimo uždavinį ir pasiūlyti sprendimus, leidžiančius padidinti klasifikavimo tikslumą bei sumažinti reikalingų požymių aibę.

Siekiant suformuoti tikslo, buvo sprendžiami šie uždaviniai:

1. pasiūlyti hierarchinę emocijų klasifikavimo schemą, kuri leistų padidinti klasifikavimo tikslumą lyginant su tiesiogine schema (kurioje visos emocijos klasifikuojamos iš karto vienu žingsniu);
2. hierarchinei klasifikavimo schemai suformuoti ir pritaikyti požymių atrankos metodus, leidžiančius padidinti klasifikavimo tikslumą bei sumažinti reikalingų požymių aibę;
3. atlikti pasiūlytos hierarchinės emocijų klasifikavimo schemas eksperimentinį tyrimą, įvertinti gaunamą emocijų klasifikavimo tikslumą, požymių atrankos metodų įtaką klasifikavimui.

### **1.4 Mokslinis darbo naujumas**

Disertacijoje yra nagrinėjamas emocionalios šnekos atpažinimo uždavinys. Yra atlikta emocionalios šnekos klasifikavimo schemų analizė ir pasiūlyta hierarchinė klasifikavimo schema. Pasiūlyta schema yra visiškai nepriklausoma nuo psichologinių, socialinių ir kitų emocijas apibūdinančių veiksnių. Ji leidžia atlikti efektyvų emocionalios šnekos įrašų klasifikavimą naudojant tik akustinius šnekos signalo požymius. Klasifikavimo procese buvo pritaikyti trys skirtingi požymių atrankos metodai, keliasdešimt kartų sumažinantys naudojamų požymių aibes. Pasiūlytoji hierarchinė klasifikavimo schema buvo pritaikyta itin didelės apimties (5 000 įrašų) emocionalios lietuvių šnekos įrašams klasifikuoti. Eksperimentiniai rezultatai patvirtino pasiūlytos schemas efektyvumą – padidėjusį klasifikavimo tikslumą bei sumažėjusią požymių aibę.

## 1.5 Tyrimo metodika

Darbo tikslui pasiekti ir uždaviniams spręsti buvo atliekama literatūros apžvalga, teorinė analizė, įvykdytas eksperimentinis žvalgomojo (angl. *exploratory*) pobūdžio tyrimas. Darbe buvo panaudotos algoritmų teorijos, duomenų gavybos, statistinės analizės, atpažinimo teorijos, skaitmeninio signalo apdorojimo žinios.

## 1.6 Ginamieji teiginiai

- Hierarchinė emocionalios šnekos klasifikavimo schema tikslumo požiūriu yra iš esmės efektyvesnė už tiesioginį (vieno etapo) klasifikavimą.
- Požymių atrankos taikymas leidžia reikšmingai sumažinti nagrinėjamų duomenų kiekį ir kartu padidinti klasifikavimo efektyvumą, palyginti su pilnais požymių rinkiniais.
- Hierarchinėje klasifikavimo schemoje skirtingi požymių atrankos metodai neturi esminės įtakos visos schemos efektyvumui.
- Didėjant nagrinėjamų emocijų skaičiui, vidutinis emocijų klasifikavimo tikslumas mažėja, o klasifikavimo tikslumui maksimizuoti reikalingas požymių kiekis didėja.

## 1.7 Darbo rezultatų aprobavimas

Pagrindiniai tyrimo rezultatai paskelbti mokslinėse publikacijose, pristatyti tarptautinėse mokslininkų bei respublikinėse konferencijose.

Pranešimai skaityti šiose konferencijose:

- Federated Conference on Computer Science and Information System. Pranešimas „Comparative Study of Multi-stage Classification Scheme for Recognition of Lithuanian Speech Emotions“. Lenkija, Gdanskas, rugsėjo 11–14 d., 2016.

- 3rd IEEE Workshop on Advances in Information, Electronic and Electrical Engineering. Pranešimas „SFS Feature Selection Technique for Multistage Speech Emotion Recognition“. Ryga, Latvija, lapkričio 13–14 d., 2015.
- Elektros, elektronikos ir informatikos mokslų konferencija „eStream 2015“. Pranešimas „Minimalios koreliacijos kriterijus daugiapakopių šnekos emocijų atpažinimo požymių atrankai“. Lietuva, Vilnius, balandžio 21 d., 2015.
- Tarpuniversitetinė magistrantų ir doktorantų konferencija „Informacinė visuomenė ir universitetinės studijos 2014“. Pranešimas „Dinaminiai pagrindinio tono dažnio požymiai kalbos emocijoms atpažinti“. Lietuva, Kaunas, balandžio 24 d., 2014.

Stendiniai pranešimai pristatyti šiose konferencijose:

- 7th International Workshop „Data Analysis Methods for Software Systems“. Stendinis pranešimas „Multistage Speech Emotion Recognition for Lithuanian: experimental study“. Lietuva, Druskininkai, gruodžio 3–5 d., 2015.
- 6th International Workshop „Data Analysis Methods for Software Systems“. Stendinis pranešimas „Low-order Multi-level Features for Speech Emotions Recognition“. Lietuva, Druskininkai, gruodžio 4–6 d., 2014.

Publikacijos periodiniuose recenzuojamuose leidiniuose:

- Liogienė T., Tamulevičius G., 2016. Multi-stage Recognition of Speech Emotion Using Sequential Forward Feature Selection. *Journal on Electrical, Control and Communication Engineering*. Vol.10 35–41, ISSN 2255-9140, e-ISSN 2255-9159.

- Tamulevičius G., Liogienė T., 2015. Low-order Multi-level Features for Speech Emotion Recognition. *Baltic Journal of Modern Computing Vol. 3, No. 4*: 234–347. ISSN 2255-8950.

Publikacijos recenzuojamuose konferencijų leidiniuose:

- Liogienė T., Tamulevičius G., 2016. Comparative Study of Multi-stage Classification Scheme for Recognition of Lithuanian Speech Emotions. *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems. ACSIS, Vol. 8*: 483–486, ISSN 2300-5963.
- Liogienė T., Tamulevičius G., 2015. SFS Feature Selection Technique for Multistage Emotion Recognition. *Proceeding of the 2015 IEEE 3th workshop on Advances in Information, Electronic and Electrical Engineering*: 1–4, ISBN 978-1-5090-1201-5.
- Liogienė T., Tamulevičius G., 2015. Minimal Cross-correlation Criterion for Speech Emotion Multi-level Feature Selection. *Proceedings of the Open Conference of Electrical, Electronic and Information sciences (eSTREAM). Washington, IEEE*: 1–4, ISBN 978-1-4673-7445-3.
- Liogienė T., 2014. Dinaminiai pagrindinio tono dažnio požymiai kalbos emocijoms atpažinti. Informacinės technologijos: 19-oji tarpuniversitetinė magistrantų ir doktorantų konferencija „Informacinė visuomenė ir universitetinės studijos“: 161–166. ISSN 2029-4832.

Santraukos konferencijų leidiniuose:

- Liogienė T., Tamulevičius G. Multistage Speech Emotion Recognition for Lithuanian: experimental study. *Data Analysis Methods for Software Systems: 7th International Workshop*: [abstracts book], Druskininkai, Lietuva, ISBN 9789986680581, gruodžio 3–5 d., 2015, p. 34.

- Liogienė T., Tamulevičius G. Low-order Multi-level Features for Speech Emotions Recognition. *Data Analysis Methods for Software Systems: 6th International Workshop*: [abstracts book], Druskininkai, Lietuva, ISBN 978-9986-680-50-5, 2014, p. 35.

## **1.8 Disertacijos struktūra**

Disertaciją sudaro 4 skyriai: Įvadas, Emocijų klasifikavimo uždavinys, Hierarchinė emocijų klasifikavimo schema, Eksperimentinis tyrimas bei bendrosios išvados su literatūros sąrašu. Disertacijos apimtis: 100 puslapių, 7 lentelės, 37 iliustracijos. Disertacijoje remtasi 90 literatūros šaltinių.

## 2. Emocijų klasifikavimo uždavinys

Emocijų atpažinimas šnekos signale kaip savarankiška mokslinių tyrimų sritis atsirado praeito amžiaus devintojo dešimtmečio viduryje. Nepaisant intensyviai atliekamų tyrimų, šnekos emocijų atpažinimo uždavinys nėra visiškai išspręstas – emocijų atpažinimo tikslumas dar nėra itin aukštas, nėra pasiūlyta unikalių ir savo efektyvumu išskirtinių požymių sistemų bei atpažinimo metodų. Tam galima rasti bent keletą priežasčių: ribotos duomenų panaudojimo galimybės (šnekos emocijų duomenų bazių nėra daug, o ir jų panaudojimo galimybes sunkina nuosavybės teisės, asmens duomenų apsauga ir kiti klausimai), nepakankamos žinios apie fiziologines žmogaus emocijų apraiškas balse, kryptingų tyrimų nebuvimas.

Nepaisant to, emocinės būsenos atpažinimas yra aktualus uždavinys ir turi didelį taikymo potencialą: emocinės būsenos identifikavimas gali būti naudojamas kriminalistikoje, skambučių centruose (vertinant skambinančiojo emocinę būklę), kuriant robotus (įgyvendinant reakciją į skirtingas emocines būkles) ir kitose srityse, sudarant prielaidas efektyvesnei žmogaus ir kompiuterio sąveikai (Jang & Kwon, 2006), (Ververidis & Kotropoulos, 2006), (Hyun, et al., 2007), (Vogt, et al., 2008), (Park, et al., 2009), (Ramakrishnan & El Emary, 2013).

Šnekos emocijos atpažinimo uždavinys savo turiniu yra klasikinis atpažinimo uždavinys. Iš esmės emocijų atpažinimo procesą sudaro 3 pagrindiniai etapai: šnekos signalo analizė, klasifikatoriaus apmokymas ir klasifikavimo etapas (Rong, et al., 2009), (Anagnostopoulos, et al., 2012), (Gjoreski & Gjoreski, 2014). Pagrindinis šnekos signalo analizės etapo tikslas – išskirti charakteringus duomenis, vadinamuosius požymius, atspindinčius šnekos signale užkoduotą emociją. Tuo tikslu šnekos signalas apdorojamas (dažniausiai filtruojamas siekiant pašalinti nepageidaujamas signalo savybes), segmentuojamas (skaidomas į persidengiančius kadrus), išskiriami pasirinktieji požymiai. Atliekant konkretų taikymą, gali būti išskiriama daugybė įvairių



požymių – akustiniai, t. y. atspindintys spektrines signalo savybes, prozodiniai, t. y. apibūdinantys šnekos intonaciją, ir kiti (Gjoreski & Gjoreski, 2014). Visi išskirtieji požymiai sudaro vadinamąjį požymių rinkinį. Požymių rinkinio dydis apibūdinamas skirtingų požymių kiekiu rinkinyje. Suformuotieji požymių rinkiniai panaudojami klasifikatoriui apmokyti (apmokymo etapas) arba nežinomai emocijai šnekos signale identifikuoti (atpažinimo etapas) (Busso, et al., 2012).

## **2.1 Emocijų poveikis šnekai**

Verbalinio bendravimo metu informacija yra perduodama šnekos signalu. Šnekos signalas – konkrečią žinią pernešantis signalas, generuojamas sąmoningais šnekos organų veiksmis. Savo prigimtimi šnekos signalas yra akustinis signalas – atmosferos aplinkoje jis sklinda kaip oro tankio svyravimas. Pakaitomis einančios sutankėjimų ir praretėjimų sritys sklinda ore kaip garso banga.

Emocijų šnekos signale analizei yra svarbu suprasti kaip stresas ir emocinė būseną daro įtaką kalbos signalo generavimui. Stresas yra psichologinė būseną, atsakanti už grėsmės suvokimą ir įprastai yra lydima specifinių emocijų (pavyzdžiui, baimės, pykčio, nerimo ir kt.) (Hansen & Patil, 2007). Streso sukelti pokyčiai gali paveikti šneką net prieš asmens valią.

Kalbančiojo asmens emocinė būseną tiesiogiai veikia šnekos generavimo procesą – pakitęs kvėpavimas, nevalingas krūtinės raumenų įsitempimas, burnos ertmės sausumas, kalbos padargų virpesiai ir kiti procesai neišvengiamai daro įtaką generuojamam šnekos signalui (Scherer, 2005), (Hyun, et al., 2007).

Kvėpavimas dažnai yra tikslus rodiklis tam tikrose emocinėse situacijose. Pavyzdžiui, kai asmuo patiria stresinę situaciją, jo kvėpavimo dažnis padidėja, o tai kalbėjimo metu balsaskylės pradžioje padidina slėgį, lemiantį pagrindinio tono dažnio padidėjimą. Padidėjus kvėpavimo dažniui, kalbėjimo trukmė tarp įkvėpimų sumažėja ir tai paveikia tarimo / artikuliacijos greitį (Hansen & Patil, 2007).

Burnos ertmės sausumas būdingas susijaudinimo, baimės, pykčio atvejais. Šios emocijos gali daryti įtaką ir pagrindiniams šnekos signalo generavimo etapams (pavyzdžiui, gerklų raumenų aktyvumui, balso klosčių būsenai). Gerklų raumenų veikla ir balso klosčių virpėjimas tiesiogiai veikia garso greitį per balsaskylę, o tai daro įtaką pagrindinio tono dažniui (Hansen & Patil, 2007).

Taigi, šnekos generavimo sistemai įtaką gali daryti skirtingi stresą sukeltantys veiksniai – stresoriai, kurie atlieka skirtingus vaidmenis šnekos generavimo metu (Hansen & Patil, 2007):

- fiziniai stresoriai – balso trakto formos pokyčiai, sukelti vibracijos ar judėjimo, kurie tiesiogiai paveikia artikulatorius;
- nesąmoningi fiziologiniai stresoriai – streso sukelti kvėpavimo dažnio ar raumenų įsitempimo pokyčiai, kuriuos gali sukelti miego trūkumas, nuovargis ir pan.;
- sąmoningi fiziologiniai stresoriai – streso sukeltas vokalių pastangų padidėjimas. Pavyzdžiui, balso pakėlimas, kad kalbantysis galėtų išgirsti save triukšmingoje aplinkoje. Taip gali nutikti būnant triukšmingoje aplinkoje, įvykus kokiam nors atsitikimui, patiriant pakitusią emocinę būseną.

Šnekamojoje kalboje emocijos buvimas yra aiškiai užkoduotas šnekos signale, bet lieka neaišku, kaip tiksliai nustatyti specifinius požymius, kurie padėtų identifikuoti emociją (Cowie, et al., 2001), (Bitouk, et al., 2010), (Origlia, et al., 2010), (Planet & Iriondo, 2012).

## **2.2 Emocijoms šnekoje atpažinti naudojami požymiai**

Populiariausi ir dažniausiai emocijoms šnekoje atpažinti yra naudojami prozodiniai ir spektro požymiai (Hansen & Patil, 2007), (Rong, et al., 2009), (Bitouk, et al., 2010), (Koolagudi, et al., 2010), (Planet & Iriondo, 2012). Požymiams išskirti yra naudojami automatizuoti įrankiai. Vienas iš jų – *OpenEAR* sistema (Eyben, et al., 2009). Ši sistema leidžia išskirti net 6 552 skirtingus šnekos požymius bei įvairias jų statistines reikšmes.

### 2.2.1 Požymių rūšys

**Prozodiniiais** vadinami požymiai, gaunami iš pagrindinio tono dažnio, formančių dažnių reikšmių, balso intensyvumo, energijos, pauzių, šnekos trukmės ir greičio, balso kokybės charakteristikų (Adell, et al., 2005), (Koolagudi & Rao, 2012). Dažnai papildomai yra naudojama įvairi šių požymių statistika: vidutinė reikšmė, mediana, standartinis nuokrypis, dispersija, minimali ir maksimali reikšmės, kvantiliai ir t. t. Statistika besiremiantys išvestiniai požymiai apibūdina ilgesnę signalo atkarpą ir yra vadinami globaliais požymiais. Visi šie požymiai yra statiniai savo prigimtimi ir atspindi momentines šnekos signalo savybes (Vogt & Andre, 2005), (Ververidis & Kotropoulos, 2006), (Bitouk, et al., 2010), (Koolagudi & Rao, 2012).

Pavyzdžiui, naudojant *OpenEAR* sistemą, galima išskirti net 234 pagrindinio tono dažnio požymius, kurie yra suskirstyti į 6 pagrindinio tono dažnio grupes: vidutinius statinius ir funkcinius požymius (pavyzdžiui, reikšmių diapazonas, vidurkis), pirmos eilės dinامينius ir funkcinius požymius, antros eilės dinامينius ir funkcinius požymius, pagrindinio tono dažnio kontūro statinius požymius, kontūro pirmos eilės dinامينius požymius, kontūro antros eilės dinامينius požymius. Kiekvieną grupę sudaro 39 požymiai (absoliuti ir aritmetinė F0 kontūrų reikšmės (angl. *the absolute and the arithmetic means of the F0 contour*), minimalios ir maksimalios F0 reikšmių pozicijos (angl. *positions of the minimal and maximal F0 values*), įvairių eilių statistiniai duomenys, kvartiliai ir kiti).

Teigiama, jog pagrindinio tono dažnio vidurkis yra puikus streso būsenos rodiklis. Šnekant garsiai ir piktai pagrindinio tono dažnio reikšmių vidurkis ir dispersija, palyginti su kitų būsenų duomenimis, labai skiriasi. Tačiau, pagrindinio tono dažnio dispersija nėra patikimas rodiklis. Be to reikėtų turėti omenyje, jog pagrindinio tono dažnis yra individuali šnekančiojo charakteristika, todėl šiuo parametru perduodama ir informacija apie šnekančiąjį, t. y. pagrindinio tono savybės priklauso nuo šnekančiojo.

Vidutinė žodžio trukmė taip pat priklauso nuo šnekančiojo emocinės būklės. Priebalsių, pusbalsių ir dvibalsių trukmė yra vienoda esant tiek švelniai, tiek garsiai, tiek neutraliai šnekai. Balsių trukmė, palyginti su neutralia šneka, nežymiai mažėja šnekant švelniai ir nežymiai didėja šnekant garsiai.

Intensyvumas, palyginti su neutralia šneka, didesnis garsioje ir piktoje šnekoje, o švelnioje ir aiškioje – mažesnis. Balsių intensyvumas lieka pastovus esant lėtai ir aiškiai šnekai, o priebalsių intensyvumas padidėja esant švelniai ir piktai šnekai. Taigi, žodžio intensyvumas gali būti naudojamas klasifikuojant streso būsenas.

Siekiant tiksliai atpažinti emocijas natūralioje šnekoje ir dialoguose, turi būti atsižvelgta ne tik į prozodinę informaciją. Dažnai kartu su pagrindinio tono dažnio ir energijos požymių statistika yra naudojami **spektro** požymiai (Ayadi, et al., 2011), (Anagnostopoulos, et al., 2012), kurie remiasi trumpalaikio spektro įverčiais, kaip antai, tiesinės prognozės koeficientai, melų skalės kepstro koeficientai, Furjė spektro koeficientai bei jų statistika.

Spektrinius požymius yra sunkiau intuityviai susieti su emocine būsena, jie detalai aprašo spektrines šnekos signalo savybes ir gali pagerinti emocijos atpažinimo tikslumą, palyginti su prozodiniais požymiais. Visgi reikėtų pažymėti, kad spektro požymiuose galima išžvelgti nemažai informacijos apie lingvistinį šnekos turinį (tai, kas sakoma), tai daro šiuos požymius priklausomus nuo kalbos turinio (Vogt & Andre, 2005), (Bitouk, et al., 2010), (Ayadi, et al., 2011), (Koolagudi & Rao, 2012). Tai reiškia, jog emocijos atpažinimo rezultatai lemia ne tik emocinė būsena, užkoduota šnekos signale, bet ir jo lingvistinis turinys. Ir tai gali būti papildomas netikslumo šaltinis.

Įvairios emocinės būklės šnekai būdingas spektrinis nuolydis (spektro dedamųjų amplitudės sumažėjimas) gerokai skiriasi nuo neutralios šnekos. Stipresnis spektrinis nuolydis (-9 dB/oktavą) yra būdingas greitai, garsiai ir piktai šnekai, o staigus (-15 dB/oktavą) – lėtai ir švelniai šnekai. Taigi, spektrinis nuolydis yra svarbus perduodant stresą ir geras identifikatorius išskiriant šnekos stilius.

Emocijų atpažinime kartu su požymių statistika taip pat yra naudojami **epochų parametrai**. Šnekos signalo generavimo metu balso klostės virpa užtverdamos ir atverdamos kelią oro srautui iš plaučių. Uždarasis balso klosčių (balsaskylės) periodas yra vadinamas epocha. Požymiai, išskirti iš signalo atkarpos, atitinkančios šį laikotarpį, yra vadinami epochų parametrais. Pasiūlyti tokie epochų parametrai kaip epochos stiprumas (angl. *strength of epoch*), momentinis dažnis, epochos ryškumas (angl. *sharpness of epoch*), epochos nuolydžio stiprumas (Koolagudi, et al., 2010).

**Balso kokybės** požymiai apibūdina balsaskylės pradžios (angl. *glottal source*) savybes ir jiems įtaką daro šnekos sužadavimo signalas. Taigi, skirtingos emocinės būsenos gali lemti balso kokybę. Kartu su pagrindinio tono dažniu ir trukme balso kokybė priskiriama prie šnekos prozodijos. Emocijoms analizuoti šnekoje yra naudojami tokie balso kokybės požymiai kaip sužadavimo signalo savybės, artikuliacijos būdas, balso tembras, pagrindinio tono dažnio svyravimas ir garsumo svyravimas (Tato, et al., 2002), (Lugger & Yang, 2006), (Lugger & Yang, 2007), (Zhang, 2008), (Ayadi, et al., 2011).

Kitas svarbus požymis emocijoms atpažinti yra **harmonikų**, susidarančių dėl netiesinių balso trakto savybių, **skaičius**. Pagrindinio tono dažnis nulemia harmonikų kiekį spektre. Kuo tono dažnis yra aukštesnis, tuo mažiau harmonikų susidaro, ir atvirkščiai – kuo tono dažnis yra žemesnis, tuo daugiau harmonikų susidaro (Ververidis & Kotropoulos, 2006), (Origlia, et al., 2010). Geresniam emocinės šnekos apibūdinimui, kartu su populiariais pagrindinio tono dažniu ir energijos požymių statistika, siūloma naudoti naujus harmonikų ir **Zipf** grindžiamus (angl. *Zipf based*) požymius. Harmonikų požymiai apibūdina signalo spektrą remdamiesi pagrindinio tono ir formančių struktūra. **Zipf** požymiai charakterizuoja vidinę signalo struktūrą, ypač ritinius ir prozodinius balso išraiškos aspektus (Xiao, et al., 2009).

Apibendrinant galima teigti, kad dažniausiai, kaip šnekos emocijų požymiai, naudojami pagrindinio tono dažnis, ištaramo trukmė ir energija bei jų statistika: minimali, vidutinė ir maksimali reikšmės, mediana, standartinis

nuokrypis, dispersija, maksimumo ir minimumo skirtumas, kiti išvestiniai požymiai (Vogt & Andre, 2005), (Ververidis & Kotropoulos, 2006), (Bitouk, et al., 2010), (Origlia, et al., 2010), (Koolagudi & Rao, 2012), (Rao, et al., 2013).

## 2.2.2 Požymių kombinacijos

Emocijoms šnekoje atpažinti naudojami požymių rinkiniai yra sudaromi iš įvairių rūšių požymių. Taip yra siekiama pagerinti emocijų atpažinimo tikslumą.

Prozodiniai požymiai (dažniausiai F0 ir energija) yra klasikiniai požymiai, naudojami atliekant daugelį emocijų atpažinimo eksperimentų. Tačiau siekiant tiksliai atpažinti emocijas (natūralioje šnekoje, dialoguose) turi būti atsižvelgta ne tik į prozodinę informaciją (Vidrascu & Devillers, 2007).

Literatūroje yra pasiūlytas šnekos emocijų požymių, remiantis fonemų informacija, išskyrimo metodas. Pirmiausia, buvo įvertinti tokie prozodiniai požymiai, kaip suvokiamas pagrindinis tonas, energija ir formančių savybės, kurie yra paprastai naudojami emocijoms atpažinti. Antra, požymiai buvo suskirstyti į emocijas atspindinčius požymius ir į dominuojančių fonemų požymius. Šis emocinių požymių išgavimo metodas remiasi fonemų informacija ir suteikia galimybę šnekos emocijoms atpažinti panaudoti požymius, kuriuose yra atsižvelgta į fonemos poveikį (Hyun, et al., 2007).

Dar yra siūloma naudoti prozodinius požymius ir kokybės požymius, priimant galutinį sprendimą apie emocinę būseną (Tato, et al., 2002).

Harmonikų ir *Zipf* požymių naudojimas emocijoms šnekoje atpažinti parodo stiprų kokybės pagerinimą, kai yra naudojamos kartu su populiariais dažnio ir energijos grindžiamais požymiais (Xiao, et al., 2009).

Literatūroje taip pat nagrinėjami požymiai, gauti iš suvokiamo tono, energijos ir MFCC požymių sekų. Požymiai buvo atrenkami siekiant palyginti emocijų atpažinimą suvaidintoje ir natūralioje emocinėje šnekoje. Tyrimas parodė, kad suvaidintas emocijas yra paprasčiau atpažinti nei natūralias. Požymių pasirinkimo įtaka yra didesnė suvaidintai šnekai. Suvaidintoje

šnekoje tono požymiai (angl. *pitch-related features*) yra dominuojantys, dažniau naudojami. Natūralioje šnekoje dėmesys sutelkiamas daugiau į MFCC požymius. Pauzės yra labai svarbus požymis suvaidintoms emocijoms, daug pauzių pastebėta liūdnose emocijose. Tačiau tai netinka natūralioms emocijoms, nes pauzės ne visada naudojamos (Vogt & Andre, 2005).

Apibendrinant galima pasakyti, kad nėra bendro susitarimo dėl pagrindinio požymių sąrašo ir galima daryti išvadą, kad požymių pasirinkimas priklauso nuo analizuojamų duomenų (Vidrascu & Devillers, 2007).

## 2.3 Požymių atrankos procesas

Tinkamo požymių rinkinio sudarymas yra vienas svarbiausių sprendimų, lemiančių emocijų atpažinimo tikslumą. Deja, galima konstatuoti, kad vienintelės ir teisingos požymių išskyrimo metodikos kol kas nėra pasiūlyta (Vogt & Andre, 2005), (Anagnostopoulou, et al., 2012).

Teorinė požymių bei jų charakteristikų (pavyzdžiui, efektyvumo) analizė yra praktiškai neįmanoma, todėl analizuojant emocijas šnekos signalo požymiai labai dažnai yra pasirenkami eksperimentiniu pagrindu (Koolagudi & Rao, 2012). Tokia požymių analizė lemia formuojamųjų požymių rinkinių suboptimalumą, t. y. gautieji požymių rinkiniai negarantuoja optimalaus klasifikavimo. Konkrečiam duomenų rinkiniui suformuotasis požymių rinkinys yra lokaliai maksimalaus efektyvumo, t. y. jis garantuoja maksimalų klasifikavimo tikslumą tik nagrinėjamam duomenų (emocijų pavyzdžių) rinkiniui. Pasikeitus duomenims, tikėtina, kad bus gaunamas nebe maksimalus galimas klasifikavimo tikslumas. Įvertinus šnekos signalo nestacionarumą, žmogaus emocijų individualumą, sunkiai formalizuojamą emocijų poveikį kalbos procesui, šnekos emocijų požymių optimizavimas yra neįmanomas.

Išskiriamų požymių kiekis yra neapibrėžtas ir jų skaičius kartais siekia keletą tūkstančių požymių (Schuller, et al., 2009), (Chiou & Chen, 2013), (Gjoreski & Gjoreski, 2014). Tokius didelius požymių rinkinius neišvengiamai tenka mažinti dėl dviejų priežasčių. Pirma, norint tinkamai apmokyti klasifikatorių, didelis požymių rinkinys reikalauja milžiniškų apmokymui

skirtų duomenų kiekių (emocingos šnekos įrašų). Antra, didelės apimties požymių rinkiniai reiškia ilgą apmokymą ir klasifikavimo procesą, todėl požymių rinkinius tenka mažinti (Lugger & Yang, 2007), (Anagnostopoulos, et al., 2012). Mažinant požymių rinkinius, kyla rizika prarasti emocijų klasifikavimui reikšmingus požymius, todėl tenka naudoti požymių rinkinio optimizavimo metodus (Origlia, et al., 2010).

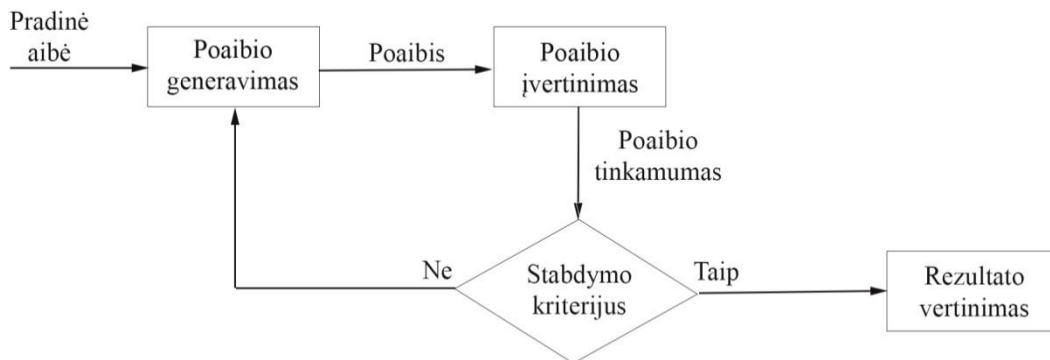
Požymių rinkinio formavimo metodus galima suskirstyti į dvi grupes: požymių atrankos ir požymių transformavimo (Rong, et al., 2009).

Požymių atrankos metodai leidžia atrinkti rinkinių poaibius atmetant klasifikavimui nereikšmingus požymius arba sujungiant požymius, lemiančius didžiausią emocijų klasifikavimo tikslumą.

Reikšmingų požymių aibėms formuoti taikomi įvairūs vertinimo kriterijai. Kadangi emocijų požymiai yra daugiamačiai duomenys (tai reiškia, jog turimų pavyzdžių kiekis yra gerokai mažesnis už duomenų dimensijų skaičių), klasifikavimo požiūriu optimalaus požymių poaibio suradimas yra sudėtinga užduotis (Kumar & Minz, 2014), (Mencattini, et al., 2014).

Apibendrintą požymių atrankos procesą sudaro keturi pagrindiniai žingsniai (1 pav.) (Dash & Liu, 1997):

1. poaibio generavimas;
2. poaibio įvertinimas;
3. stabdymo kriterijaus nustatymas;
4. galutinio rezultato vertinimas.



1 pav. Požymių atrankos procesas (Kumar & Minz, 2014)



Poaibio generavimas yra paieškos procedūra, naudojanti tam tikrą paieškos strategiją (Liu & Motoda, 1998). Atsižvelgiant į tam tikrus vertinimo kriterijus, sugeneruotas požymių poaibis lyginamas su geriausiu požymių poaibiu. Jei naujasis požymių poaibis yra geresnis, tuomet jis išsaugomas kaip geriausias požymių poaibis. Ši procedūra tęsiama tol, kol neįvykdoma stabdymo sąlyga. Po paieškos procedūros gautas rezultatas – gautasis požymių rinkinys – yra įvertinamas. Vertinimas gali būti atliekamas naudojant generuotus (dirbtinius) duomenis arba tikrus duomenis (angl. *real-word data*). Apibendrintas požymių atrankos algoritmas yra pateiktas 1 lentelėje.

1 lentelė. Požymių atrankos algoritmas (Kumar & Minz, 2014)

<i>ĮVESTIS:</i>	$X$ : Duomenų rinkinio požymių aibė, turinti $N$ požymių $SG$ : Kito poaibio (paveldėtojo (angl. <i>successor generation</i> )) generatorius $E$ : Įvertinimo matas $\theta$ : Algoritmo nutraukimo sąlyga
<i>IŠVESTIS:</i>	$X_{opt}$ : Optimalus požymių rinkinys, kurio tikslumas yra didžiausias, arba svoriniai požymiai
<i>Inicijavimas:</i>	$X' := Pradžios\_taškas(X)$ ; $X_{opt} := \{Geriausias\ X' \text{ naudojant } E\}$ ;
<i>Kartojimas:</i>	$X' := Paieškos\_strategija(X', SG(E), X)$ ; $X_{opt} := \{Geriausias\ X' \text{ atsižvelgiant į } E\}$ ; <i>Jeigu</i> $E(X') \geq E(X_{opt})$ <i>arba</i> $(E(X') == E(X_{opt}) \&  X'  <  X_{opt} )$ <i>Tada</i> $X_{opt} = X'$ ;
<i>Iki</i>	Nutraukimo sąlyga nerasta

Poaibio generavimo etapą sudaro paieškos organizavimas (angl. *search organization*) ir paveldėtojo (kito poaibio) generavimas (angl. *successor generation*) (Kumar & Minz, 2014).

**Paieškos organizavimas.** Duomenų rinkiniui  $D$ , turinčiam  $N$  požymių, egzistuoja  $2^N$  galimų požymių poaibių. Netgi esant nedideliame  $N$ , paieškos

erdvė eksponentiškai didėja ir tai trukdo atlikti išsamią paiešką, t. y. pilnąjį perrinkimą. Todėl yra pasiūlytos kelios požymių paieškos strategijos, tokios kaip nuosekli paieška (angl. *sequential search*), eksponentinė paieška, atsitiktinė paieška.

**Kitam požymių poaibiui (angl. *successor*) generuoti** gali būti naudojami tokie operatoriai kaip pirmyn (poaibis generuojamas didinant ankstesnįjį), atgal (pradinis poaibis mažinamas), atsitiktinis (požymių poaibis generuojamas dalimi ar visiškai), svertinis ar junginio (poaibių sujungimo būdu).

Naujai sugeneruotas **požymių rinkinys** turi būti **įvertintas** naudojant konkrečius vertinimo kriterijus. Yra pasiūlyti tokie vertinimo kriterijai ( Dash & Liu, 1997), (Liu & Motoda, 1998), (Bhargava & Polzehl, 2012), (Kumar & Minz, 2014):

- atstumo matas, kuris įvertina atstumą tarp klasių požymių arba tikimybinį atstumą tarp požymių sąlyginių tikimybių tankių (Liu & Motoda, 1998);
- informacijos arba neapibrėžtumo matas, paremtas požymių informacijos pokyčiu tarp skirtingų klasių (Liu & Motoda, 1998), (Azhagusundari & Selvadoss Thanamani, 2013);
- klaidų tikimybės matas, minimizuojantis klasifikavimo klaidų tikimybę (Duda, et al., 2000);
- priklausomybės matas (angl. *dependency measures*), dar žinomas kaip panašumo arba koreliacijos matas (Liu & Motoda, 1998), (Asad, et al., 2012), (Bhargava & Polzehl, 2012);
- atstumo tarp klasių matas, iš kurių dažniausiai naudojamas Euklido atstumas (Liu & Motoda, 1998);
- nuoseklumo matas (angl. *consistency measures*), įvertinantis, kuris minimalus požymių rinkinys identifikuoja klasę taip pat kaip ir pilnas požymių rinkinys (Liu & Motoda, 1998).

Reikėtų pabrėžti, kad požymių rinkinys, suformuotas remiantis vienu kriterijumi, gali nesutapti su požymių rinkiniais, gautais naudojant kitus vertinimo kriterijus.

Požymių atrankos procese taip pat turi būti apibrėžtas **stabdomo kriterijus**. Yra pasiūlyti tokie stabdomo kriterijai, kaip iš anksto nustatytas didžiausias iteracijų skaičius, mažiausias požymių skaičius, minimalus klasifikavimo klaidų lygis, aiški paieškos pabaiga. Atrankos stabdomo kriterijumi taip pat gali būti nepakankamas fiksuojamos charakteristikos kitimas – atranka stabdoma, kai generuojant naujus požymių poaibius, nebegaunamas užsibrėžtas klasifikavimo charakteristikos pokytis (Kumar & Minz, 2014). Tokio kriterijaus pavyzdžiu galėtų būti atrankos stabdymas, kai naujų požymių rinkinių lemiamas klasifikavimo tikslumas pakinta mažiau nei užsibrėžta.

## 2.4 Požymių rinkinio formavimo metodai

Dažniausiai yra naudojami tokie požymių atrankos metodai, kaip nuoseklus aibės didinimas, nuoseklus aibės mažinimas, geriausiojo pirmo pasirinkimo metodas, įvairūs genetiniai algoritmai ir kiti (Dash & Liu, 1997), (Liu & Motoda, 1998), (Guyon & Elisseeff, 2003), (Planet & Iriondo, 2012).

Taikant nuoseklus aibės didinimo (angl. *Sequential Forward Selection*) metodą, požymių rinkinys formuojamas nuo pradžių, kiekvienoje iteracijoje pasirenkant požymį, lemiantį didžiausią naujojo požymių rinkinio efektyvumą, t. y. didžiausią emocijų klasifikavimo tikslumą (Marill & Green, 1963), (Guyon & Elisseeff, 2003), (Ververidis, et al., 2004), (Casale & Russo, 2007), (Origlia, et al., 2010). Šio metodo patobulintas variantas (angl. *Sequential Floating Forward Selection*) leidžia po kiekvieno poaibio didinimo žingsnio atlikti poaibio mažinimą – pašalinti poaibio efektyvumą mažinančius požymius (Pudil, et al., 1994), (Lugger, et al., 2009), (Chen, et al., 2012).

Nuoseklus aibės mažinimo (angl. *Sequential Backward Selection*) metodo esmė yra mažinti pradinį (pilnąjį) požymių rinkinį nuosekliai pašalinant tą požymį, be kurio rinkinys tampa efektyvesniu (Whitney, 1971),

(Guyon & Elisseeff, 2003), (Casale & Russo, 2007). Naudojant geriausiojo pirmojo pasirinkimo (angl. *Promising First Selection*) metodą, pirmiausia įvertinamas kiekvieno požymio individualus efektyvumas – klasifikavimo klaida, gaunama naudojant vienintelį nagrinėjamąjį požymį. Požymiai surikiuojami klaidos didėjimo tvarka, o požymių rinkinys formuojamas nuosekliai, pasirenkant geriausią pirmąjį. Galutiniu variantu paskelbiama požymių rinkinio versija, lemianti mažiausią atpažinimo klaidą (Dellaert, et al., 1996). Dar vienas požymių atrankai naudojamas, maksimalaus reikšmingumo ir minimalaus pertekliško, metodas (angl. *Maximum Relevance – Minimum Redundancy*) atrenka požymius, kurie yra reikšmingiausi analizuojamai emocijų klasei. Reikšmingumas apibrėžiamas abipusiu ryšiu tarp dviejų požymių (Guyon & Elisseeff, 2003), (Peng, et al., 2005). Požymių rinkinys gali būti formuojamas ir iš tiesiškai nepriklausomų požymių. Tokiu būdu tikimasi gauti geresnį klasifikavimo rezultatą. Tiesinis požymių nepriklausomumas yra įvertinamas naudojant kryžminę koreliaciją (angl. *Cross-correlation*) (Asad, et al., 2012), (Bhargava & Polzehl, 2012), (Mencattini, et al., 2014). Požymių rinkinys taip pat gali būti formuojamas genetiniais algoritmais generuojant ir optimizuojant naujus požymių rinkinius (Siedlecki & Sklansky, 1989), (Schuller, et al., 2006), (Casale & Russo, 2007), (Origlia, et al., 2010).

Taikant požymių transformavimo metodiką, sudarytieji požymių rinkiniai optimizuojami transformuojant požymių reikšmes ir mažinant rinkinio dydį (požymių kiekį rinkinyje). Tokiu būdu bandoma spręsti „daugiamatiškumo prakeiksmo“ problemą, kuomet dėl didelio požymių kiekio rinkinyje ir nepakankamo duomenų kiekio nepavyksta tinkamai apmokyti klasifikatoriaus (ir tuo pačiu tikėtis aukšto klasifikavimo tikslumo). Šnekos emocijų požymių rinkinio efektyvumui maksimizuoti yra naudojami tokie duomenų transformavimo metodai kaip pagrindinių komponentų analizė (Chiou & Chen, 2013), (Arias, et al., 2014), tiesinė diskriminantinė analizė (You, et al., 2006), daugiamatį skalų metodas (Rong, et al., 2009), Lipšitzo erdvės metodas (angl. *Lipschitz Spacing Method*) (You, et al., 2007), Fišerio

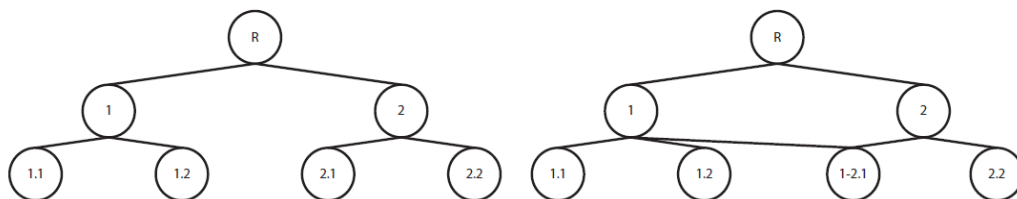
diskriminantinė analizė (Zhang, et al., 2010), apdorojimas neuronų tinklais (Gharavian, et al., 2012), sprendimų medžiai (Rong, et al., 2009) ir kiti.

## 2.5 Hierarchinio klasifikavimo analizė

Duomenų tyrybos, mašininio mokymo ir atpažinimo uždaviniuose dažniausiai koncentruojamasi į vieno etapo sprendimą arba tiesioginiu klasifikavimu (angl. *Flat Classification*) pagrįstą sprendimą. Tokio klasifikavimo metu yra sprendžiamos binarinio arba aukštesnės eilės skaidymo į klases uždavinys. Esminis tokio klasifikavimo argumentas – paprastumas: uždavinys sprendžiamas vieno klasifikatoriaus pagalba, neretai naudojant pilną požymių rinkinį. Tokios sistemos efektyvumo didinimas galimas tik klasifikatoriaus arba naudojamo požymių rinkinio pakeitimu (tai iš esmės reiškia naują klasifikavimo sistemos sudarymą).

Kita vertus, realiame pasaulyje egzistuoja nemažai pavyzdžių, kur nagrinėjama hierarchinio klasifikavimo problema. Hierarchiniame klasifikavime skirstymas į klases vyksta keliais etapais, kiekviename etape išskiriant klasę ar abstrahuotą jų grupę. Tokia klasifikavimo schema reiškia keleto klasifikatorių naudojimą ir netgi skirtingų požymių poabių taikymą kiekvienam klasifikatoriui.

Hierarchiniame klasifikavime klasės yra organizuotos į klasių hierarchiją. Hierarchijai atvaizduoti gali būti panaudota medžio struktūra arba kryptinis aciklinis grafas (KAG). Esminis skirtumas tarp šių hierarchinių struktūrų tipų yra tai, kad KAG struktūroje mazgas gali turėti daugiau nei vieną tėvinį mazgą (2 pav.) (Costa, et al., 2007), (Silla Jr. & Freitas, 2011).



2 pav. Paprasti medžio struktūros (kairė) ir KAG struktūros (dešinė) pavyzdžiai (Silla Jr. & Freitas, 2011)

Hierarchinis klasifikavimas gali būti atliekamas dvejopai. Pirmasis būdas – visada klasifikuoti naujai gautus mazgus – vadinamas privalomąja prognoze. Kitas būdas numato galimybę sustabdyti klasifikavimą bet kuriame mazge ir bet kuriame lygyje ir yra vadinamas neprivalomąja prognoze.

### **2.5.1 Klasifikavimo būdai**

Literatūroje yra siūlomi tokie trys klasifikavimo būdai (Costa, et al., 2007), (Silla Jr. & Freitas, 2011), (Hernandez, et al., 2014):

1. tiesioginis (ang. *flat*);
2. lokalus;
3. globalus.

Lokalus klasifikavimas dar yra vadinamas „iš viršaus į apačią“, kai kiekvienai hierarchinei klasei (ar lygmeniui) naudojami lokalūs klasifikatoriai, o hierarchija organizuota skaidymo kryptimi (nuo klasių grupių link pavienių klasių). Globalaus klasifikavimo atveju naudojamas tik vienas klasifikatorius visoms hierarchijos klasėms, o patį klasifikavimo procesą tenka organizuoti etapais (pirmiau atskiriama viena klasė ar jų junginys, po to – kita ir t. t.). Klasifikuojant tiesioginiu būdu, sąryšiai tarp klasių yra ignoruojami.

#### **2.5.1.1 Tiesioginis klasifikavimas**

Tiesioginis klasifikavimo uždavinio sprendimas yra pats paprasčiausias būdas išspręsti hierarchinio klasifikavimo problemą. Klasių hierarchija yra tiesiog ignoruojama ir vieno klasifikavimo (įgyvendinamas tik vienas klasifikavimo lygis) pagrindu yra nurodomos visos klasės. Tiesioginis įgyvendinimas dažniausiai naudojamas kaip tradicinis klasifikavimo principas apmokymo ir testavimo metu. Tačiau esant klasių hierarchijai, toks klasifikavimas neduoda gerų rezultatų. Šis labai paprastas klasifikavimo būdas turi labai rimtą trūkumą – šiai schemai tenka sukurti klasifikatorių, gebantį atskirti visas klases nenagrinėjant klasių hierarchijoje esančios tėvų-vaikų klasių sąryšio informacijos (Silla Jr. & Freitas, 2011).

Tokio klasifikavimo privalumas yra jo paprastumas, o trūkumas – klasių hierarchijos ignoravimas.

### **2.5.1.2 Lokalus klasifikavimas**

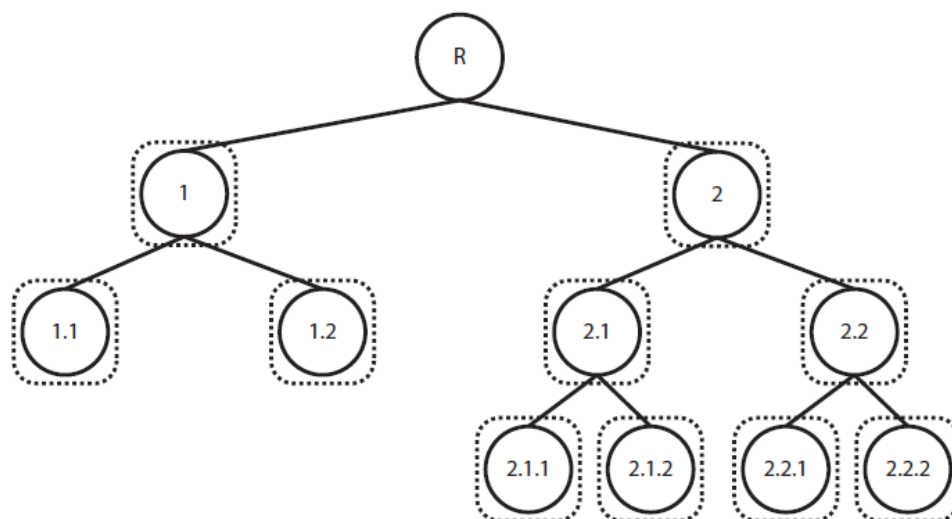
Pirmasis lokalus klasifikatorius buvo pasiūlytas 1997 metais (Koller & Sahami, 1997). Svarbus šio klasifikavimo aspektas yra tai, kad hierarchijoje yra atsižvelgiama į lokalią informacijos naudojimo perspektyvą. Egzistuoja trys įprastiniai būdai naudoti lokalią informaciją hierarchiniame klasifikavime:

- lokalus klasifikavimas per mazgą (angl. *per node*);
- lokalus klasifikavimas per tėvinį mazgą;
- lokalus klasifikavimas per lygį.

Reikėtų paminėti, kad išvardyti trys hierarchinio klasifikavimo būdai atsiskiria mokymo etape, o testavimo etape supanašėja ir pasižymi labai panašiu „iš viršaus į apačią“ principu. Remiantis „iš viršaus į apačią“ principu, pirmiausia yra numatoma aukščiausio lygio (bendriausia) klasė. Paskui kiekviena nagrinėjama klasė susiaurina galimus žemesnio lygio klasių pasirinkimus, pavyzdžiui, antrojo lygio klasės kandidatės yra pirmojo lygio klasės „vaikai“. Tokiu principu kiekviename lygyje klasės yra skaidomos, kol gaunama vertinamos klasės prognozė. Tokioje klasifikavimo schemoje bet kurio klasės lygio klaida tolimesnėje hierarchinėje analizėje dauginasi iš žemesnio lygio klaidų (Silla Jr. & Freitas, 2011).

#### **Lokalus klasifikavimas per mazgą**

Tai dažniausiai literatūroje minimas klasifikavimo metodas. Lokalaus klasifikavimo per mazgą principas apima vieno binarinio klasifikatoriaus apmokymą kiekvienam klasių hierarchijos mazgui (įskaitant šakninį mazgą) (3 pav.) (Silla Jr. & Freitas, 2011).



**3 pav. Lokalus klasifikavimas per mazgą (apskritimai reprezentuoja klases, o punktyriniai kvadratai, apvalintais kampais, reprezentuoja binarinius klasifikatorius) (Silla Jr. & Freitas, 2011)**

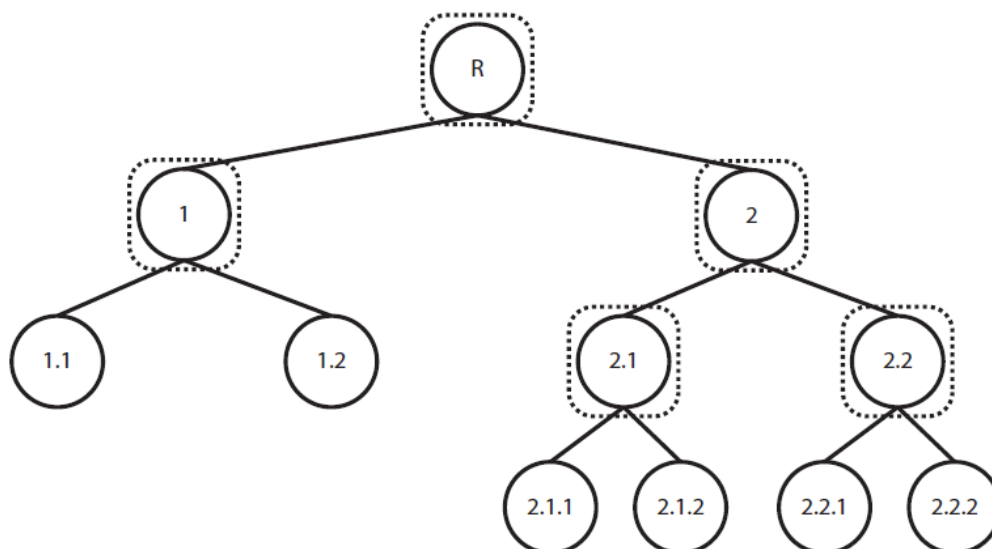
Tokio klasifikavimo privalumas yra jo paprastumas ir atsižvelgimas į klasių hierarchiją, o trūkumai – didelis naudojamų klasifikatorių skaičius ir polinkis į nenuoseklumą.

### **Lokalus klasifikavimas per tėvinį mazgą**

Tai yra dar vienas klasifikavimo būdas, kuriame yra naudojama lokali informacija. Kiekvienam tėviniam mazgui klasių hierarchijoje yra apmokomas klasifikatorius, leidžiantis atskirti jo dukterinius mazgus (Silla Jr. & Freitas, 2011).

Sakykime, testavimo etape pirmojo lygio klasifikatorius priskiria pavyzdį prie klasės „2“. Tada, antrojo lygio klasifikatorius, kuris buvo apmokytas tik su 2 klasės mazgo vaikais, t. y. su 2.1 ir 2.2, atlieka priskyrimą savo klasėms. Procesas kartojamas tol, kol gaunamas galutinis sprendimas (4 pav.).





**4 pav. Lokalus klasifikavimas per tėvinį mazgą (apskritimai reprezentuoja klases, o punktyriniai kvadratai, apvalintais kampais, reprezentuoja daugiaklasius klasifikatorius – prognozuoja jų vaikų klases) (Silla Jr. & Freitas, 2011)**

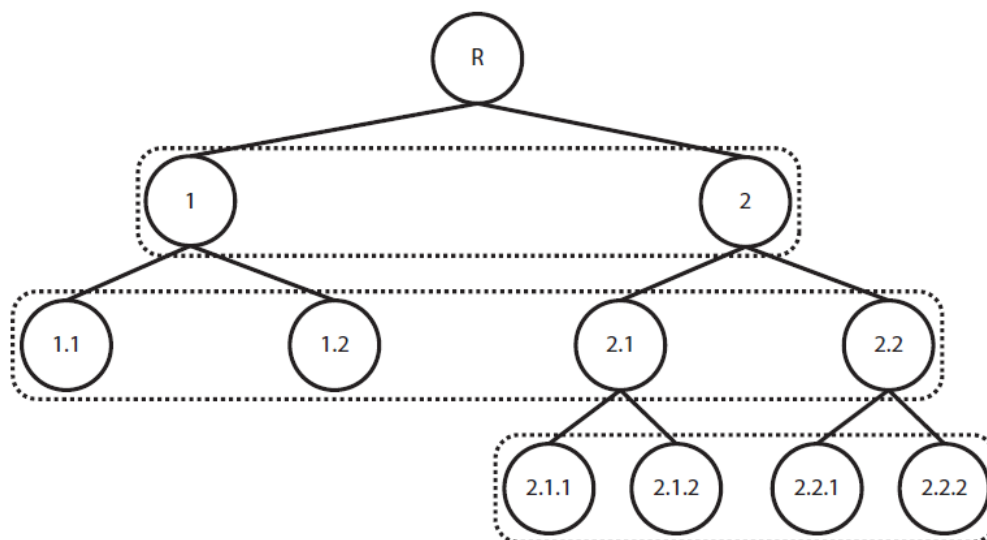
Tokio klasifikavimo privalumas yra jo paprastumas, mažesnis klasifikatorių skaičius, nei taikant lokalaus klasifikavimo per mazgą požiūrį, o trūkumas – polinkis į nenuoseklumą.

### **Lokalus klasifikavimas per lygį**

Tai rečiausiai naudojamas klasifikavimo būdas. Šio metodo esmė yra vieno klasifikatoriaus apmokymas kiekvienam lygiui klasių hierarchijoje (Silla Jr. & Freitas, 2011).

5 pav. pavaizduotoje klasifikavimo schemeje yra naudojami trys klasifikatoriai – po vieną klasifikatorių kiekvienam klasių hierarchijos lygiui.

Tokio klasifikavimo privalumas yra jo paprastumas, mažesnis klasifikatorių skaičius, nei taikant lokalaus klasifikavimo per mazgą ar lokalaus klasifikavimo per tėvinį mazgą požiūrius, o trūkumai – polinkis į nenuoseklumą, didelis klasių skaičius vienam klasifikatoriui, „tėvas-vaikas“ sąryšio tarp klasių ignoravimas apmokymo metu.



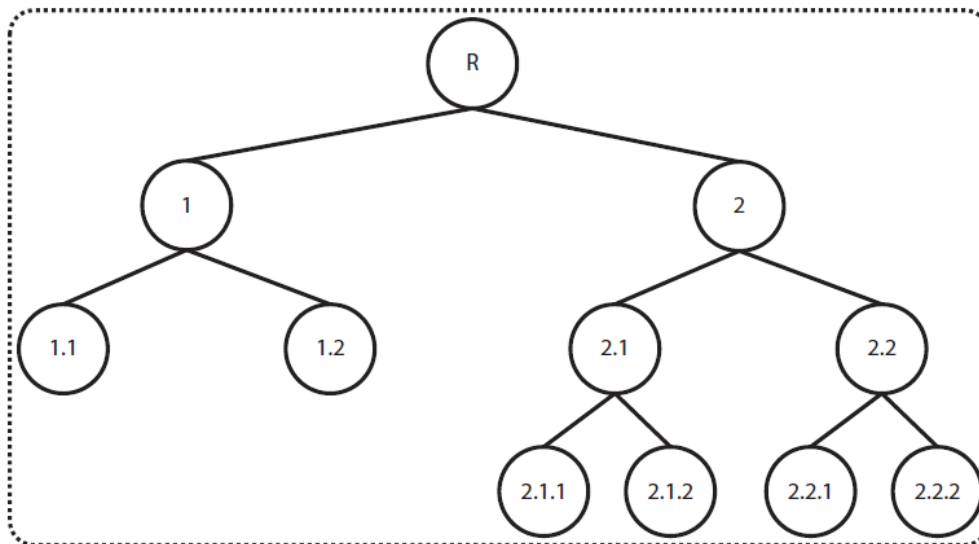
5 pav. Lokalus klasifikavimas per lygį (apskritimai reprezentuoja klases ir kiekvienas punktyrinis stačiakampis, apvalintais kampais, reprezentuoja daugiaklasio klasifikatoriaus prognozuojamas klases) (Silla Jr. & Freitas, 2011)

### 2.5.1.3 Globalus klasifikavimas

Hierarchinio klasifikavimo problemą galima išspręsti apmokant vieną globalų modelį visoms klasėms. Toks požiūris į klasifikavimą yra žinomas kaip „globalus“ mokymas ir kartais įvardijamas „didžiuoju sprogitu“ (angl. *big-bang*).

Naudojant globalaus klasifikavimo principą, vienas klasifikavimo modelis (klasifikatorius) yra naudojamas viso klasifikavimo metu (visoms hierarchijos klasėms).

Literatūroje globalus klasifikavimas kartais yra klaidingai vadinamas lokaliuoju klasifikavimu „iš viršaus į apačią“ (Xiao, et al., 2007), (Xiao, et al., 2009), (Silla Jr. & Freitas, 2011). Globaliai klasifikuojant vyksta vienas vieno klasifikatoriaus apmokymo etapas visoms klasėms (6 pav.).



6 pav. „Didžiojo sprogimo“ klasifikavimo požiūris naudojant klasifikavimo algoritmą visam globaliam klasifikavimo modeliui visoje klasių hierarchijoje (Silla Jr. & Freitas, 2011)

Tokio klasifikavimo privalumu laikoma tai, kad išsaugomi natūralūs sąryšių tarp klasių apribojimai, atsižvelgiama į klasių hierarchiją klasifikatoriaus apmokymo bei testavimo metu, įgyvendinamas vieno (tačiau kompleksinio) sprendimo modelis. Tokios schemos trūkumas – visas klasifikavimo procesas yra priklausomas nuo vieno klasifikatoriaus (angl. *classifier-specific*).

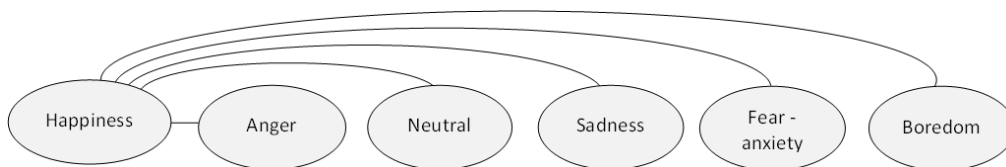
Eksperimentų rezultatai atskleidžia lokalaus bei globalaus klasifikavimų pranašumą prieš tiesioginį klasifikavimą (Silla Jr. & Freitas, 2011). Pavyzdžiui, tiesioginis ir lokalusis hierarchinis klasifikavimas buvo naudojami emocionalios šnekos įrašams klasifikuoti į šešias emocijų klases. Lokalusis trijų lygių klasifikavimas per tėvinį mazgą davė 3 proc. geresnį rezultatą nei tiesioginis klasifikavimas (Xiao, et al., 2007), (Xiao, et al., 2009).

## 2.6 Emocijų klasifikavimo hierarchinių schemų apžvalga

Pats paprasčiausias būdas atlikti emocionalios šnekos įrašų klasifikavimą – vieno etapo arba tiesioginis klasifikavimas. Įvairios hierarchinės emocijų klasifikavimo schemas yra pasiūlytos kaip alternatyva vieno etapo klasifikavimui, siekiant pagerinti emocijų atpažinimo tikslumą.

Pasiūlytuose metoduose klasifikatoriai yra kombinuojami hierarchinių būdu, darant prielaidą apie emocijų tarpusavio ryšius.

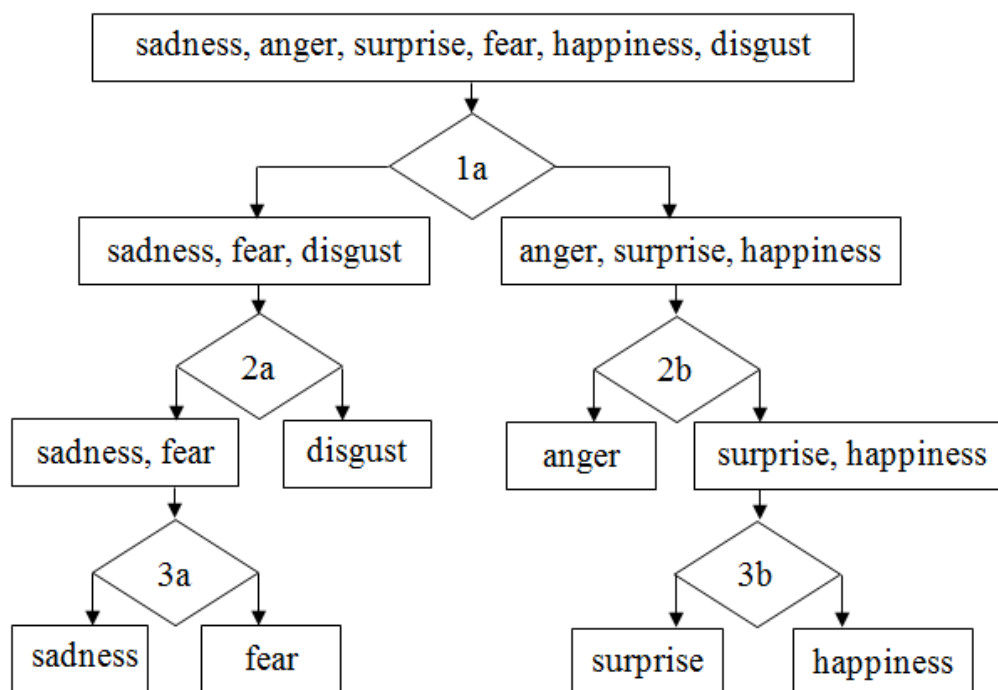
Hierarchinė posistemių struktūra yra pasiūlyta emocijų analizei poromis (Giannoulis & Potamianos, 2012). Eksperimentui atlikti buvo pasirinkta Berlyno emocionalios šnekos duomenų bazė. Šešios emocijos (pyktis, džiaugsmas, baimė, liūdesys, nuobodulys ir neutrali būseną) buvo sugrupuotos į 15 skirtingų emocijų porų (7 pav.). 15 atskirų posistemių buvo apmokyta emocijų atpažinimui poroje, naudojant kiekvienai porai konkretų ir vis kitokį požymių rinkinį. Tokie požymių rinkiniai gaunami iš bendro požymių rinkinio, kurį sudaro 112 skirtingų tipų požymių. Pavyzdžiui, džiaugsmui atpažinti buvo analizuojamos tokios poros: džiaugsmas / pyktis, džiaugsmas / neutrali būseną, džiaugsmas / liūdesys, džiaugsmas / baimė, džiaugsmas / nuobodulys. Kiekvienai porai buvo apmokamas SVM klasifikatorius, kurio užduotis buvo identifikuoti vieną iš dviejų nagrinėjamų emocijų. Galutinis klasifikavimo rezultatas buvo gautas naudojant balsų daugumos principą atsižvelgiant į posistemių rezultatus. Bendras emocijų atpažinimo tikslumas siekė 85,2 proc. atsižvelgiant į kalbančiojo lytį (eksperimento sąlygos – priklausomybė nuo kalbančiojo lyties) ir 80,1 proc. nuo lyties nepriklausančius eksperimentus.



7 pav. Penkios posistemės „happiness“ atpažinimui (Giannoulis & Potamianos, 2012)

Trijų etapų klasifikavimo schema buvo pasiūlyta šešioms emocijoms (liūdesys, pyktis, nuostaba, baimė, džiaugsmas ir neutrali būseną) identifikuoti (Chen, et al., 2012). Tuo tikslu buvo panaudoti penki klasifikatoriai emocijų poroms išanalizuoti. Emocijų klasės buvo sudaromos taikant Fišerio kriterijų emocijų požymių reikšmėms (8 pav.). Pradinį požymių rinkinį kiekvienam etapui sudarė 288 požymiai. Pirmajame etape visos emocijos buvo

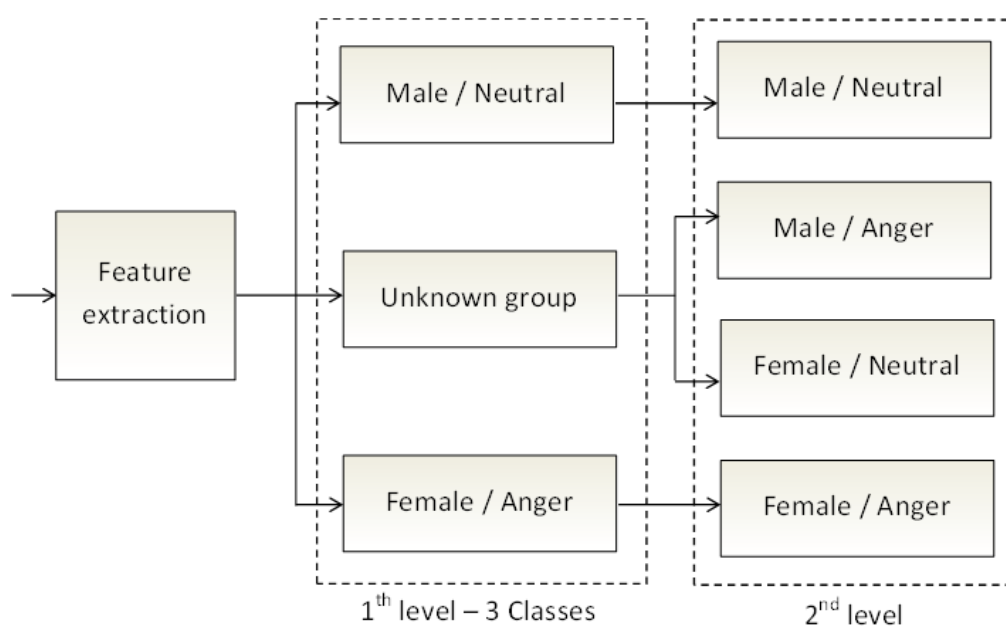
klasifikuojamos į dvi klases naudojant klasifikatorių 1a. Liūdesys, baimė ir išgąstis priklausė pirmajai klasei, o pyktis, nuostaba ir džiaugsmas priklausė antrajai klasei. Antrajame etape visos pirmosios klasės emocijos buvo klasifikuojamos į dvi grupes naudojant 2a ir 2b klasifikatorius. Liūdesys ir baimė priklausė pirmajai grupei, o išgąstis priklausė kitai grupei. Visos antrosios klasės emocijos buvo klasifikuojamos taip pat į dvi grupes. Pirmajai grupei priklausė nuostaba ir džiaugsmas, o antrajai – pyktis. Trečiajame etape likusios emocijų poros buvo skirstomos į atskiras emocijas naudojant 3a ir 3b klasifikatorius. Emocijų atpažinimo vidurkis kiekviename etape sudarė atitinkamai 86,5 proc., 68,5 proc. ir 50 proc. Bendras nagrinėjamų požymių skaičius – 288.



8 pav. Trijų lygių klasifikavimo schema (Chen, et al., 2012)

Dviejų etapų klasifikavimo į dvi klases (pyktis ir neutrali būsena) schema remiasi akustiniais vyrų ir moterų balsų skirtumais (Yoon & Park, 2011). Pirmajame etape visi emocionalios šnekos įrašai buvo klasifikuojami naudojant pagrindinio tono vidurkio reikšmę ir kalbančiųjų lyčių klasifikavimo algoritmą

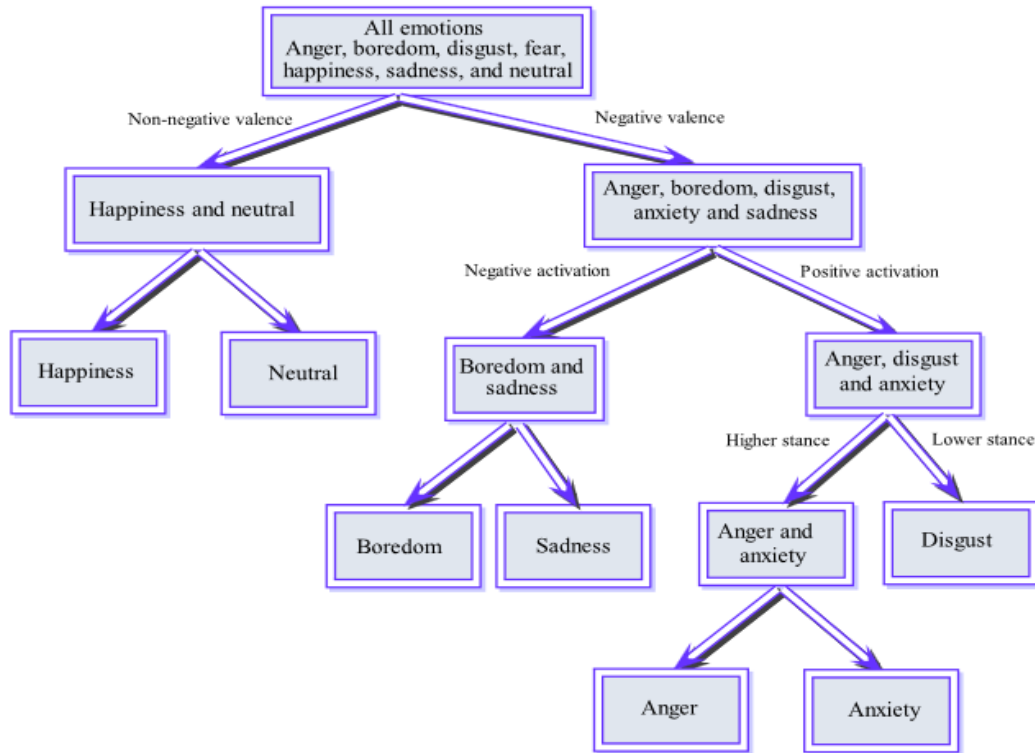
į tris emocijų grupes: vyrai arba neutrali būsena, moterys arba neutrali būsena ir nežinomų emocijų grupė. Būtent tokioms trimis grupėms buvo gautos skirtingos pagrindinio tono reikšmės. Antrajame klasifikavimo etape nenustatytų emocijų grupė buvo klasifikuojama į dvi emocijų grupes: vyrai arba pyktis ir moterys arba neutrali būsena (9 pav.). Skirtingi požymių rinkiniai buvo naudojami kiekviename klasifikavimo etape. Bendrą eksperimentui naudojamą požymių vektorių sudarė 56 požymiai. Emocijų atpažinimo vidurkis siekė 80,7 proc.



9 pav. Dviejų etapų hierarchinio klasifikavimo į dvi klases schema (Yoon & Park, 2011)

Vyrų ir moterų skirtumo informacija taip pat buvo naudojama sustiprinto bendro mokymo algoritmui (angl. *Enhanced Co-training Algorithm*). Du skirtingi požymių rinkiniai ir du klasifikatoriai buvo naudojami šešioms emocijoms (pyktis, baimė, džiaugsmas, liūdesys, nuostaba ir neutrali būsena) atskirti (Liu, et al., 2007). Šiame algoritme vyrų ir moterų šnekos įrašus pasiūlyta klasifikuoti atskirai. Jeigu abu klasifikatoriai kurį nors pavyzdį identifikuodavo vienodai, jis buvo perkeliamas į aibę, kuri būdavo klasifikuojama pakartotinai. Galutinis sprendimas apie šnekančiojo emociją būklę gaunamas sujungiant abiejų klasifikatorių rezultatus. Autorių

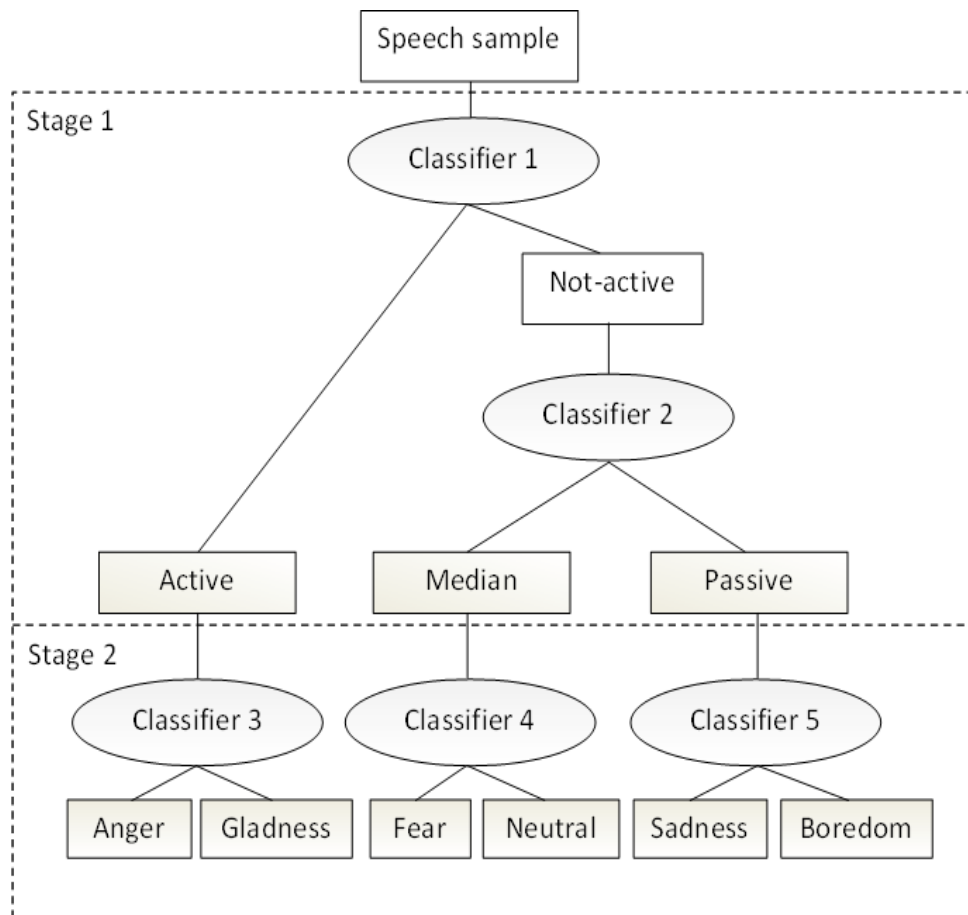
deklaruojamų emocijų atpažinimo tikslumo vidurkis siekė 75,9 proc. moterims ir 80,9 proc. vyrams.



10 pav. Hierarchinė binarinė klasifikavimo schema (Kotti & Paterno, 2012)

Dar viena hierarchinė binarinė klasifikavimo schema remiasi psichologų pasiūlytu trimatės emociinės erdvės modeliu, kuriame emocijos aprašomos trimatėje erdvėje: valentingumas – aktyvacija – pozicija (angl. *valence, activation, and stance*) (Kotti & Paterno, 2012). Pirmajame etape visos septynios emocijos yra klasifikuojamos į dvi emocijų klases: teigiamas valentingumas (priskiriamas džiaugsmas ir neutrali būseną) ir neigiamas valentingumas (priskiriamas pyktis, nuobodulys, šleikštulys, nerimas ir liūdesys). Antrajame etape visos teigiamo valentingumo grupės emocijos yra suskirstomos į dvi pavienes emocijas: džiaugsmas ir neutrali būseną. Visos neigiamo valentingumo grupės emocijos yra toliau klasifikuojamos į kitas dvi grupes: neigiamos aktyvacijos (nuobodulys ir liūdesys) ir teigiamos aktyvacijos (pyktis, šleikštulys ir nerimas). Neigiamos aktyvacijos grupė

trečiajame etape yra skirstoma į dvi pavienes emocijas: nuobodulys ir liūdesys, o teigiamos aktyvacijos grupė yra klasifikuojama į dvi klases: žemesnės pozicijos (šleikštulys) ir aukštesnės pozicijos (pyktis ir nerimas) (10 pav.). Skelbiamas tokios klasifikavimo schemas emocijų atpažinimo tikslumo vidurkis siekė 87,7 proc., naudojant 75 požymių vektorių.



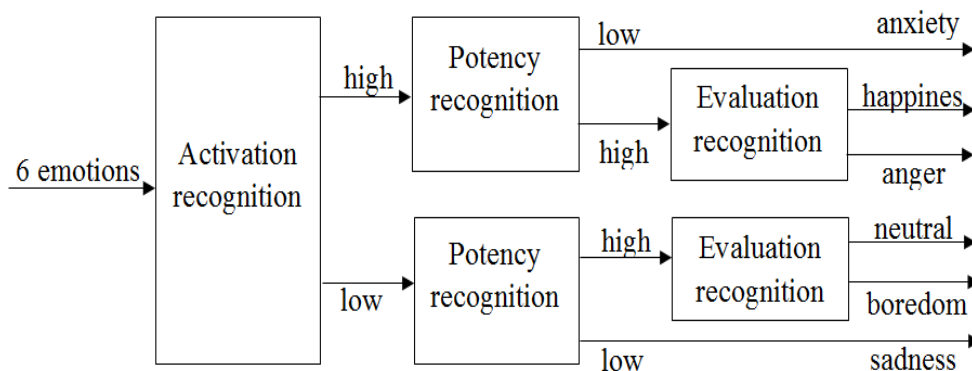
11 pav. Daugiaetapė (kelių etapų) klasifikavimo schema (Xiao, et al., 2009)

Kitas hierarchinis požiūris į emocijų atpažinimą šnekoje yra pagrįstas kelių etapų klasifikavimo schema (angl. *Multistage Classification Scheme*) (Xiao, et al., 2009). Pirmajame etape emocijos yra klasifikuojamos į tris grupes pagal emocinio pakilimo lygį (angl. *arousal dimension*): aktyvios, vidutinės, pasyvios. Antrajame etape, naudojant skirtingus klasifikatorius, kiekviena emocijų grupė yra klasifikuojama į dvi konkrečias emocijų grupes: aktyvių emocijų grupei priskiriamas pyktis ir džiaugsmas, vidutinių emocijų grupei –



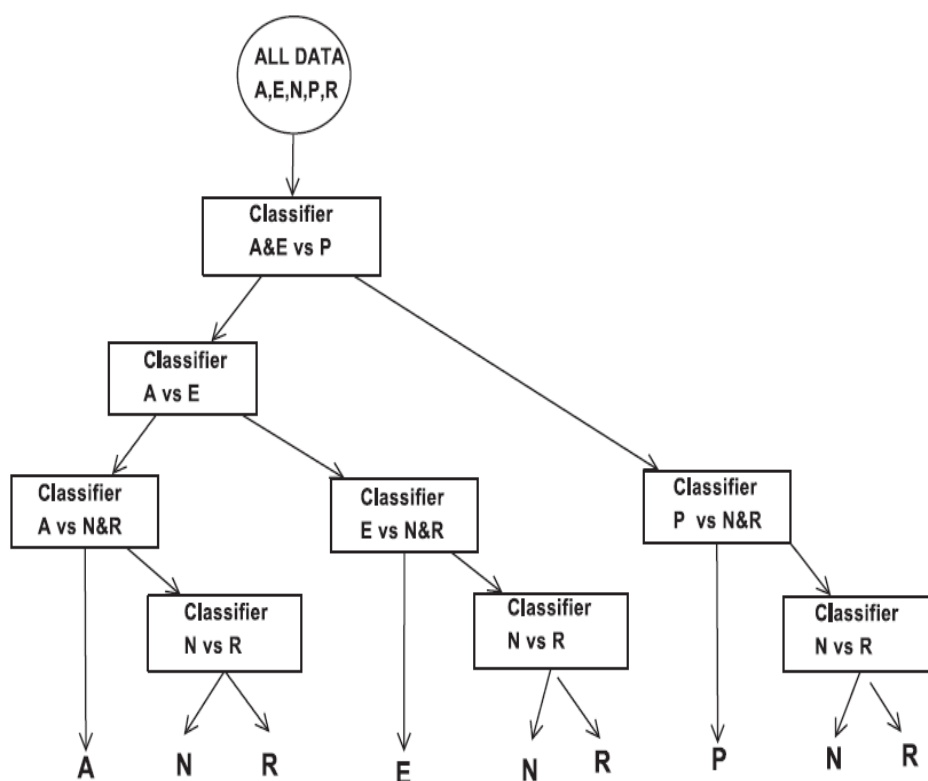
baimė ir neutrali būseną, pasyvių emocijų grupei – liūdesys ir nuobodulys (11 pav.). Kiekviename klasifikavimo etape yra naudojami skirtingi požymiai ir binariniai klasifikatoriai. Požymių rinkiniai gaunami naudojant SFS požymių atrankos metodą (angl. *Sequential Forward Selection*). Eksperimente panaudoti 68 požymiai, atpažinimo tikslumas siekė 76,4 proc.

Dar vienas hierarchinis emocijų klasifikavimo būdas yra pagrįstas trijų dimensijų emocijų modeliu (angl. *3 dimensions emotions model*): aktyvacija (angl. *activation*), stiprumas (angl. *potency*) ir vertinimas (angl. *evaluation*) (Lugger, et al., 2009 ). Šešios emocijos (pyktis, džiaugsmas, nerimas, liūdesys, nuobodulys ir neutrali būseną) yra klasifikuojamos trimis lygiais, naudojant binarinį Bajeso klasifikatorių (12 pav.). Požymių atrankai buvo naudojamas nuoseklaus svyruojančio aibės didinimo (angl. *Sequential Floating Forward Selection*) algoritmas. Kiekviename klasifikavimo lygyje naudojami požymiai yra apribotai emocijų modelio. Pirmajame ir antrajame klasifikavimo lygiuose visos emocijos yra klasifikuojamos į aukštos aktyvacijos ir žemos aktyvacijos klases. Trečiajame lygyje kiekvienos klasės įrašai yra klasifikuojami į atskiras pavienes emocijas. Klasifikuojant emocijas pagal šią klasifikavimo schemą, buvo gautas 88,8 proc. emocijų atpažinimo tikslumas optimizuotam požymių rinkiniui.



12 pav. Hierarchinis emocijų klasifikavimas (Lugger, et al., 2009 )

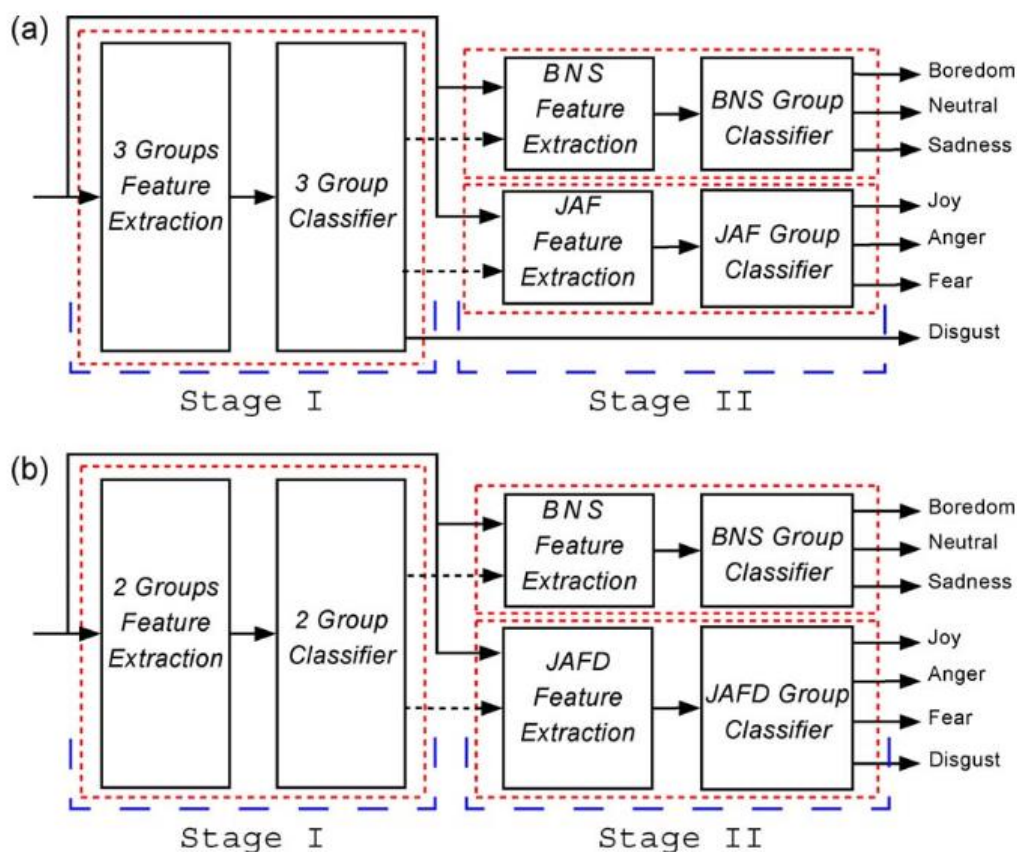
Penkios emocijos yra klasifikuojamos naudojant dar vieną hierarchinio binarinio klasifikavimo schemą (Lee, et al., 2011). Klasifikavimas organizuotas hierarchinių medžių principu, padalijant penkių klasių klasifikavimo uždavinį į šešis binarius klasifikavimo uždavinius (13 pav.). Emocijų klasės sudaromos remiantis Emocijų vertinimo teorija (angl. *Appraisal Theory of Emotions*). Eksperimente nagrinėjamos tokios penkios emocijų klasės: A – piktas, E – jausmingas, P – pozityvus, N – neutralus ir R – atsipalaidavęs. Eksperimentai buvo atlikti naudojant AIBO vaikų emocionalios šnekos duomenų bazę. Vidutiniškai vienos klasės emocijų atpažinimo tikslumas sudaro 48,27 proc. Šie rezultatai turėtų būti vertinami itin atsargiai, nes vaikų šnekos emocijų atpažinimo klausimas yra itin sudėtingas ir plačiau nenagrinėtas.



13 pav. Hierarchinė binarinė klasifikavimo schema (Lee, et al., 2011)

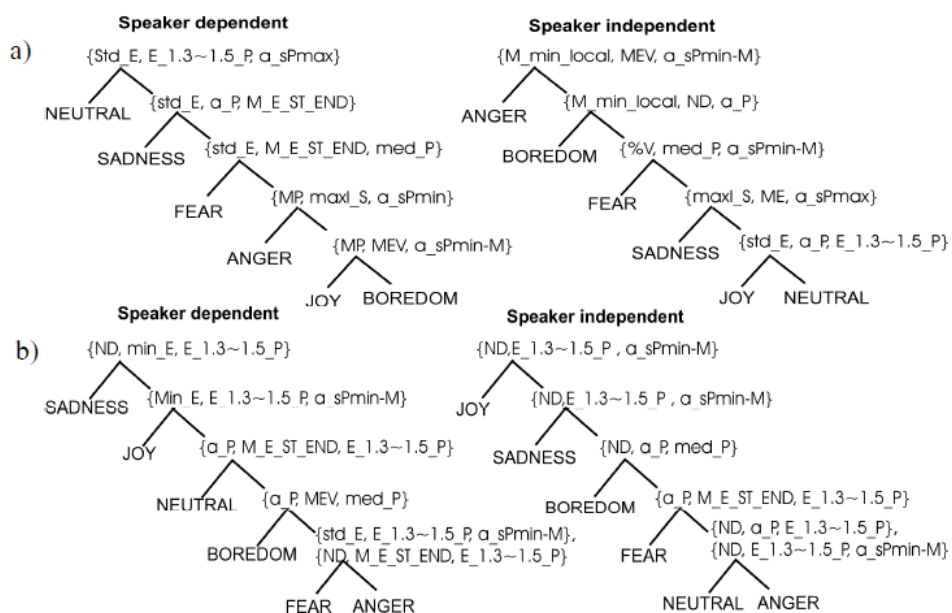
Dviejų etapų hierarchinė klasifikavimo schema buvo pasiūlyta septynioms emocijoms (nuobodulys, liūdesys, džiaugsmas, pyktis, baimė,

šleikštulys ir neutrali būseną) klasifikuoti (Albornoz, et al., 2011). Tyrime nagrinėjami du klasifikavimo schemas variantai. Pirmojoje stadijoje visos emocijos yra skirstomos į tris grupes (pirmas variantas) (14 pav. (a)). arba į dvi grupes (antras variantas) (14 pav. (b)). Tam naudojami spektro ir prozodijos požymiai bei SOM (angl. *Self-Organizing Map*) tipo neuronų tinklas. Antrojoje stadijoje kiekvienos grupės įrašai yra klasifikuojami į atskiras emocijas (14 pav.). Kiekvienos grupės įrašams klasifikuoti naudojami skirtingi požymių rinkiniai. Toks hierarchinis klasifikavimo būdas davė 3,18 proc. geresnį klasifikavimo rezultatą, palyginti su vieno etapo klasifikavimu. Bendras emocijų klasifikavimo tikslumo vidurkis naudojant pasiūlytą hierarchinę schemą siekė 71,5 proc.



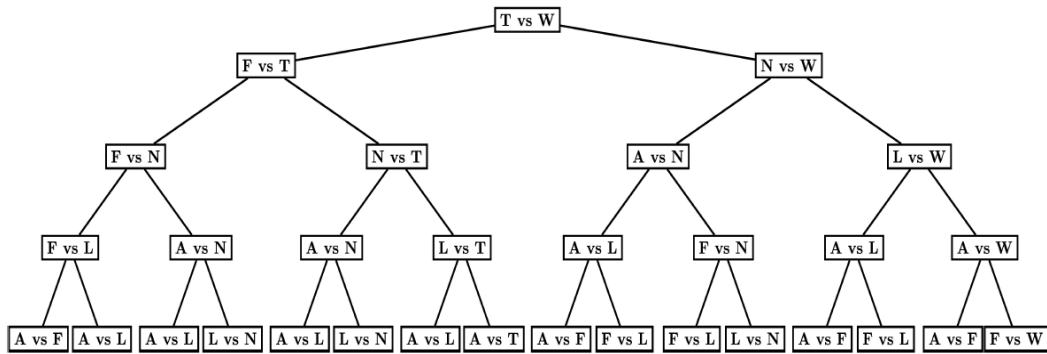
14 pav. Dviejų etapų hierarchinė klasifikavimo schema: (a) emocijų klasifikavimas į tris grupes pirmojoje stadijoje; (b) emocijų klasifikavimas į dvi grupes pirmojoje stadijoje (Albornoz, et al., 2011)

Binariniais sprendimų medžiais paremtas klasifikatorius pasiūlytas šešioms emocijoms klasifikuoti (Cichosz & Slot, 2007). Emocijos yra išskiriamos nuosekliai kiekviename mazge (15 pav.), t. y. kiekviename medžio mazge yra identifikuojama konkreti emocija. Sprendimas priimamas kiekviename mazge. Emocijų atpažinimo vidurkis siekė 72 proc.



15 pav. Binariniais sprendimų medžiais paremtas klasifikatorius: (a) Lenkų emocionalios šnekos duomenų bazė; (b) Berlyno emocionalios šnekos duomenų bazė (Cichosz & Slot, 2007)

Taip pat yra pasiūlyta dar viena binariniais sprendimų medžiais paremta hierarchinė emocionalios šnekos klasifikavimo schema (16 pav.), skirta 7 emocijoms (džiaugsmas (F), pyktis (W), nuobodulys (L), liūdesys (T), šleikštulys (E), išgąstis (A) ir neutrali būseną (N)) atpažinti (Garg, et al., 2013). Deklaruotas emocijų teisingo klasifikavimo vidurkis – 83 proc.



16 pav. Binariniais sprendimo medžiais paremta hierarchinė emocionalios šnekos klasifikavimo schema (Garg, et al., 2013)

Apibendrinant galima teigti, kad hierarchinis emocijų klasifikavimo organizavimas leidžia pasiekti nuo 50 proc. iki 88 proc. vidutinį tikslumą. Jis yra pranašesnis už vieno etapo klasifikavimą ir lemia geresnį emocijų atpažinimo rezultatą. Emocijų atpažinimo tikslumą taip pat pagerina papildomos informacijos, tokios kaip vyrų bei moterų balsų skirtumai bei psichologiniai emocijų aspektai, naudojimas.

## 2.7 Antro skyriaus apibendrinimas

Atlikus emocijų klasifikavimo uždavinio analizę, hierarchinio klasifikavimo apžvalgą bei hierarchinių klasifikavimo schemų analizę, apibendriname:

- Šnekančiojo emocinė būseną daro įtaką šnekos signalo generavimo procesui. Yra nemažai fiziologinių aspektų, kurie būdingi skirtingoms emocinėms būsenoms. Šnekos generavimo procesui įtaką daro fiziniai, nesąmoningi ir sąmoningi stresoriai, kurie atlieka skirtingus vaidmenis šnekos generavimo metu. Emocionali šneka daro įtaką tam tikroms šnekos signalo charakteristikoms. Deja, detalios emocinės būklės daromos įtakos akustinėms balso savybėms analizės nėra atlikta.

- Emocionaliai šnekai atpažinti yra naudojami įvairūs požymiai bei jų rinkiniai. Eksperimentiškai nagrinėtų požymių rinkinių dydis kai kuriais atvejais siekia keletą tūkstančių požymių. Visgi nėra gauta rezultatų, liudijančių fiksuoto požymių rinkinio egzistavimą emocijoms šnekos signale identifikuoti. Dažniausiai šnekos emocijoms nagrinėti yra naudojami prozodiniai ir spektro požymiai.
- Greta tiesioginio klasifikavimo principo šnekos emocijoms klasifikuoti pasiūlytas ir hierarchinis principas. Hierarchinė šnekos emocijų analizė leidžia padidinti emocijų klasifikavimo tikslumą (palyginti su vieno etapo emocijų klasifikavimu), sumažinti naudojamų požymių skaičių.
- Kai kuriais atvejais emocijų atpažinimo tikslumui padidinti yra naudojama papildoma informacija (vyrų ir moterų balsų skirtumas, psichologiniai emocijų aspektai ir kt.) bei skirstymas į tarpines klases.
- Hierarchinis klasifikavimas gali būti atliekamas lokaliai klasifikuojant per mazgą, per tėvinį mazgą arba lokaliai per lygį. Lokalus bei globalus hierarchinis klasifikavimas turi pranašumą prieš tiesioginį (vieno etapo) klasifikavimą.

## 3. Hierarchinė emocijų klasifikavimo schema

### 3.1 Sprendimų medžiai

Klasifikavimo uždaviniui išspręsti yra naudojami įvairūs modeliai: sprendimų medžiai, dirbtinių neuronų tinklai, Bajeso tinklai ir kiti (Rokach & Maimon, 2014).

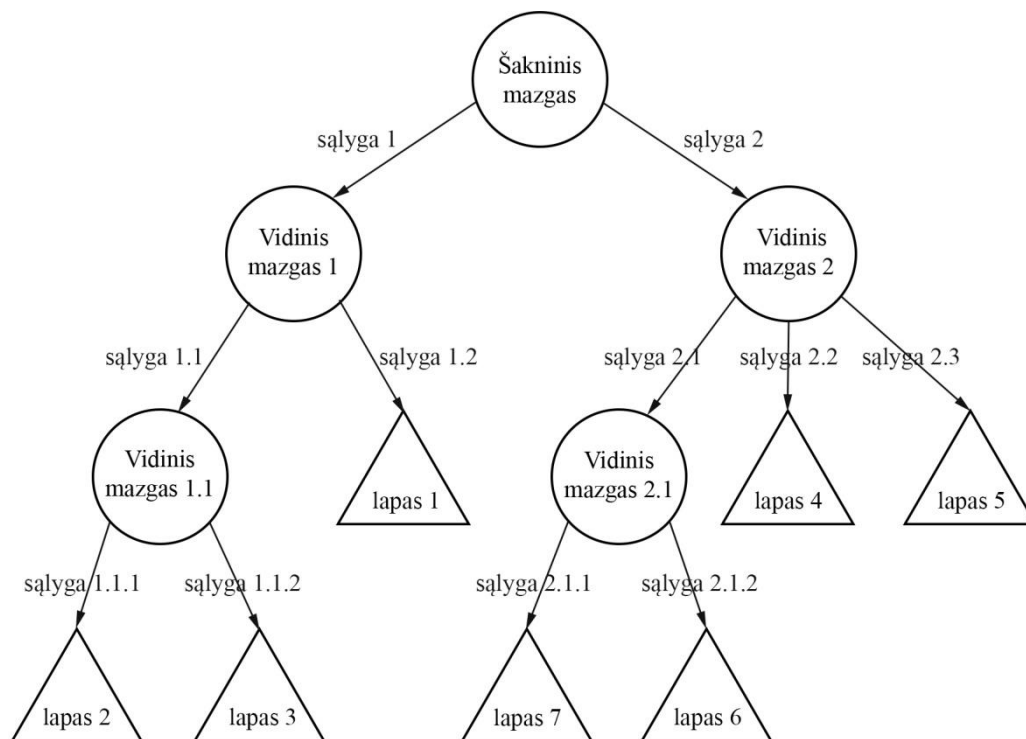
Sprendimų medžiai yra daugiapakopės (angl. *multistage*) sprendimų sistemos, kuriose klasės yra nuosekliai atmetamos tol, kol yra gaunamas galutinis sprendimas apie pavyzdžio klasę (Theodoridis & Koutroumbas, 2008). Klasifikavimui taikant sprendimų medžius yra gaunamas klasifikavimo medis (Breiman, et al., 1984).

Naudojantis klasifikavimo medžiais, pavyzdžiai skirstomi į iš anksto žinomą klasių rinkinį, atsižvelgiant į atributų (požymių, savybių) reikšmes. Sprendimų seka yra taikoma kiekvienam požymiui individualiai atsakant į klausimus „ar požymis  $x_i \leq \alpha$ ?“, kur  $\alpha$  yra iš anksto apibrėžta ribinė reikšmė. Kai kuriais atvejais pavyzdžio priklausomybę konkrečiai klasei yra įmanoma priimti nepanaudojus visų turimų požymių (Theodoridis & Koutroumbas, 2008).

Paprastai klasifikavimo medžiai grafiškai yra pateikiami kaip hierarchinės struktūros, kurios palengvina jų interpretavimą ir yra pranašesnės už kitus klasifikavimo medžių atvaizdavimo būdus. Jeigu klasifikavimo medis tampa painus (turi daug mazgų), tai jo tiesioginė grafinė reprezentacija tampa beprasmė ir tokios struktūros medžiams turi būti pritaikytos juos supaprastinančios grafinės procedūros.

Sprendimų medis yra klasifikatorius, išreikštas kaip rekursinis pavyzdžių (arba objektų) erdvės skaidinys. Sprendimų medis yra sudarytas iš mazgų, kurie formuoja medį su šaknimis (angl. *rooted tree*). Šaknimis (angl. *root*) yra vadinami medžio mazgai, neturintys įeinančių briaunų. Visi kiti mazgai turi lygiai po vieną įeinančią briauną. Mazgas su išeinančiomis briaunomis yra vadinamas vidiniu mazgu (angl. *internal node*) arba testiniu mazgu (angl. *test*

*node*). Visi kiti mazgai yra vadinami lapais (dar vadinami sprendimų mazgais). Sprendimų medyje kiekvienas vidinis mazgas (sprendimo priėmimo mazgas) padalina pavyzdžių erdvę į dvi arba daugiau poerdvių, atsižvelgdamas į atributų (požymių) funkcijos reikšmes bei sau užduotą sąlygą (Quinlan, 1996).



17 pav. Sprendimų medis

Pavyzdžiai yra klasifikuojami judant nuo medžio šaknies žemyn link lapų, atsižvelgiant į „pakeliui“ gautus rezultatus (Quinlan, 1986). Kiekvienai šakai yra pasirenkami konkretūs požymių rinkiniai. Kitas žingsnis – nagrinėjamas mazgas, kuriame pasirodo atšaka. Ciklas yra kartojamas šitame mazge (mazgas 1, mazgas 1.1), kol yra priartėjama prie lapų (lapai 1–3) (17 pav.).



### 3.1.1 Sprendimų medžio sudarymas

Tarkime, turime apmokymo aibę  $T$ , sudarytą iš objektų (pavyzdžių), kurių kiekvienas yra apibūdinamas atributais (požymiais)  $m$ . Vienas iš požymių arba tam tikrų požymių rinkinys parodo nagrinėjamo pavyzdžio priklausomybę konkrečiai klasei  $\{C_1, C_2, \dots, C_k\}$ .

Galimos trys situacijos:

1. Aibėje  $T$  yra vienas arba daugiau pavyzdžių, priklausančių vienai klasei  $C_k$ . Tada aibės  $T$  sprendimų medis – lapas, apibrėžiantis klasę  $C_k$ .
2. Aibėje  $T$  nėra nei vieno pavyzdžio, t. y. aibė yra tuščia. Tada turime lapą ir su lapu susijusi klasė yra atrenkama iš kitos aibės nei  $T$ , tarkime, iš aibės, susijusios su tėvine klase.
3. Aibėje  $T$  yra skirtingoms klasėms priklausančios pavyzdžiai. Tokiu atveju aibė  $T$  yra suskaidoma į poaibius. Tam yra pasirenkamas vienas iš požymių (arba požymių rinkinys), turintis skirtingas reikšmes  $O_1, O_2, \dots, O_n$ . Aibė  $T$  yra padalinama į poaibius  $T_1, T_2, \dots, T_n$ , kur kiekvienas poaibis  $T_i$  turi visus pavyzdžius, kurių pasirinkto požymio (arba požymių rinkinio) reikšmės yra  $O_i$ . Ši procedūra bus rekursyviai kartojama tol, kol baigtinė aibė nebus sudaryta iš pavyzdžių.

Aprašytos trys situacijos yra daugumos šiuolaikinių sprendimų medžių kūrimo algoritmų pagrindas. Šis būdas dar yra vadinamas „skaldyk ir valdyk“. Naudojant šią metodiką, sprendimų medis yra sudaromas principu „iš viršaus į apačią“ (Quinlan, 1992).

Toks sprendimų medžių sudarymo procesas yra vadinamas mokymu su mokytoju, kadangi visi egzemplioriai buvo iš anksto priskirti žinomoms klasėms. Apmokymo procesas yra vadinamas indukciniu mokymusi arba medžio indukcija.

Apmokant sprendimų medį yra svarbu pasirinkti padalijimo kriterijų (angl. *splitting criterion*). Daug įvairių padalijimo kriterijų gali būti apibrėžta

geriausiam padalijimui kiekviename mazge. Dažniausiai naudojami ir literatūroje minimi yra šie kriterijai (Rokach & Maimon, 2014): Gini priemaiša, Gini indeksas, informacijos pokytis, pokyčių koeficientas (angl. *Gain Ratio*), atstumo matas, binarinis kriterijus ir kiti.

### 3.1.2 Sprendimų medžius realizuojantys algoritmai

Yra žinoma nemažai sprendimų medžius realizuojančių algoritmų: CART, C4.5, NewId, ITrule, CHAID, CN2 ir t. t. Bet populiariausiais, sprendžiant klasifikavimo uždavinį, yra laikomi CART, C4.5 ir ID3.

**CART (Classification and Regression Tree)** – tai binarinio medžio konstravimo algoritmas (Breiman, et al., 1984), (Rokach & Maimon, 2014). Kiekvienas vidinis mazgas po skaidymo gali turėti tik du palikuonius. Padalijimui yra naudojamas Gini indeksas.

**ID3** yra laikomas labai paprastų sprendimų medžių algoritmu (Quinlan, 1986). Šis algoritmas gali persimokyti didelės apimties mokymo duomenimis, todėl jis yra naudojamas sudarant mažus medžius. ID3 skirtas nominaliems atributams, todėl tolydinės kilmės duomenys gali būti naudojami tik po jų konvertavimo į nominalias reikšmes. Kaip padalijimo kriterijus yra naudojamas informacijos pokytis, kurio vertinimas remiasi entropijos samprata. Šiuo algoritmu remiasi dauguma sprendimų medžių konstravimo metodų.

**C4.5** – tai sprendimų medžio konstravimo algoritmas su neribojamu palikuonių kiekiu. Kaip padalijimo kriterijus yra naudojamas informacijos pokytis. Šis algoritmas yra patobulintas ID3 variantas ir praktikoje naudojamas dažniau nei jo pirmtakas (Quinlan, 1992), (Rokach & Maimon, 2014).

Dauguma iš žinomų sprendimų medžių konstravimo algoritmų yra vadinami godžiaisiais (angl. *greedy*) algoritmais. Jeigu buvo pasirinktas konkretus atributas (požymis) ir jį naudojant buvo atliktas aibės skaidymas į poaibius, tai algoritmas nebesudaro galimybės grįžti atgal ir pasirinkti kitą atributą (požymį), kuris duotų geresnį aibės skirstymo į poaibius (klasifikavimo) rezultatą. Todėl sprendimų medžio konstravimo metu negalima

pasakyti, ar pasirinktas konkretus atributas (požymis) garantuos maksimaliai efektyvų skirstymą (klasifikavimą) galutiniame etape.

### **3.1.3 Sprendimų medžių metodo privalumai ir trūkumai**

Apibendrinant galima išskirti tam tikrus sprendimų medžių metodo privalumus ir trūkumus.

Metodo privalumai:

- metodą lengva suprasti ir interpretuoti. Klasifikavimo medžiai idealiai pritaikyti grafiniam vizualizavimui, todėl jų pagrindu priimtus sprendimus yra žymiai lengviau interpretuoti, nei tik skaitmenine forma išreikštus sprendimus;
- intuityviai suprantamas klasifikavimo modelis;
- nereikalauja specialaus duomenų paruošimo, pavyzdžiui, duomenų normalizavimo, fiktyvių duomenų pridėjimo, praleistų duomenų pašalinimo;
- metodas veikia pagal „baltos dėžės“ principą. Tam tikrą modelio situaciją galima aprašyti pasitelkus Būlio logiką;
- yra galimybė įvertinti modelio patikimumą statistinių testų pagalba;
- patikimas metodas, kuris gerai veikia net tuo atveju, kai pradinės prielaidos buvo pažeistos;
- modelis leidžia dirbti su didelės apimties duomenimis be specialių paruošiamųjų procedūrų;
- sprendimų medžiai gali dirbti tiek su nominaliais, tiek su skaitiniais atributais (Rokach & Maimon, 2014);
- sprendimų, paremtų medžio struktūromis, privalumas yra tai, kad medžiai leidžia apdoroti didelį kiekį požymių-kandidatų ir atrinkti tinkamiausius, kuriais remiantis atliekamas duomenų padalijimas (Breiman, et al., 1984).

Metodo trūkumai:

- dauguma sprendimų medžius realizuojančių algoritmų (kaip ID3 ir C4.5) reikalauja, kad atributai turėtų tik diskrečias reikšmes (Rokach & Maimon, 2014);
- praktinis sprendimų medžių algoritmo taikymas yra paremtas euristiniais algoritmais, tokiais kaip „godumo“ algoritmas, kai vienintelis optimalus sprendimas pasirenkamas lokaliai kiekviename mazge;
- gali būti sudarytos ir dažnai yra sudaromos labai sudėtingos medžio konstrukcijos. Tai yra vadinama medžio permokymu (Bramer, 2007);
- metodas remiasi „skaldyk ir valdyk“ principu, todėl jis gerai veikia naudojant glaudžiai susijusių ir reikšmingų atributų (požymių) rinkinius. Jei naudojamų atributų (požymių) tarpusavio sąveika sudėtinga, klasifikavimo medžių efektyvumas mažėja (Rokach & Maimon, 2014).

### 3.2 Pasiūlytoji hierarchinė emocijų klasifikavimo schema

Šiame darbe yra pristatoma hierarchinė šnekos emocijų klasifikavimo schema, kuri remiasi klasifikavimo medžio principu. Toks klasifikavimo organizavimas leidžia naudoti skirtingus požymių rinkinius kiekviename klasifikavimo etape, t. y. individualiai parinktą požymių rinkinį kiekvienai emocijai ar jų grupei.

Žemiau yra pateiktos trys pagrindinės hierarchinio emocionalios šnekos klasifikavimo prielaidos:

1. Atliekant emocijų klasifikavimą vienu etapu iškyla tokia problema, kaip nagrinėjamų emocijų savybių (akustinių, prozodinių ir kitų požymių reikšmių) **persidengimas**. Ši klasifikavimo problema gali būti supaprastinta mažinant vienu metu analizuojamų emocijų kiekį. Tai galima padaryti emocijų klasifikavimo procesą organizuojant lygiais, kiekviename jų analizuojant ribotą emocijų kiekį.

2. Kiekviena emocija yra charakterizuojama tik jai būdingais akustiniais bei prozodiniais požymiais. Šie požymiai skirtingoms emocijoms gali būti skirtingi arba vienodi. Todėl didinant emocijų atpažinimo **vidurkį nebūtinai bus didinamas** ir kiekvienos emocijos individualus klasifikavimo tikslumas. Šios problemos sprendimas būtų analizuoti kiekvieną emociją arba emocijų grupę, charakterizuojamą vienodais požymiais, atskirai. Tuomet maksimalus tokios grupės klasifikavimo efektyvumas garantuotų kiekvienos tos grupės emocijos klasifikavimo tikslumą atskirai.
3. Emocijos, priklausomai nuo pasirinktųjų požymių, gali būti suskirstytos į įvairias skirtingo abstraktumo klases. Pastarosios pagal kitus požymius gali būti suskirstytos į žemesniojo lygio klases ir t. t. iki pavienių emocijų klasių. Pavyzdžiui, skirstant emocijas pagal pagrindinio tono dažnį, galima sudaryti žemo tono ir aukšto tono klases. Pirmajai būtų galima priskirti liūdesį, nuobodulį, neutralią būseną. Prie aukšto tono emocijų priskirtume džiaugsmą, pyktį. Kiekviena iš šitų klasių gali būti suskirstyta į atskiras emocijas naudojant trukmės, energijos ir kitus klasifikavimo požymius. Dar kitoks skirstymas į klases būtų galimas pagal prozodinius požymius (pavyzdžiui, kylančio tono, pastovaus tono ir krentančio tono).

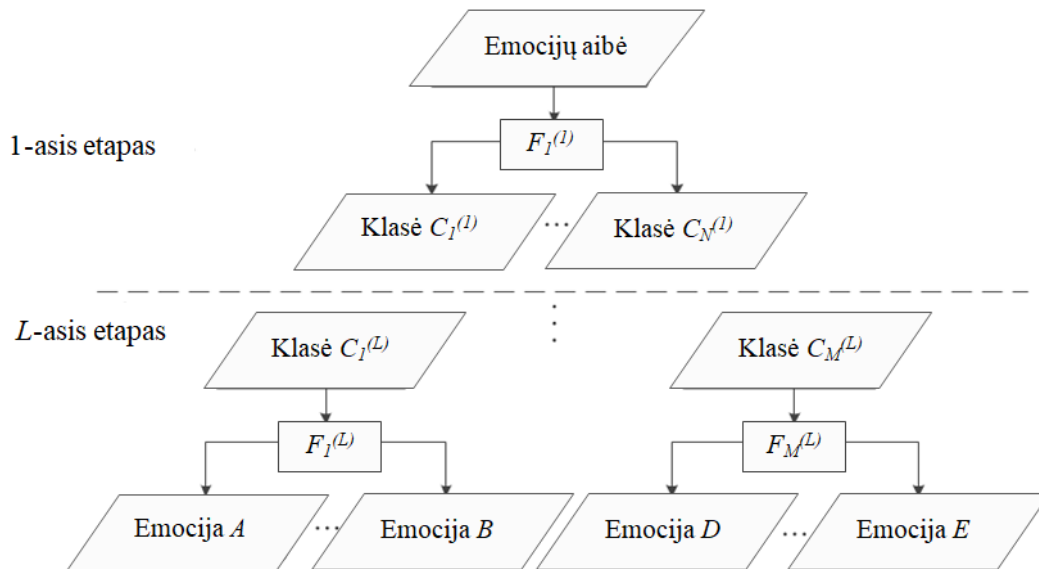
Taigi atsižvelgiant į iškeltas prielaidas, galima:

- organizuoti emocijų atpažinimą keliais lygiais;
- kiekvienam emocijų klasifikavimui naudoti kitą požymių rinkinį.

Numatomas **privalumas** – galimybė kiekvienam klasifikavimo lygiui sudaryti tam lygiui efektyviausią požymių rinkinį.

### 3.2.1 Apibendrinta hierarchinio klasifikavimo schema naudojant keletą lygių požymių

Atsižvelgiant į išvardytas prielaidas, emocijoms šnekoje klasifikuoti yra siūloma hierarchinė emocionalios šnekos klasifikavimo schema.



18 pav. Apibendrinta hierarchinio klasifikavimo schema naudojant keletą lygių požymių

Pagrindinė hierarchinio klasifikavimo schemos idėja yra emocionalios šnekos įrašų klasifikavimas keliais etapais naudojant skirtingus požymių rinkinius kiekviename etape (18 pav.). Pirmajame etape visi emocionalios šnekos įrašai yra skirstomi į  $N$  klasių  $\{C_1^{(1)}, \dots, C_N^{(1)}\}$ , kurias lemia pirmo etapo požymių rinkinys  $F_1^{(1)}$ . Šis požymių rinkinys yra sudarytas taip, kad klasifikavimas į klases  $\{C_1^{(1)}, \dots, C_N^{(1)}\}$  duotų maksimalų tikslumą. Antrajame etape kiekviena emocijų klasė yra skirstoma į žemesnio lygio klases  $\{C_1^{(2)}, \dots, C_K^{(2)}\}$  arba į pavienes atskiras emocijas, naudojant antrojo etapo požymių rinkinius  $F_k^{(2)}$ ,  $k = 1, \dots, K$ . Naudojant šią klasifikavimo schemą gali būti bet koks klasifikavimo etapų  $L$  skaičius su skirtingais požymių rinkiniais kiekvienai klasei. Emocijų klasių skaičius kiekviename etape taip pat yra

neribojamas. Tokiu būdu galima užtikrinti tinkamą požymių rinkinį kiekvienam klasifikavimo lygiui, t. y. kiekvienai emocijai.

Pavyzdžiui, nagrinėjant energiją kaip požymius, pirmajame klasifikavimo etape būtų vykdomas klasifikavimas į žemos ir aukštos energijos klases (žemos energijos emocijų pavyzdžiai – nuobodulys, neutrali būseną, o aukštos energijos – džiaugsmas, pyktis). Antrajame klasifikavimo etape žemos ir aukštos energijos klasės būtų klasifikuojamos į žemesnio lygio klases arba į atskiras emocijas, kurias apibrėžtų naudojami požymiai (visiškai kitokie, nei naudoti pirmojo etapo klasifikavimo metu).

Kiekvienas dalinis požymių rinkinys  $F_m^{(l)}$  yra naudojamas konkrečiai emocijai ar emocijų grupei klasifikuoti, tokiu būdu organizuojant emocijų klasifikavimą keliais lygiais. Pagrindiniai pasiūlytos klasifikavimo schemas principai yra tokie:

- Klasifikavimas yra organizuotas atskirais lygiais. Siekiant atpažinti emocijas ar emocijų grupes, kiekviename klasifikavimo etape yra naudojamas vis kitoks individualus ir efektyviausias požymis ar požymių rinkinys  $F_m^{(l)}$ . Tokie požymių rinkiniai gali būti sudaryti naudojant jau minėtas nuoseklaus aibės didinimo, nuoseklaus aibės mažinimo, geriausiojo pirmojo pasirinkimo arba kitas požymių atrankos technikas.
- Visas hierarchinis klasifikavimo procesas yra apibūdinamas bendru visų klasių ir emocijų požymių junginiu.

$$F = \{F_m^{(l)}\}, \quad m = 1, \dots, M; \quad l = 1, \dots, L. \quad (1)$$

Čia  $M$  yra emocijų klasių skaičius konkrečiame klasifikavimo lygyje,  $L$  yra klasifikavimo lygių skaičius.

Apskritai, emocijų rinkinys (arba klasių rinkinys gautas iš aukštesnio lygio klasių) gali būti klasifikuojamas į bet koki klasių skaičių. Analizuojamų klasių skaičius (vieno lygio) ir klasifikavimo lygių skaičius yra apibrėžiami bendru analizuojamų emocijų skaičiumi ir pasirinktais požymių rinkiniais.

Paprasčiausias hierarchinio emocijų klasifikavimo atvejis yra klasifikavimas į dvi emocijų klases arba emocijas.

Požymių rinkiniai gali būti nevienalyčiai – sudaryti iš skirtingų požymių (laiko, spektro, kepstro, energijos, balsu kokybės ir kt.) naudojant įvairius požymių atrankos metodus bei kriterijus.

Pagrindinis pasiūlytojo hierarchinio emocijų klasifikavimo metodo pranašumas yra tai, kad skirtingų lygių klasifikavimo procesai yra nepriklausomi nuo bendro požymių rinkinio. Klasifikavimo procesą galima optimizuoti gerinant pasirinktos emocijų klasės atpažinimo tikslumą nedarant įtakos kitų emocijų klasių atpažinimo tikslumui.

### 3.3 Klasifikavimo klaidos vertinimas

Klasifikavimo klaida (arba jam priešinga charakteristika – tikslumas), kaip ir tiesioginio klasifikavimo schemos atveju, išreiškiama klaidingai klasifikuotų ir visų pavyzdžių santykiu. Tačiau hierarchinės schemos atveju klasifikavimo klaida savo struktūra yra kompleksinė – ją sudaro daugelio lygio klasifikavimo klaidos.

**Klasifikavimo klaida** konkrečiai  $i$ -ajai emocijai išreiškiama:

$$E_i = \prod_{l=1, m=1}^{L, M} E(F_m^l(i)) \quad (2)$$

čia  $E(F_m^l(i))$  – klasifikavimo klaida naudojant požymių rinkinį  $F_m^l$   $l$ -ajame lygmenyje,  $m$ -ajame mazge.

Taigi norėdami minimizuoti hierarchinio klasifikavimo klaidą, mes turėtume minimizuoti klaidą kiekviename lygyje (mazge). Tai galime pasiekti maksimizuodami kiekviename klasifikavimo mazge naudojamo požymių rinkinio  $\{F_m^l\}$  efektyvumą klasifikavimo tikslumo atžvilgiu. Natūralus sprendimas tam pasiekti – požymių atranka kiekviename schemos mazge.



### 3.4 Požymių atrankos kriterijai

Ankstesniame skyriuje buvo paminėtos dvi požymių aibės mažinimo strategijos: transformavimo ir atrankos. Kadangi transformavimas yra susijęs su naujos požymių erdvės formavimu (tai nėra mūsų darbo tikslo dalis), tai apsiribosime požymių atrankos metodais.

Požymių atrankos metodai skirstomi į eksponentinės, nuosekliosios ir atsitiktinės paieškos metodus.

Eksponentinės paieškos metodams priskiriami pilnojo perrinkimo, šakų ir ribų, spindulio paieškos algoritmai. Esminis šių metodų trūkumas – didelis skaičiavimo operacijų kiekis, kurį lemia platus požymių atrankos laukas. Šiuose metoduose lygiagrečiai nagrinėjami keli arba visi įmanomi (pilnojo perrinkimo atveju) potencialių rinkinių variantai. Literatūroje galima rasti teiginių, liudijančių eksponentinį šių algoritmų sudėtingumą, t. y. nagrinėjamų požymių rinkinių skaičius eksponentiškai priklauso nuo nagrinėjamų požymių kiekio. Tai gali tapti rimta kliūtimi įgyvendinant šiuos algoritmus, turint omenyje, galimų požymių kiekį (siekiantį kelis tūkstančius). Dėl šios priežasties eksponentinės paieškos kriterijai nenagrinėti šiame darbe.

Atsitiktinės paieškos idėja – daugiau ar mažiau atsitiktinai generuoti požymių rinkinius ir išsirinkti efektyviausią iš jų. Prie atsitiktinių priskiriami genetinis algoritmas, atkaitinimo modeliavimas (angl. *simulated annealing*), atsitiktinio generavimo nuoseklaus atrinkimo algoritmas. Šiuose metoduose nagrinėjamų požymių rinkinių skaičius yra iš anksto apribojamas maksimaliu galimu iteracijų skaičiumi. Be to, atsitiktinė požymių atranka yra neatkartojama ir klasifikavimas taikant šį principą bus nestabilus efektyvumo požiūriu. Įvertinus šias aplinkybes, atsitiktinės požymių atrankos taikymo taip pat atsisakyta.

Prie nuoseklios atrankos grupės priskiriami metodai, kuriuose požymių rinkinys formuojamas nuosekliai pridedant (kai kuriais atvejais ir atimant) vieną ar daugiau požymių, tenkinančių tam tikrą kriterijų. Prie šių metodų priskiriami nuoseklaus aibės didinimo, nuoseklaus aibės mažinimo, nuoseklaus

svyruojančio aibės didinimo ir kiti metodai. Nagrinėjant  $N$  skirtingų požymių atrankos metodais, rinkinio padidinimas vienu požymiu reikalauja  $N$  skirtingų rinkinių analizės. Taigi  $M$  dydžio požymių rinkinys pareikalaus apytiksliai  $N \times M$  rinkinių-kandidatų analizės. Nuosekliosios atrankos metodams būdingas toks pats trūkumas – garantuojamas lokalus, bet ne globalus požymių rinkinio efektyvumas. Būtent šį nuoseklųjį principą pritaikysime formuodami požymių atrankos metodus hierarchinei klasifikavimo schemai.

Šiame darbe buvo suformuluoti ir pritaikyti 3 požymių atrankos kriterijai, kuriais remiantis hierarchiniam klasifikavimui pritaikyti trys požymių atrankos metodai:

1. maksimalaus efektyvumo kriterijus;
2. minimalios koreliacijos kriterijus;
3. nuoseklaus aibės didinimo metodas.

Tolimesniuose poskyriuose bus pristatyti šie atrankos metodai.

### 3.4.1 Maksimalaus efektyvumo požymių atranka

Maksimalaus efektyvumo atrankos metodas suformuotas remiantis prielaida apie tiesinę požymių efektyvumo savybę. Jeigu mes priimsime prielaidą apie efektyvumo tiesiškumą, požymių rinkinio efektyvumas bus garantuojamas rinkinį sudarant iš individualiai efektyviausių požymių. Individualus požymio efektyvumas įvertinamas klasifikuojant šnekos emocijų pavyzdžius.

Taigi maksimalaus efektyvumo (ME) požymių atrankos kriterijus leidžia atrinkti požymius su mažiausia klasifikavimo klaida

$$F_m^{(l)} = \left\{ \arg \min_j E(F_j^{(l)}) \right\}, \quad j = 1, \dots, J. \quad (3)$$

Čia  $E(F_j^{(l)})$  –  $l$ -tojo lygio klasifikavimo klaida naudojant  $j$ -ąjį požymių rinkinį  $F_j^{(l)}$ .  $J$  žymi bendrą požymių skaičių  $l$ -tame klasifikavimo lygyje.

Efektyviausi požymiai  $F_j^{(l)}$  buvo pakartotinai pridedami į požymių rinkinį  $F_m^{(l)}$ . Požymių rinkinio didinimas sustabdomas išnagrinėjus visus  $J$  požymius-kandidatus.

Maksimalaus efektyvumo požymių atranka iš esmės panaši į geriausiojo pirmojo pasirinkimo metodą. Tačiau šiame darbe siūlomame metode požymių rinkinys sudaromas nagrinėjant visus požymius-kandidatus, t. y. stabdymo kriterijumi naudojamas ne požymių rinkinio efektyvumo nebedidėjimas, o požymių-kandidatų skaičius.

Naudojant šį atrankos metodą teks nagrinėti  $J$  skirtingų požymių rinkinių ir iš jų išrinkti efektyviausią.

### **3.4.2 Minimalios koreliacijos požymių atranka**

Prielaida, nulėmusi šio požymių atrankos kriterijaus pasirinkimą yra tiesinė požymių tarpusavio nepriklausomybė. Tiesiškai nepriklausomi požymiai turi didesnę diskriminantinę galią palyginus su tiesiškai priklausomais požymiais. Tokiu būdu sudarant tiesiškai nepriklausomų ir maksimaliai efektyvų požymių rinkinį yra tikėtina gauti tikslesnį emocijų klasifikavimo procesą.

Paprasčiausias būdas įvertinti tiesinę duomenų tarpusavio priklausomybę yra kryžminės koreliacijos naudojimas (Guyon & Elisseeff, 2003), (Oflazoglu & Yildirim, 2013), (Mencattini, et al., 2014). Bendruoju atveju maksimali absoliuti kryžminės koreliacijos reikšmė parodo pilną tiesinę duomenų priklausomybę, o lygi nuliui kryžminės koreliacijos reikšmė patvirtina visišką tiesinę duomenų tarpusavio nepriklausomybę.

Taigi minimalios koreliacijos atrankos principą įgyvendinsime požymių rinkinį formuodami iš minimalia tarpusavio koreliacija pasižyminčių požymių. Kadangi koreliacija vertinama paporiui, t. y. požymių porai, rinkiniui formuoti atrinksime požymius, turinčius minimalią koreliaciją su pirmuoju, rinkinį inicializuojančiu požymiu.

Siekdami gauti kiek įmanoma efektyvesnį požymių rinkinį, pasirinkome jį inicializuoti individualiai efektyviausiu požymiu (kaip ir maksimalaus efektyvumo atrankoje).

$$F_0^l = \arg \min_i E(F_i^l), i = 1, \dots, P, \quad (4)$$

$E(F_i^l)$  -  $l$ -ojo lygio klasifikavimo klaida naudojant vieną iš  $P$  požymių  $F_i^l$ .

Požymių vektorių didiname pridėdami nuo inicializuojančiojo tiesiškai nepriklausomus požymius  $F_k^l$ . Tiesinė požymių tarpusavio priklausomybė leidžia tikėtis, jog tokie požymiai turės maksimalią diskriminantinę galią. Paprasčiausias tiesinės priklausomybės įvertinimo būdas – požymių tarpusavio koreliacijos skaičiavimas.

$$F_k^l = \arg \min_j R(F_0^l, F_j^l), j = 1, \dots, P, \quad (5)$$

$R(F_0^l, F_j^l)$  pradinio  $F_0^l$  ir  $l$ -ojo lygmens  $j$ -ojo požymio tarpusavio koreliacija.

Kiekvieną kart pridėjus požymį  $F_k^l$ , gaunamo požymių rinkinio efektyvumas vertintas klasifikuojant emocijas. Požymių rinkinio  $\{F_m^l\}$  didinimas sustabdomas, kai išnagrinėjama  $P$  mažiausią koreliaciją su efektyviausiuoju turinčių požymių.

Požymiai nėra pasiskirstę pagal Gauso dėsnį, todėl tiesinei požymių tarpusavio priklausomybei vertinti pasirinktas Spearman koreliacijos koeficientas, nereikalaujantis išankstinių prielaidų apie nagrinėjamų duomenų pasiskirstymo dėsnį. Be to, šis koreliacijos metodas apdoroja duomenis netiesiogiai, o nagrinėja jų klases (kategorijas), tai yra šio metodo išskirtinumas lyginant jį su Pearson koreliacijos koeficientu.

Spearman koreliacijos koeficientas skaičiuojamas:

$$R_s = 1 - 6 \frac{\sum_i d_i^2}{n^2(n-1)}, \quad (6)$$

$d_i$  – nagrinėjamų požymių rangų skirtumas,  $n$  – nagrinėjamų požymių porų skaičius.

Požymių atrankai naudojant šį atrankos metodą teks nagrinėti  $P$  skirtingų požymių rinkinių ir iš jų išrinkti efektyviausią.

Reikėtų pastebėti, jog nagrinėjamų minimalios koreliacijos požymių skaičių  $P$  padidinę iki maksimalaus nagrinėjamų požymių skaičiaus  $J$ , minimalios koreliacijos atrankos metodo rezultatai gali sutapti su maksimalaus efektyvumo požymių atrankos rezultatais (3).

### 3.4.3 Nuoseklaus aibės didinimo metodas

Nuoseklaus aibės didinimo (SFS) metodas yra „godusis“ paieškos algoritmas, turintis palyginti žemą skaičiuojamąją galią. Šis metodas atranka požymių rinkinį pagal jo maksimalų efektyvumą klasifikavimo tikslumo požiūriu. Efektyvumas yra vertinamas nuosekliai didinant požymių rinkinį. Šis paieškos būdas duoda sprendimą, nutolusį nuo optimalaus taško, kadangi ne visi galimi požymių rinkiniai yra analizuojami (Guyon & Elisseeff, 2003).

Formuojant požymių rinkinį nuoseklaus aibės didinimo algoritmu, pradedama nuo tuščio požymių rinkinio  $F_0$ .

$$F_0 = \{\emptyset\}, i = 0. \quad (7)$$

Požymių rinkinys  $F_0$  yra didinamas tol, kol naujai gautas požymių rinkinys  $F_{i+1}$  duoda geresnį klasifikavimo rezultatą:

$$f_m = \arg \max [E(F_i + f_m) - E(F_i)], \quad m = 1, 2, \dots, M \quad (8)$$

$$F_{i+1} = \{F_i + f_m\}, \quad i = 1, 2, \dots, J \quad (9)$$

Požymių rinkinio išplėtimo žingsnis yra kartojamas tol, kol naujai sudaryto požymių rinkinio  $F_{i+1}$  klasifikavimo tikslumas didėja arba kol  $i < J$ , čia  $J$  yra bendras analizuojamų šnekos emocijų požymių skaičius.

Naudojant nuoseklaus aibės didinimo metodą hierarchiniam emocijų klasifikavimui pirmame etape, visi emocionalios šnekos įrašai yra skirstomi į anksto numatytas klases. Šias klases nustato požymių rinkinys, inicijuotas pirmajam klasifikavimo lygiui. Tokiu pat būdu yra analizuojamas

klasifikavimo uždavinys antrajame etape ir yra gaunami antrojo etapo požymių rinkiniai.

Apskritai nuoseklus aibės didinimas reikalauja  $J!/M!$  rinkinių analizės  $M$  dydžio efektyviausiam rinkiniui nustatyti. Taigi potencialiai tai daugiausiai skaičiavimo operacijų reikalaujantis požymių atrankos metodas iš pasiūlytųjų. Tikėtina, kad jis leis gauti efektyvesnį požymių rinkinį nei kiti du metodai.

### 3.5 Trečio skyriaus apibendrinimas

Atlikę sprendimų medžių sistemos taikymą klasifikavimo uždaviniams spręsti, galime suformuoti tokius apibendrinimus:

- Sprendimų medžiai yra hierarchinės sprendimų sistemos ir gali būti taikomos emocionalios šnekos klasifikavimo uždaviniui spręsti. Taikant sprendimų medžius, duomenys klasifikuojami į iš anksto žinomas klases pagal atrinktus požymius.
- Egzistuoja įvairūs sprendimų medžius realizuojantys algoritmai, klasikiniai yra CART, ID3 ir C4.5.

Remiantis atlikta analize bei gautomis išvadomis yra pasiūlyta:

- Hierarchinė emocionalios šnekos klasifikavimo schema, paremta sprendimų medžiais, leidžianti naudoti skirtingus požymių rinkinius kiekviename klasifikavimo etape. Esminis šios schemos privalumas – klasifikavimo proceso organizavimas lygiais, iš kurių kiekvienas gali būti realizuotas jam individualiu klasifikatoriumi ir požymių rinkiniu.
- Hierarchinės klasifikavimo sistemos efektyvumas maksimizuojamas maksimizuojant klasifikavimo tikslumą atskiruose sistemos lygiuose. Tą galima atlikti atliekant požymių atranką kiekvienam hierarchinės klasifikavimo schemos lygiui atskirai.
- Požymių rinkiniams sudaryti yra suformuoti maksimalaus efektyvumo, minimalios koreliacijos metodai bei pritaikytas nuoseklus aibės didinimo metodas.

## 4. Eksperimentinis tyrimas

Tyrimo tikslas – ištirti pasiūlytų požymių atrankos kriterijų bei hierarchinės emocijų klasifikavimo schemos efektyvumą.

Eksperimentinis tyrimas bus vykdomas trimis etapais. Pirmiausia nagrinėsime požymių atrankos kriterijus – tarpusavyje palyginsime šnekos emocijų klasifikavimo tikslumą naudojant maksimalaus efektyvumo, minimalios koreliacijos ir nuoseklaus aibės didinimo metodus. Efektyvumui įvertinti kartu pateiksime tiesioginio emocijų klasifikavimo, naudojant pilnus požymių rinkinius, rezultatus.

Antrajame eksperimentinės patikros etape vertinsime hierarchinės schemos efektyvumą bei lyginsime ją su tiesioginio klasifikavimo schema. Požymių atrankai naudosime pirmojo etapo metu naudotus atrankos metodus.

Galiausiai bandysime palyginti keletą skirtingų hierarchinio emocijų klasifikavimo schemų ir įvertinti jų tarpusavio pranašumus.

### 4.1 Tyrimo duomenys

Emocijoms šnekoje atpažinti (kitiems emocionalios šnekos analizės tikslams) yra naudojami emocionalios šnekos duomenų bazių įrašai. Tokiose bazėse yra konkrečias emocijas atspindintys šnekos įrašai. Dauguma emocionalios šnekos duomenų bazių apima 5–6 emocijas, kurios dažnai yra įvardijamos pagrindinėmis emocijomis (Cornelius, 1996), (Cowie & Cornelius, 2003), (Ververidis & Kotropoulos, 2006). Realiame gyvenime emocijų kategorijų yra daugiau.

Dažniausiai pasitaikančios emocijos, kurias galima rasti duomenų bazėse, yra pyktis, liūdesys, laimė, baimė, pasišlykštėjimas, nuostaba, nuobodulys ir džiaugsmas. Neutrali šneka taip pat yra dažnai nagrinėjama (Koolagudi & Rao, 2012).

Esamos emocionalios šnekos duomenų bazės yra labai skirtingos. Jos yra įrašytos įvairiomis kalbomis (anglų, vokiečių, kinų, ispanų, japonų, rusų, švedų, italų, indų ir kt.), skiriasi jų apimtys, emocijos bei jų skaičius,

šnekančiųjų skaičius, kūrimo tikslas (atpažinimui, sintezei), įrašų kokybė, emocijų išraiškingumas ir kt. (Ververidis & Kotropoulos, 2003), (Koolagudi & Rao, 2012).

Atlikus palyginamąją emocionalios šnekos duomenų bazių analizę buvo nuspręsta eksperimentiniame tyrime naudoti Berlyno emocionalios šnekos duomenų bazę (Schuller, et al., 2009) ir Lietuvių emocionalios šnekos duomenų bazę (Matuzas J., et al., 2015). Kitos nagrinėtos emocionalios šnekos duomenų bazės buvo atmestos dėl tokių priežasčių, kaip netinkamas duomenų bazės tipas, nepakankamas emocijų rinkinys, žema įrašų kokybė, silpnai išreikštos emocijos, triukšmas, kelių kalbėtojų balsai viename įrašė ir nepakankamas emocionalios šnekos įrašų kiekis.

**Berlyno emocionalios šnekos duomenų bazė** (*Berlin Database of Emotional Speech*)

*Kalba:* vokiečių.

*Tipas:* suvaidintos emocijos.

*Kalbantieji:* 10 aktorių (5 vyrai ir 5 moterys).

*Emocijos:* džiaugsmas, pyktis, baimė, liūdesys, išgąstis, nuobodulys ir neutrali būseną.

*Įrašai:* 10 sakinių (5 trumpi ir 5 ilgesni), kurie gali būti naudojami kasdieniam bendravimui.

Bendra duomenų bazės apimtis yra 500 įrašų. Įrašant Berlyno emocionalios šnekos duomenų bazę dalyvavo 10 aktorių. Buvo atrinkta 10 sakinių, kurie yra sukonstruoti taip, kad juos būtų galima perteikti visomis pasirinktomis emocijomis. Duomenų bazės sakiniai yra sudaryti taip, kad juose būtų kiek įmanoma daugiau balsių, tam, kad būtų galima atlikti formančių analizę. Norint pasiekti aukštą įrašų kokybę, įrašinėjimas buvo atliekamas beaidėje kameroje su aukštos kokybės įrašymo įranga. Kalbantieji stovėjo priešais mikrofoną maždaug 30 cm atstumu ir esant poreikiui galėjo naudoti kūno kalbą. Tyrimas parodė, kad klausytojai atpažįsta 80 proc. emocijų, o 60 proc. duomenų bazės emocijų yra artimos natūralioms emocijoms. Ši duomenų bazė yra plačiausiai naudojama ir yra laisvai prieinama



(<http://emodb.bilderbar.info/docu/>) (Vogt & Andre, 2005), (Burkhardt, et al., 2005), (Xiao, et al., 2009), (Bitouk, et al., 2010), (Koolagudi, et al., 2010).

**Lietuvių emocionalios šnekos duomenų bazė** (*Lithuanian Spoken Language Emotions Database*)

*Kalba:* lietuvių.

*Tipas:* suvaidintos emocijos.

*Kalbantieji:* 10 neprofesionalių aktorių (5 vyrai ir 5 moterys).

*Emocijos:* džiaugsmas, pyktis, baimė, liūdesys, nuostaba, pasišlykštėjimas ir neutrali būseną.

*Įrašai:* 7 000 įrašų.

Kiekvienas iš 10 neprofesionalių aktorių skaitė po 100 neutralaus konteksto sakinių kiekvienai emocijai. Buvo sukurta po 1 000 įrašų kiekvienai iš septynių emocinių būsenų (Matuzas J., et al., 2015).

## 4.2 Tyrimo priemonės

Emocijų atpažinimo eksperimentas modeliuotas MATLAB aplinkoje. Tuo tikslu buvo sukurtas hierarchinio šnekos emocijų klasifikavimo sistemos prototipas, leidžiantis vykdyti eksperimentus bei vertinti jų rezultatus.

Požymiams išskirti buvo panaudotas automatizuotas įrankis *OpenEAR* (Eyben, et al., 2009), (Eyben, et al., 2013). Ši sistema leidžia išskirti net 6 552 skirtingus šnekos požymius bei įvairias jų statistines reikšmes.

## 4.3 Tyrimo sąlygos

Eksperimentiniam tyrimui pasirinktos 5 emocinės būsenos: neutralumas, džiaugsmas, pyktis, liūdesys ir baimė. Šis emocijų rinkinys laikomas baziniu ir pakankamai iliustruojančiu šnekos emocijų atpažinimo uždavinį.

Eksperimentiniame tyrime buvo nagrinėjami 3 skirtingi emocijų identifikavimo šnekoje uždaviniai: 3 emocijų (pyktis, džiaugsmas ir neutrali būseną), 4 emocijų (pyktis, džiaugsmas, neutrali būseną ir liūdesys) ir 5 emocijų (pyktis, džiaugsmas, neutrali būseną, liūdesys ir baimė). 2 emocijų

uždavinys nebuvo nagrinėjamas, nes tokiu atveju hierarchinis klasifikavimas nebeįmanomas. Didesnis emocijų skaičius buvo atmestas dėl kai kurių emocijų nepagrįstumo (pavyzdžiui, nuobodulio).

Emocijos nebuvo vertinamos nei psichologiniu, nei socialiniu požiūriu ir nagrinėtos tik kaip skirtingos duomenų klasės.

Kadangi Berlyno emocionalios šnekos duomenų bazėje sukauptų įrašų kiekis skirtingoms emocijoms skiriasi, siekiant suvienodinti eksperimento sąlygas, nagrinėjamų emocijų įrašų kiekis yra apribojamas iki 60 vienetų kiekvienai emociinei būsenai. Toks pavyzdžių kiekis nėra labai didelis, todėl eksperimentai bus vykdomi naudojant kryžminio vertinimo (angl. *k-fold cross-validation*) metodiką (Kohavi, 1995), (Gjoreski & Gjoreski, 2014). Kaip galutinis rezultatas bus fiksuojamas vidutinis visų testinių aibių rezultatas.

Vokiečių kalbos atveju pasirinktas duomenų skaidymo į 3 aibes variantas. Toks skaidymas užtikrina 20 pavyzdžių testinę aibę ir 5 proc. skiriamąją gebą vertinant klasifikavimo tikslumą.

Lietuvių kalbos atveju bus naudojama 1 000 kiekvienos emocijos pavyzdžių (10 žmonių po 100 ištarimų kiekvienai emocijai). Duomenų aibė bus skaidoma į 10 aibių – eksperimente bus taikomas *10-fold* kryžminis vertinimas.

Pasirinkti kryžminio vertinimo parametrai užtikrina, kad visi vykdomi eksperimentai bus nepriklausomi nuo kalbėtojo. Tai reiškia, jog apmokymui ir testavimui bus naudojami skirtingų kalbėtojų įrašai. Tokios eksperimento sąlygos yra sudėtingesnės, nes taip prarandama informacija apie kalbantįjį asmenį, tai gali sumažinti klasifikavimo tikslumą, tačiau, mūsų nuomone, toks testavimas leis geriau atskleisti siūlomos klasifikavimo schemos efektyvumą.

Emocionalios šnekos įrašams klasifikuoti pasirinktas *k*-artimiausių kaimynų (KNN) klasifikatorius (Dudani, 1976), (Khan, et al., 2011), (Gadhe, et al., 2015), (Lanjewar, et al., 2015). Pagrindiniai pasirinkimo argumentai buvo šie:

- Naudojamoje vokiečių kalbos emocijų bazėje bus naudojama po 60 kiekvienos emocijų pavyzdžių (40 iš jų skiriamas klasifikatoriui

apmokėti). Toks kiekis yra nepakankamas apibendrinantiems klasifikatoriams (pavyzdžiui, Bajeso, dirbtinių neuronų tinklo ir pan.) apmokėti. Todėl pasirinktas pavyzdžiais grįstas klasifikatorius.

- KNN klasifikatorius priklauso apmokymo su mokytoju klasifikatorių tipui. Tai leis valdyti nagrinėjamų emocijų kiekį.
- Klasifikatoriaus nepriklausomybė nuo paieškos erdvės. Nagrinėjama požymių erdvė apibrėžiama naudojamo atstumo tipu. Bus naudojamas Euklido atstumas.

Vokiečių kalbos atveju bus naudojamas klasifikatorius su 5 artimiausių kaimynų analize (5-NN), lietuvių kalbos atveju – su 7 kaimynų analize (7-NN). Tokie parametrai pasirinkti įvertinus apmokymo medžiagos apimtį. Tokiu būdu gausime, jog nagrinėjamųjų artimiausių kaimynų dalis sudarys ne daugiau kaip 25 proc. apmokymo aibės.

#### **4.4 Tyrimo eiga ir rezultatai**

Vertindami siūlomos hierarchinės klasifikavimo schemos efektyvumą, ją lyginsime su tiesioginiu klasifikavimo atveju (kai visos emocijos klasifikuojamos vienu žingsniu, naudojant maksimaliai galimą požymių aibę).

Hierarchinį klasifikavimą vertinsime tokiais būdais:

- bendru vidutiniu klasifikavimo tikslumu;
- kiekvienos emocijos klasifikavimo tikslumu;
- atitikties matrica, kurioje pateikiami duomenys apie klaidas;
- kiekvienu atveju gaunamo požymių rinkinio dydžiu.

Hierarchinę klasifikavimo schemą sudarysime naudodamiesi „aukšto tono“ ir „žemo tono“ kategorijomis. Prie pirmųjų priskirsime emocijas, kurioms būdingas pakeltas tonas, didesnis kalbėjimo tempas, aukštesnis garso lygis. Prie jų priskirsime pyktį ir džiaugsmą. Prie „žemo tono“ emocijų priskirsime neutralią būseną, liūdesį ir baimę. Pasirinkimo motyvai:

- Šnekos signalo pagrindinio tono dažnis (kuris yra tiesiogiai susijęs su mūsų suvokiamu tonu) yra stipriai susietas su šnekos emocijomis.

Todėl „aukšto tono“ ir „žemo tono“ kategorijų naudojimas mums leis grupuoti emocijas į mums patogias klases.

- „Aukšto tono“ emocijoms (džiaugsmui ir pykčiui) būdingas ir didesnis kalbėjimo tempas, padidintas garso lygis, didesnė energija. Tą turėtų puikiai įvertinti šnekos signalui apdoroti naudojami akustiniai požymiai.

Konkrečios klasifikavimo schemas bus pristatomos atskirai, vertinant jų efektyvumą, kitame poskyryje.

## **4.5 Eksperimentų rezultatai**

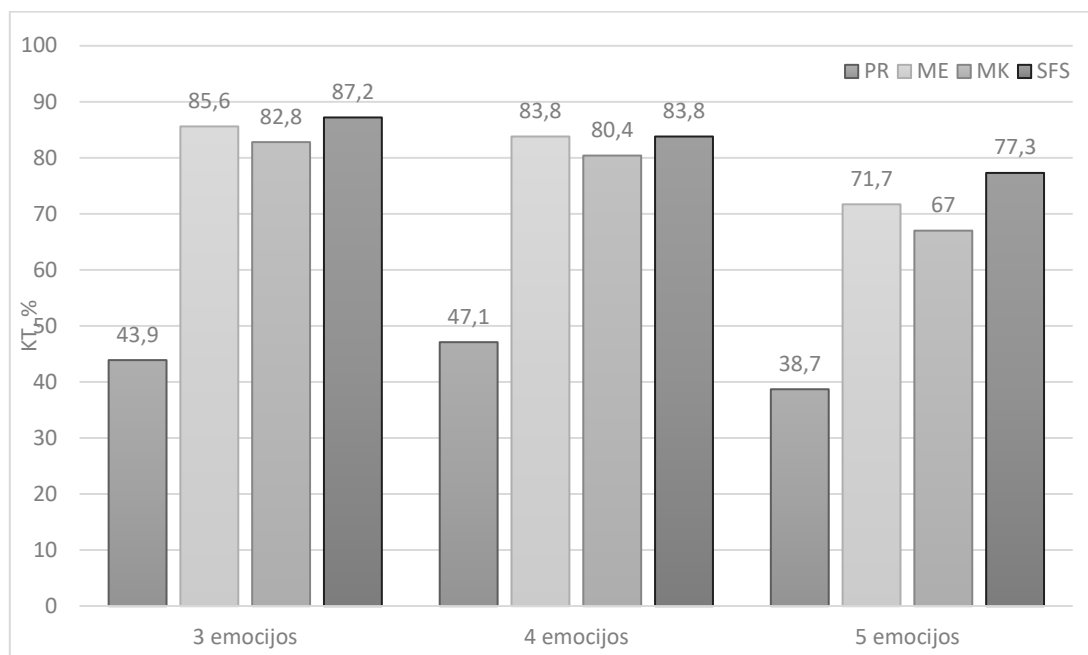
### **4.5.1 Atrankos kriterijų palyginimas**

Šiame etape suformuotus požymių atrankos kriterijus palyginome klasifikuodami emocijas. Atrankos kriterijus lyginome remdamiesi vidutiniu tiesioginio emocijų klasifikavimo, naudojant skirtingus požymių atrankos kriterijus, tikslumu (KT) ir atrankos kriterijaus suformuoto požymių vektoriaus ilgiu, išreikštu vektoriaus elementų skaičiumi.

Efektyvumui vertinti panaudojome abi duomenų bazines. Pirmosios – vokiečių kalbos – duomenų bazės įrašų panaudojome po 60 kiekvienos emocijos. Lietuvių kalbos pasirinkome po 300 kiekvienos emocijos (10 kalbėtojų × 30 ištarimų) pavyzdžių.

Siekdami įvertinti siūlomus požymių atrankos kriterijus, maksimalaus efektyvumo (ME) ir minimalios koreliacijos (MK) požymių paiešką apribojome iki pusės pilnos aibės dydžio, t. y. bus nagrinėjami 3 276 individualiai efektyviausi ir minimalią tarpusavio koreliaciją turintys požymiai. Pilnąją požymių aibę sudarys 6 552 akustiniai požymiai.

Vokiečių kalbos emocijų klasifikavimo rezultatai pateikiami 19 pav., o lietuvių kalbos – 20 pav.



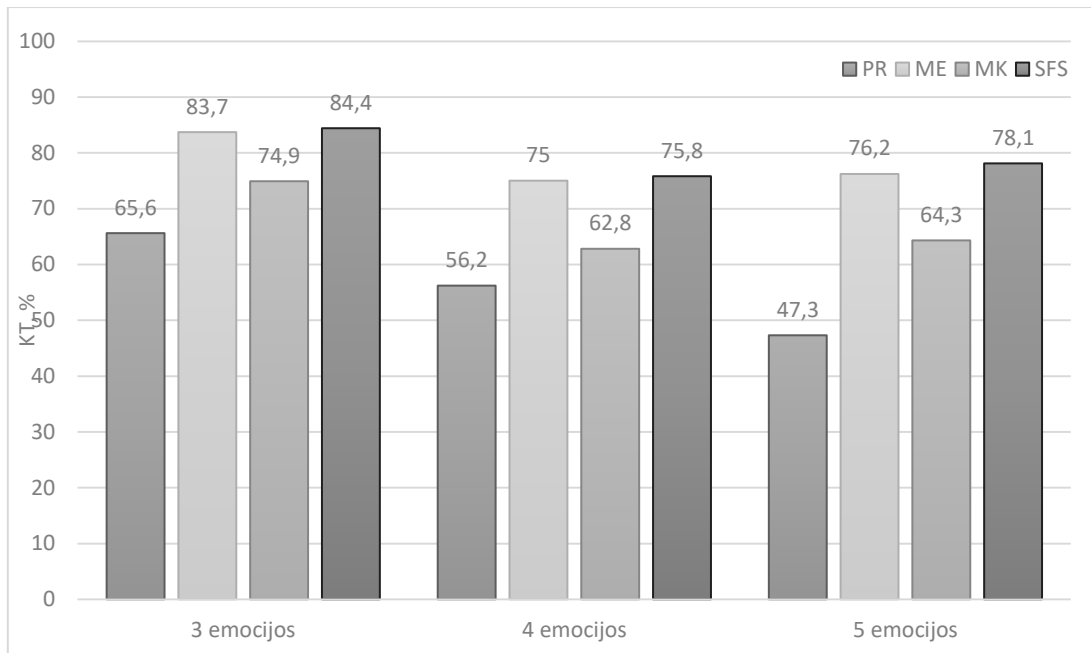
**19 pav. Vokiečių kalbos emocijų vidutinis klasifikavimas naudojant pilną požymių rinkinį (PR), Maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aibės didinimo (SFS) atrankos kriterijus**

Kaip matome 19 pav., pilno požymių rinkinio atveju gautas mažiausias tikslumas nesiekė ir 50 proc. Didžiausias tikslumas pasiektas naudojant nuoseklaus aibės didinimo metodą (SFS) – vidutinis tikslumas kito nuo 77,3 proc. penkių emocijų atveju iki 87,2 proc. trijų emocijų atveju. Pastaruoju atveju klasifikavimo tikslumas maždaug dvigubai viršijo pilno požymių rinkinio duodamą tikslumą. Taigi požymių atrankos taikymo efektas – akivaizdus.

ME ir MK atrankos metodai lėmė nedaug nutolusius tikslumo rezultatus. ME kriterijaus rezultatai atsiliko 1,6 proc., MK kriterijaus – 3,4–9,7 proc. Reikia pastebėti, jog 4 emocijų atveju ME ir SFS kriterijų rezultatai buvo tokie patys, tačiau jie gauti iš esmės suformavus skirtingus požymių rinkinius, taigi sutapimas – atsitiktinis.

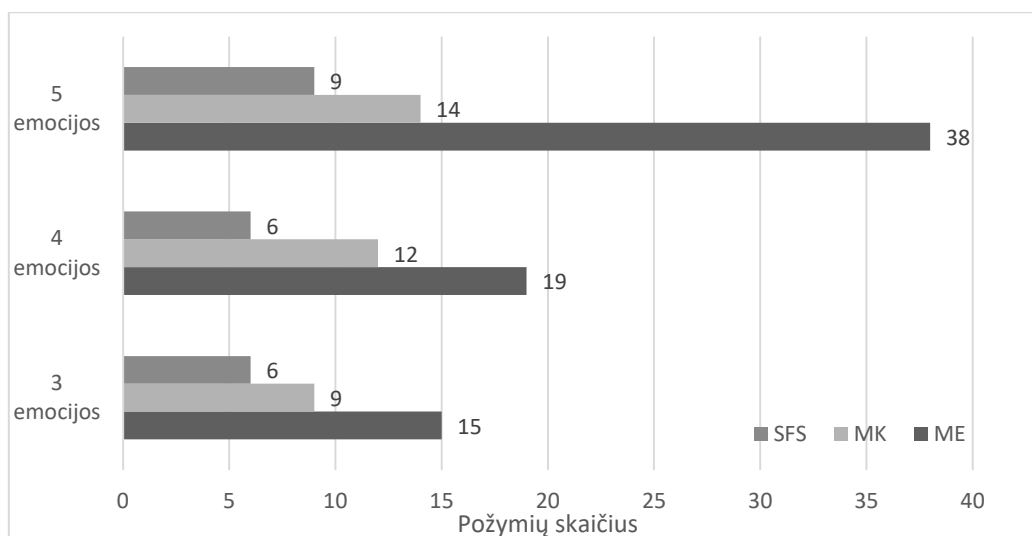
Lietuvių kalbos emocijų klasifikavimo rezultatuose, pateiktuose 20 pav., matome iš esmės analogiškus rezultatus. Klasifikavimo rezultatai, naudojant skirtingus požymių atrankos kriterijus, iš esmės gauti panašūs. Mažiausias tikslumas gautas naudojant pilnus požymių rinkinius, didžiausias – pritaikius nuoseklaus aibės didinimo metodą. Maksimalaus efektyvumo ir minimalios koreliacijos metodai šiuo atveju parodė ne identiškus rezultatus, priešingai nei

vokiečių kalbos atveju. 5 emocijų atveju minimalios koreliacijos metodo taikymas lėmė prastesnius klasifikavimo rezultatus.



**20 pav. Lietuvių kalbos emocijų vidutinis klasifikavimas naudojant pilną požymių rinkinį (PR), Maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aibės didinimo (SFS) atrankos kriterijus**

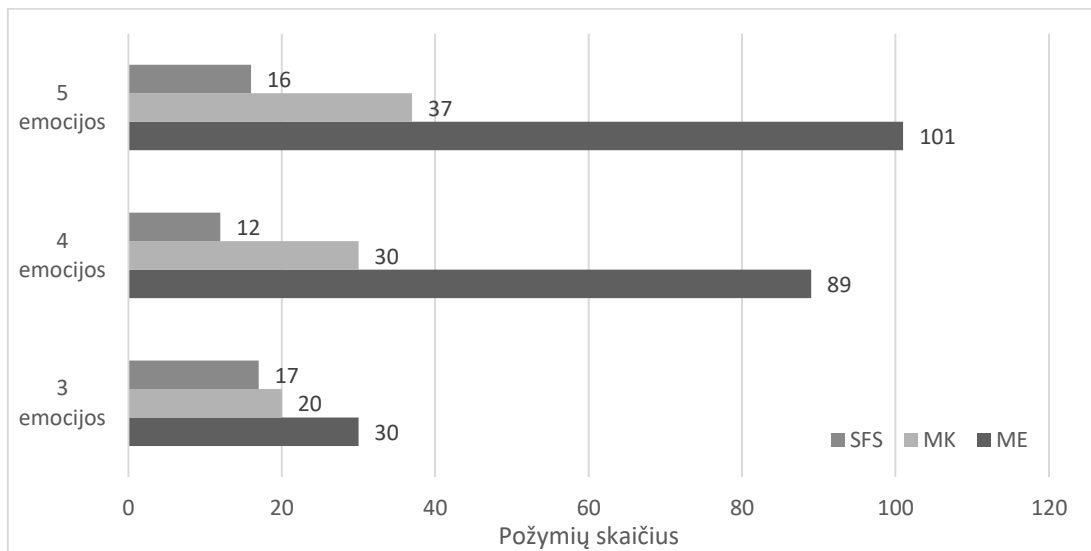
Taigi tikslumo požiūriu nuoseklaus aibės didinimo metodas buvo efektyviausias tiek lietuvių, tiek vokiečių kalbos atvejais.



**21 pav. Atrankos metodų grąžinamas vektorių dydis vokiečių kalbos emocijų atveju**

Visiems atrankos metodams (taip pat ir pilno požymių rinkinio atveju) abiejų kalbų atveju pastebėtas klasifikavimo tikslumo mažėjimas augant nagrinėjamų emocijų skaičiui. Rezultatas lauktas – didėjant nagrinėjamų emocijų kiekiui, klasifikavimo uždavinio sudėtingumas auga.

21 pav. ir 22 pav. pateikti požymių vektorių dydžiai, gauti naudojant skirtingus atrankos metodus. Tiek vokiečių, tiek lietuvių kalbos atveju mažiausias požymių vektorius gautas SFS metodo atveju – jo dydis kito nuo 6 (3 emocijų vokiečių kalba) iki 17 (3 emocijų lietuvių kalba). Didžiausias požymių rinkinio dydis gautas ME kriterijaus atveju, jis kito nuo 15 (3 emocijų uždavinyje) iki 101 (5 emocijų lietuvių kalba uždavinyje). Taigi požymių rinkinių dydžio skirtumai sudarė 2–6 kartus. Kadangi požymių skaičius išaugdavo pasikeitus nagrinėjamų pavyzdžių kiekiui (vokiečių ir lietuvių kalbos įrašų apimtys šiame eksperimente skyrėsi 5 kartus), galime daryti prielaidą, jog augant nagrinėjamų pavyzdžių kiekiui, reikalingų požymių skaičius taip pat didėja.



22 pav. Atrankos metodų grąžinamas vektorių dydis lietuvių kalbos emocijų atveju

Požymių sudėtis buvo panaši abiejų kalbų atvejais. 3 emocijų uždavinyje tiek vokiečių, tiek lietuvių kalbos atveju tarp atrinktųjų požymių dominavo amplitudės spektro juostų (ASJ) ir melų skalės kepstro (MSK) požymiai bei jų

statistiniai parametrai. Panašūs rezultatai gauti ir 4 emocijų uždavinio atveju. Nagrinėjant 5 emocijas greta minėtųjų taip pat atrinkti signalo energijos logaritmo požymiai bei jų statistiniai parametrai.

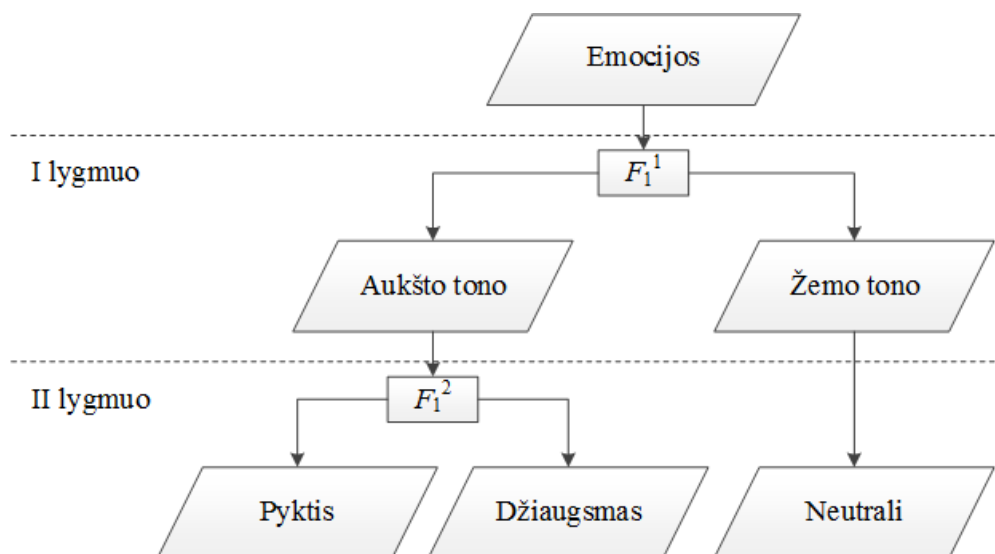
Skirtingi atrankos metodai lėmė skirtingus gaunamus požymių rinkinius. SFS atrankos atveju rinkinyje dominavo ASJ požymiai. Jų dalis rinkinyje kito nuo 17 proc. (4 vokiečių kalbos emocijų atveju) iki 71 proc. (3 lietuvių kalbos emocijų atveju). Naudojant ME ir ME ir MK atrankos kriterijus rinkiniuose dominavo MSK požymiai – jų dalis kito nuo 16 proc. iki 74 proc.

Vertinant klasifikavimo trukmę, efektyviausias atrankos metodas pasirodė esantis pats lėčiausias. Nuoseklaus aibės didinimo metodas lėmė 4–8 kartus didesnę klasifikavimo trukmę vokiečių kalbos atveju bei apie 20 kartų – lietuvių kalbos atveju.

Remdamiesi požymių atrankos metodų rezultatais, galime manyti, jog nuoseklaus aibės didinimo metodas yra pranašesnis už kitus metodus. Pranašumą šiam metodui suteikia jo orientacija į bendrą požymių poaibio efektyvumą, o ne į individualų kiekvieno požymio atskirai.

#### 4.5.2 Hierarchinis 3 emocijų klasifikavimas

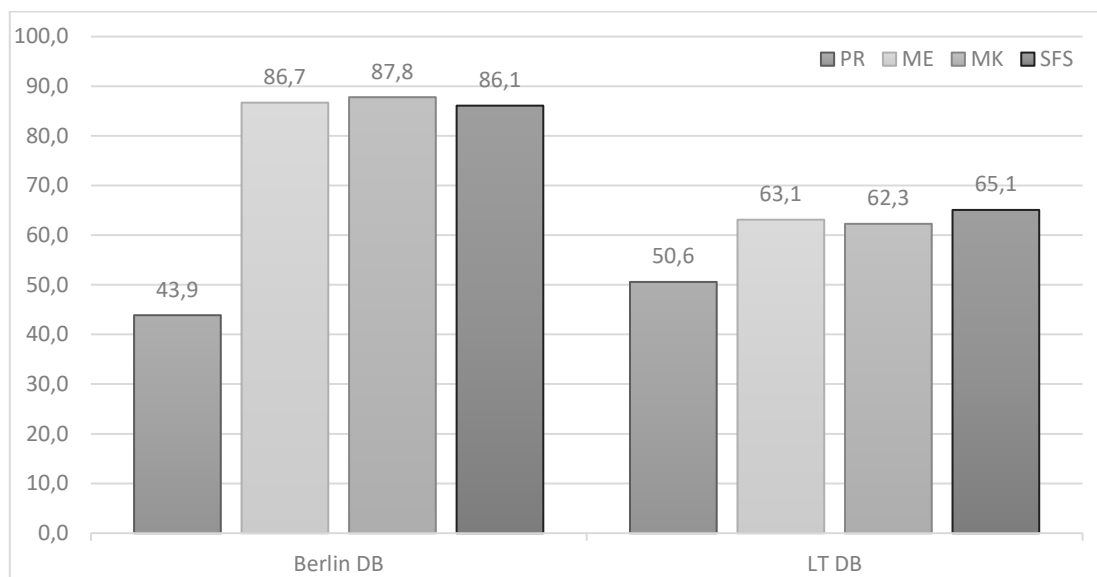
Šiame eksperimente buvo nagrinėjamas 3 emocijų klasifikavimo uždavinys. 23 pav. pateikiama tiriamoji hierarchinio klasifikavimo schema 3 emocijoms.



23 pav. Hierarchinio 3 emocijų klasifikavimo schema



Vidutiniai eksperimento rezultatai, naudojant skirtingus atrankos kriterijus, pateikti 24 pav.



**24 pav. Vidutinis 3 emocijų klasifikavimo tikslumas naudojant pilną požymių rinkinį (PR(TS)), Maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aibės didinimo (SFS) atrankos kriterijus**

Matome, jog skirtingai nuo ankstesniųjų eksperimentų rezultatų, šiuo atveju lietuvių kalbos emocijų klasifikavimo tikslumas žymiai mažesnis nei vokiečių kalbos (skirtumas svyruoja apie 20 proc.). Tokį pokytį galima paaiškinti išaugusia duomenų apimtimi (naudojama 1 000 kiekvienos emocijos pavyzdžių vietoj 300). Dėl duomenų apimties padidėjimo sumažėjęs klasifikavimo tikslumas rodo duomenų išsibarstymą (arba persidengimą klasifikavimo požimiūriu).

Esminių skirtumų tarp atrankos metodų negalima išvelgti – visi jie lėmė panašius klasifikavimo rezultatus. Palyginti su tiesioginės schemos rezultatais, matomas akivaizdus hierarchinio klasifikavimo pranašumas – tikslumo skirtumas siekia 15–40 proc. Pilno požymio rinkinio atveju gauti rezultatai – prasti – teisingai klasifikuota tik maždaug kas antra emocija.

2 lentelėje pateikti individualūs kiekvienos emocijos klasifikavimo rezultatai.

2 lentelė. Individualūs 3 emocijų klasifikavimo rezultatai

Bazė	Kriterijus	Klasifikavimo tikslumas, %		
		Pyktis	Džiaugsmas	Neutrali
Berlin DB	ME	81,7	81,7	96,7
	MK	85	81,7	96,7
	SFS	86,7	75	96,7
LT DB	ME	56,6	67,3	65,5
	MK	54	62,6	70,3
	SFS	57,6	71,7	66

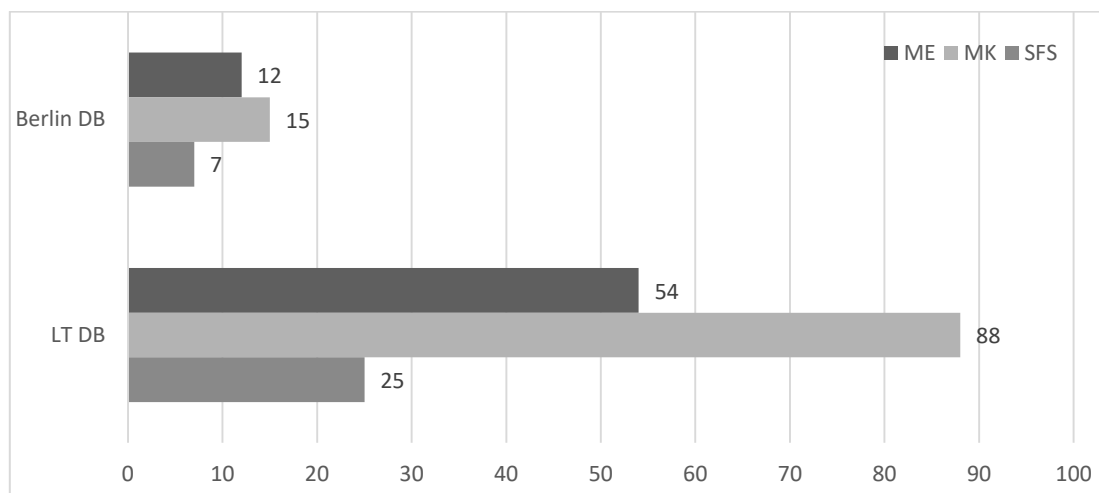
Lentelės duomenys rodo, jog tiksliausiai klasifikuojama emocija buvo neutrali būseną abiejų duomenų bazių atveju. Didžiausias klaidų kiekis gautas klasifikuojant džiaugsmą vokiečių kalbos atveju ir pyktį – lietuvių kalbos atveju. Galima to priežastis – skirtinga emocijų vaidyba. Vokiečių kalbos atveju pyktis išreiškiamas itin pakeltu balsu ir tonu, lietuvių kalbos atveju – pakitusia šnekos prozodija.

3 lentelė. Lietuvių kalbos emocijų klasifikavimo klaidų matrica

Emocijos	Atpažinta, %		
	Pyktis	Džiaugsmas	Neutrali
Pyktis	57,6	32,1	10,3
Džiaugsmas	15,1	71,7	13,2
Neutrali	18,3	15,7	66

Klasifikavimo klaidų analizei pateikiame klaidų matricą (3 lentelė). Analizei pasirinkome lietuvių šnekos emocijų klasifikavimą naudojant SFS atranką atveji. Kaip matome, dažniausia klaida – pyktis klasifikuojamas kaip džiaugsmas. Šios emocijos akustiniu požiūriu yra panašiausios, todėl ir gaunamas toks didelis klasifikavimo klaidų kiekis.

25 pav. pateikiami atrinktų požymių vektorių dydžiai skirtingiems atrankos kriterijams.



25 pav. Atrankos metodų grąžinamas vektorių dydis 3 emocijų klasifikavimo atveju

Grafike matome, jog lietuvių kalbos atveju gaunami 4–6 kartus didesni požymių rinkiniai. To priežastis – gerokai didesnis emocijų pavyzdžių kiekis. Lyginant atrankos kriterijus, priešingai nei ankstesniame eksperimente, mažiausias požymių vektorius gaunamas naudojant SFS atranką, didžiausias – MK (minimalios koreliacijos). To priežastimi galime įvardyti hierarchinį klasifikavimo modelį.

Didžiąją dalį atrinktųjų požymių sudarė melų skalės amplitudės spektro (MSAS) požymiai (tiek pirmajame, tiek antrajame klasifikavimo lygmenyje). Šių požymių dalis rinkiniuose kito nuo 40 proc. (taikant MK atranką vokiečių kalbos atveju) iki 100 proc. (taikant SFS atranką vokiečių kalbos atveju). Be šių požymių į rinkinius taip pat buvo atrinkti signalo energijos logaritmo, MSK, amplitudės spektro nuolydžio, pagrindinio tono dažnio ir jo kontūro požymiai bei jų statistiniai parametrai.

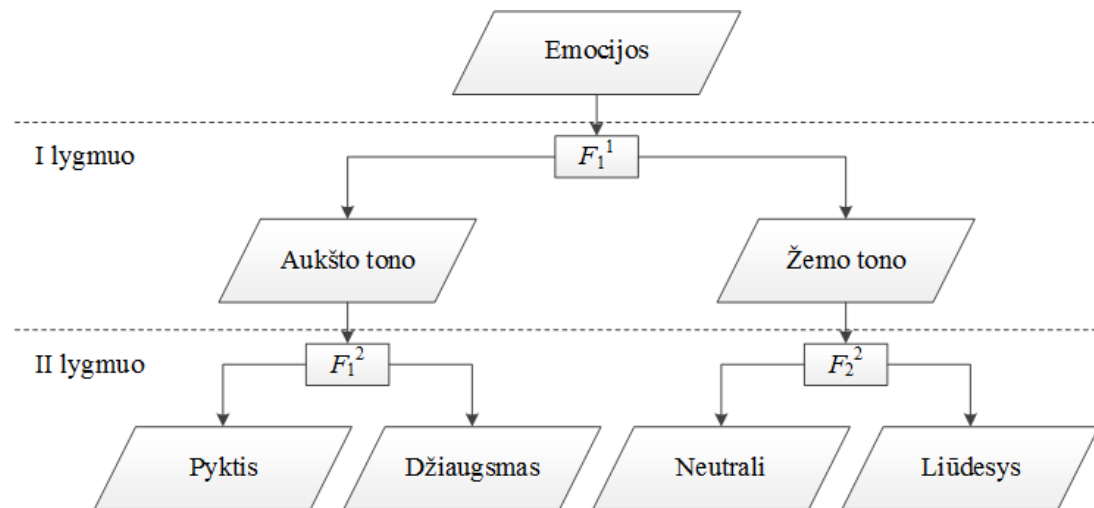
Požymių persidengimo tarp atskirų klasifikavimo lygmenų neužfiksuota, visuose lygmenyse buvo suformuoti skirtingi požymių rinkiniai. Tai liudija gerai atskirtus klasifikavimo uždavinius skirstant emocijas į žemo ir aukšto tono lygio grupes.

Apibendrinant 3 emocijų klasifikavimo rezultatus galima teigti, jog nors klasifikavimo tikslumu nė vienas atrankos kriterijus nebuvo dominuojantis, hierarchinis klasifikavimas (naudojant bet kurį atrankos kriterijų) gerokai lenkė

tiesioginį emocijų klasifikavimą. Dažniausia klasifikavimo klaida – atskiriant pyktį nuo džiaugsmo. Mažiausias gautas požymių rinkinys – naudojant SFS atranką.

### 4.5.3 Hierarchinis 4 emocijų klasifikavimas

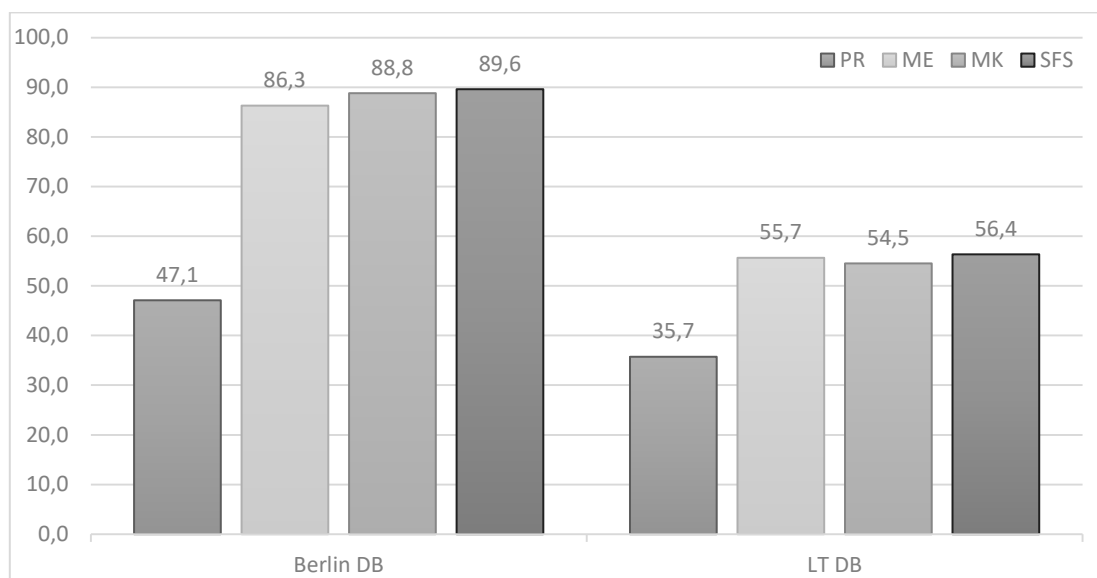
26 pav. pateikiama tiriamoji hierarchinė schema, skirta 4 emocijoms klasifikuoti.



26 pav. Hierarchinio 4 emocijų klasifikavimo schema

Emocijoms klasifikuoti teks atlikti tris klasifikavimo procesus dvejuose lygmenyse. Vidutiniai emocijų klasifikavimo rezultatai visiems trims kriterijams pateikti 27 pav.

Vėlgi visų trijų požymių atrankos atvejais gautieji klasifikavimo tikslumo rezultatai yra labai panašūs ir nė vienas iš kriterijų negalėtų būti išskirtas. Vokiečių kalbos emocijų rezultatai išliko panašūs į 3 emocijų uždavinio rezultatus. Tuo tarpu lietuvių kalbos emocijų klasifikavimo tikslumas sumažėjo ir svyravo apie 55 proc. lygį. Bet tai buvo 20 proc. daugiau negu tikslumas, gaunamas tiesioginės klasifikavimo schemos atveju.



27 pav. Vidutinis 4 emocijų klasifikavimo tikslumas naudojant pilną požymių rinkinį (PR(TS)), Maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aišės didinimo (SFS) atrankos kriterijus

Tai tik dar kartą patvirtina išvadą apie didesnę lietuvių kalbos emocijų požymių išsibarstymą. Klasifikavimo uždaviniui tampant sudėtingesniu, t. y. augant nagrinėjamų klasių skaičiui, požymių išsibarstymas lemia vis prastesnius klasifikavimo rezultatus.

4 lentelė. Individualūs 4 emocijų klasifikavimo rezultatai

Bazė	Kriterijus	Klasifikavimo tikslumas, %			
		Pyktis	Džiaugsmas	Neutrali	Liūdesys
Berlin DB	ME	80	76,7	91,7	96,7
	MK	85	81,7	93,3	95
	SFS	88,3	73,3	98,3	98,3
LT DB	ME	53,5	62,2	50,1	56,8
	MK	49,6	58,3	46,4	63,8
	SFS	52	67,6	48	57,8

4 lentelėje pateikti individualūs kiekvienos emocijos klasifikavimo rezultatai.

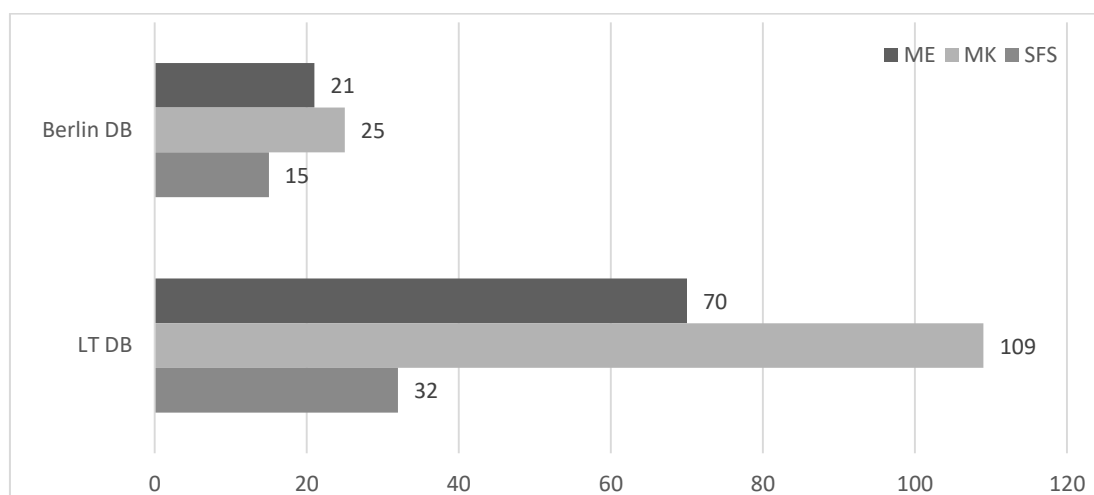
Galime matyti, jog vokiečių kalbos atveju tiksliausiai klasifikuojamas liūdesys, lietuvių kalbos atveju – džiaugsmas ir liūdesys. Didžiausias klaidų kiekis gautas klasifikuojant džiaugsmą vokiečių kalbos atveju ir neutralią būseną lietuvių kalbos atveju. Taigi rezultatai skiriasi nuo 3 emocijų uždavinio – papildomai prisidėjusi emocija iš esmės pakeičia emocijų požymių persidengimą. Detalesnis klaidų vaizdas matosi 5 lentelėje.

5 lentelė. Lietuvių kalbos emocijų klasifikavimo klaidų matrica

	Atpažinta, %			
	Pyktis	Džiaugsmas	Neutrali	Liūdesys
Pyktis	52	36	6,2	5,8
Džiaugsmas	19,1	67,5	11,1	2,2
Neutrali	15,6	10,4	48	26
Liūdesys	3,7	7,1	31,4	57,8

Kaip ir 3 emocijų uždavinyje pykčio atskyrimas nuo džiaugsmo lemia bene daugiausiai klaidų. Su papildoma emocija klasifikavimo uždavinyje atsirado ir naujas klaidų šaltinis – liūdesio ir neutralios būsenos pora. Klaidų lygis atskiriant šias emocijas siekė 26–30 proc. Taigi naudojami akustiniai požymiai yra persidengę šių dviejų emocijų požiūriu.

28 pav. pateikiami gautų požymių vektorių dydžių priklausomybė nuo atrankos metodo.



28 pav. Atrankos metodų grąžinamas vektorių dydis 4 emocijų klasifikavimo atveju

SFS atrankos metodo gražinamas požymių vektorius vėlgi gautas mažiausias – 15 ir 32 vokiečių ir lietuvių kalbos atvejais. Minimalios koreliacijos kriterijus, kaip ir 3 emocijų atveju, lėmė didžiausią požymių vektorių. Skirtumas tarp mažiausio ir didžiausio požymių skaičiaus siekė 7. Dėl duomenų kiekio padidėjimo, požymių vektoriaus dydis pasikeitė 2–4 kartus – šiek tiek mažiau nei 3 emocijų atveju.

Kaip ir 3 emocijų atveju žymių rinkinių dalį sudarė MSAS požymiai. Jų dalis rinkiniuose kito nuo 29 proc. (taikant MK atranką lietuvių kalbos atveju) iki 73 proc. (taikant SFS atranką vokiečių kalbos atveju). Kiti į rinkinius įtraukti požymiai buvo analogiški 3 emocijų uždavinio atvejui.

Požymių persidengimo tarp atskirų klasifikavimo lygmenų taip pat neužfiksuota.

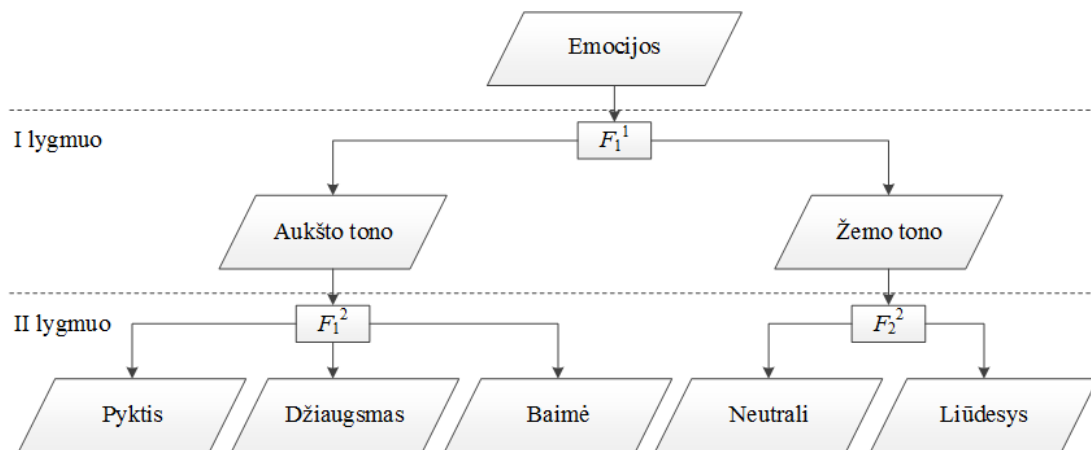
Apibendrinami 4 emocijų klasifikavimo rezultatus, galime teigti, jog vėl nėra vienas atrankos kriterijus nebuvo dominuojantis, o hierarchinis klasifikavimas buvo tikslesnis nei tiesioginis klasifikavimas. Mažiausias gautas požymių vektorius – naudojant SFS atranką.

#### **4.5.4 Hierarchinis 5 emocijų klasifikavimas**

Nagrinėjant 5 emocijų atvejį, prisidėjo penktoji emocija – baimė. Pagrindinio tono dažnio reikšmių analizė parodė, jog pagal tono dažnio kitimą baimė turėtų būti priskirta aukšto tono klasei. Tuomet žemo tono emocijų klasė apima neutralią būseną ir liūdesį, taigi šiuo atveju taip pat užteko dviejų klasifikavimo lygmenų.

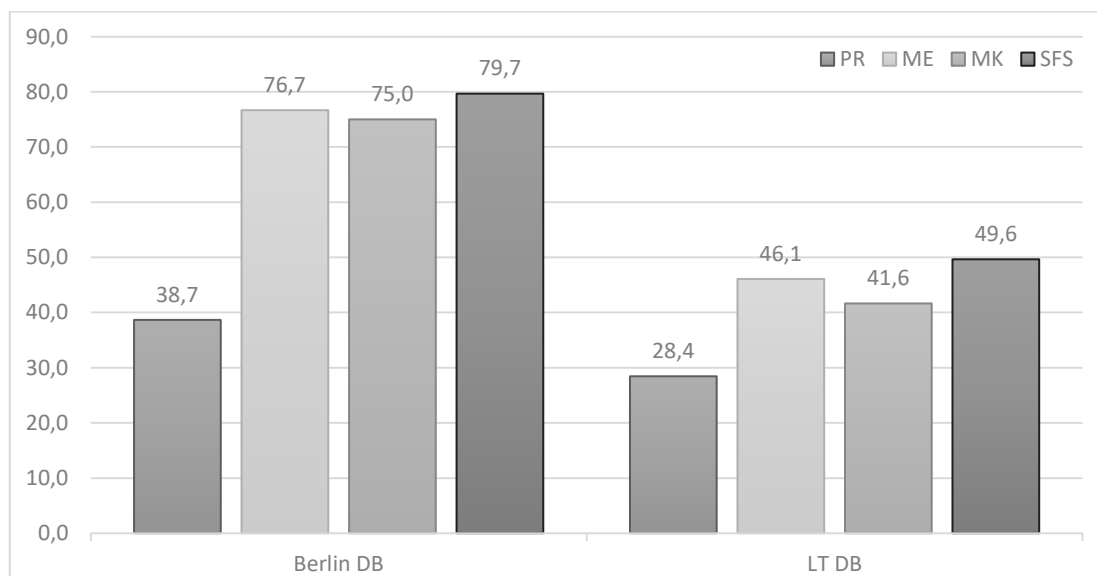
29 pav. pateikiama tiriamoji hierarchinė schema, skirta 5 emocijoms klasifikuoti.

Vidutiniai emocijų klasifikavimo rezultatai visiems trimis kriterijams pateikti 30 pav.



29 pav. Hierarchinio 5 emocijų klasifikavimo schema

5 emocijų atveju, nors ir nežymiai, tačiau geriausius rezultatus lėmė SFS atrankos metodas – jo rezultatai buvo 3–8 proc. geresni negu kitų dviejų atrankos metodų. Palyginti su 4 emocijų uždaviniu, bendras vidutinis tikslumas sumažėjo apie 10 proc., taigi sudėtingėjant klasifikavimo uždaviniui, bendras tikslumas, kaip ir galima tikėtis, mažėja. Lietuvių kalbos atveju tikslumas tesiekė 40–50 proc. ir buvo pakankamai žemas, palyginti su vokiečių kalbos atveju. Augantis klasifikavimo uždavinio sudėtingumas išryškino skirtumą tarp nagrinėtų vokiečių ir lietuvių kalbų emocijų bazių – pirmoji pasižymi daug didesniu atsparumu nagrinėjamų emocijų skaičiui.



30 pav. Vidutinis 5 emocijų klasifikavimo tikslumas naudojant pilną požymių rinkinį (PR(TS)), Maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aibės didinimo (SFS) atrankos kriterijus



Nagrinėdami kiekvienos emocijos klasifikavimo tikslumą (6 lentelė), galime matyti, jog vokiečių kalbos atveju tiksliausiai klasifikuojama neutrali būseną ir liūdesys, lietuvių kalbos atveju – liūdesys. Taigi šiame eksperimente abejoms kalboms gauti panašūs rezultatai. Mažiausias tikslumas gautas džiaugsmo (vokiečių k.), baimės bei neutralios būsenos (lietuvių k.) atvejais.

6 lentelė. Individualūs 5 emocijų klasifikavimo rezultatai

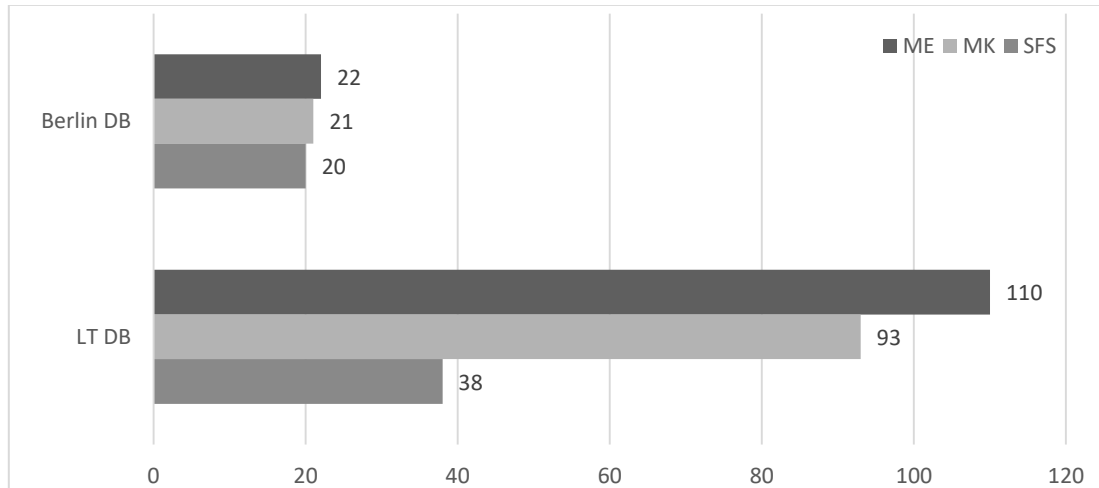
Bazė	Kriterijus	Klasifikavimo tikslumas, %				
		Pyktis	Džiaugsmas	Neutrali	Liūdesys	Baimė
Berlin DB	ME	80	55	91,7	93,3	63,3
	MK	85	58,3	85	75	71,7
	SFS	81,7	63,3	90	90	73,3
LT DB	ME	42,7	48,1	45,1	49,3	45,2
	MK	42,4	38,1	40,3	52,8	34,6
	SFS	49,4	60,6	41,2	50,3	46,7

Klasifikavimo klaidos panašios į ankstesniųjų eksperimentų metu gautas klaidas (7 lentelė). Didžiausias persidengimas pasireiškė tarp džiaugsmo ir pykčio bei neutralios būsenos ir liūdesio porų. Didžiausias atskiriamumas gautas tarp pykčio ir liūdesio poros.

7 lentelė. Lietuvių kalbos emocijų klasifikavimo klaidų matrica

	Atpažinta, %				
	Pyktis	Džiaugsmas	Neutrali	Liūdesys	Baimė
<b>Pyktis</b>	49,4	29,6	12,4	6,9	1,7
<b>Džiaugsmas</b>	16,4	60,6	8,5	6,1	8,4
<b>Neutrali</b>	7,2	10,9	41,2	33,5	7,2
<b>Liūdesys</b>	1,1	5,5	35,7	50,3	7,4
<b>Baimė</b>	4,7	5,1	33,4	10,1	48,7

31 pav. matome, jog SFS atranka lėmė mažiausią požymių rinkinį 5 emocijoms klasifikuoti. Didžiausią – ME atranka, tačiau iš esmės ji nedaug atsiliko nuo MK atrankos.



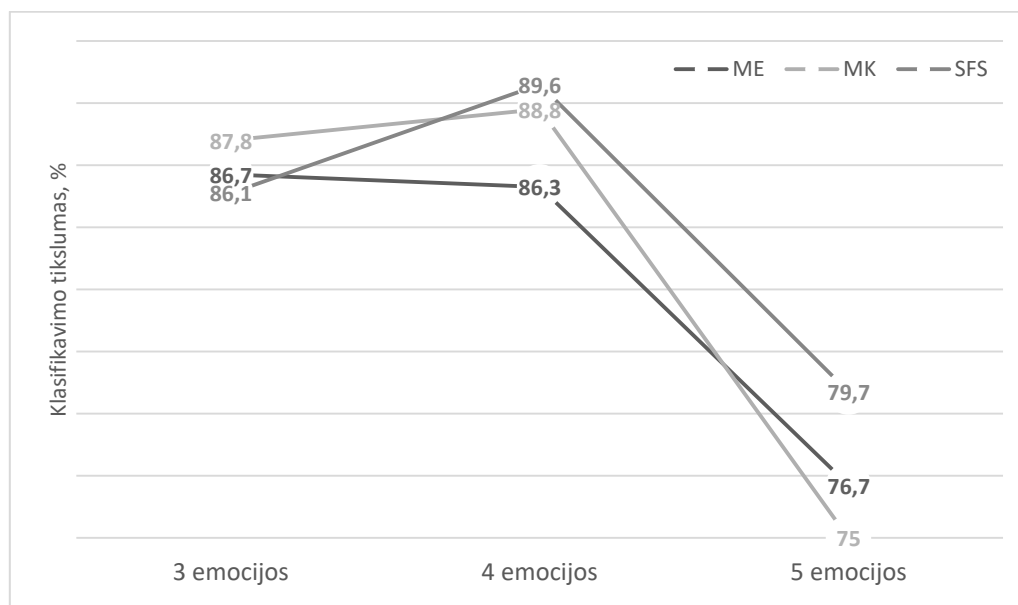
31 pav. Atrankos metodų grąžinamas vektorių dydis 5 emocijų klasifikavimo atveju

Nagrinėjant 5 emocijas požymių rinkiniai iš esmės pasikeitė – juose pradėjo dominuoti kiti požymiai. Vokiečių kalbos atveju ME atranka lėmė amplitudės spektro juostų (ASJ) požymių dominavimą (požymiai sudarė 45 proc. viso rinkinio), MK atranka – melų skalės kepstro (MSK) požymių dominavimą (33 proc. viso rinkinio). Lietuvių kalbos atveju ME atranka grąžino požymių rinkinį, kurio 38 proc. sudarė MSAS požymiai, 22 proc. – MSK požymiai, dar 22 proc. – amplitudės spektro nuolydžio požymiai. Pagal MK kriterijų atrinktą požymių rinkinį sudarė MSAS (45 proc. viso rinkinio), MSK (18 proc.) bei ASJ (18 proc.) požymiai. Tai rodo, jog nagrinėjant didesnę emocijų skaičių, hierarchinė klasifikavimo schema lemia gerokai sudėtingesnę požymių rinkinį.

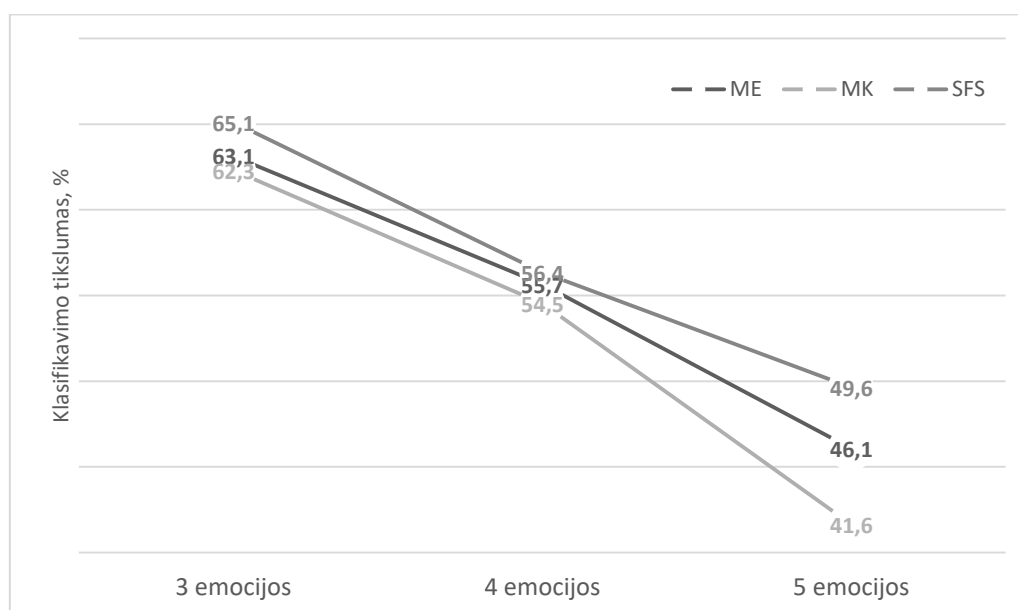
Sudėtingesnis klasifikavimo uždavinys lėmė ir požymių persidengimą tarp skirtingų lygmenų. Lietuvių kalbos atveju taikant ME atranką, 2 (iš 59) pirmojo klasifikavimo lygmens požymiai pasikartojo sekančiuose lygmenyse. Tai sudarė tik 3,4 proc. pirmojo lygmens rinkinio, taigi toks persidengimas

nepaneigia aukščiau suformuluoto teiginio apie gerai atskirtus klasifikavimo uždavinio etapus.

3, 4 ir 5 emocijų klasifikavimo eksperimentų rezultatams apibendrinti pateikiame vidutinio klasifikavimo tikslumo ir požymių vektoriaus dydžio priklausomybes nuo nagrinėjamų emocijų skaičiaus.



**32 pav. Vidutinio vokiečių kalbos emocijų klasifikavimo tikslumo priklausomybė nuo nagrinėjamų emocijų skaičiaus naudojant maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aibės didinimo (SFS) atrankos kriterijus**

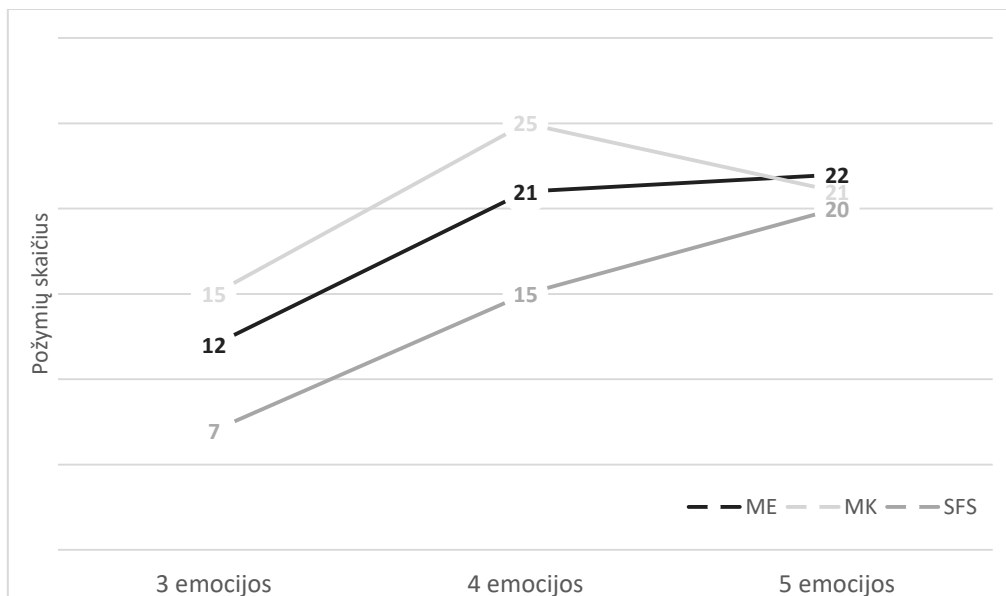


**33 pav. Vidutinio lietuvių kalbos emocijų klasifikavimo tikslumo priklausomybė nuo nagrinėjamų emocijų skaičiaus naudojant maksimalaus efektyvumo (ME), minimalios tarpusavio koreliacijos (MK) ir nuoseklaus aibės didinimo (SFS) atrankos kriterijus**

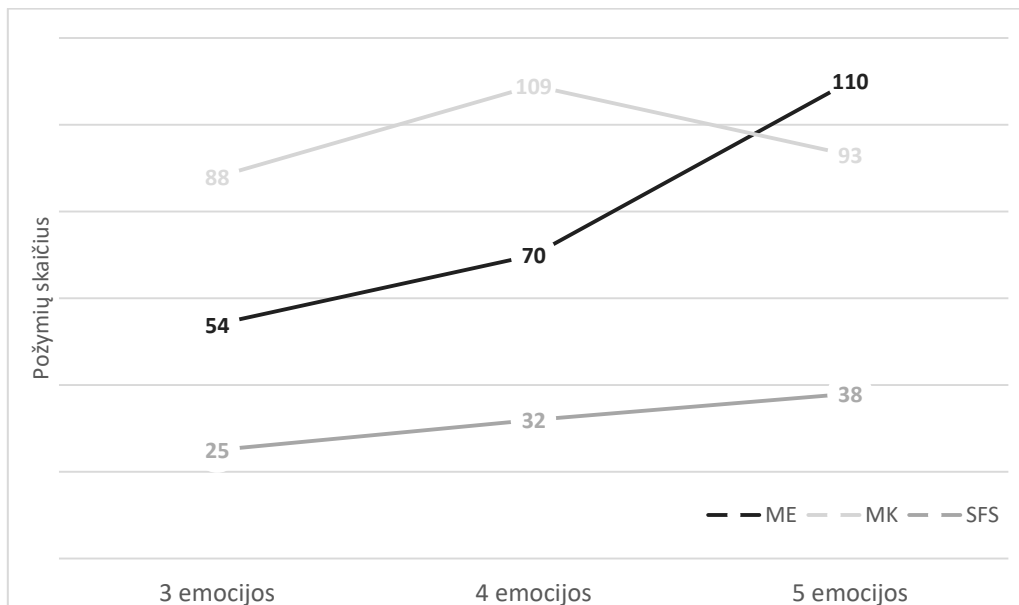
32 pav. ir 33 pav. matome vokiečių ir lietuvių kalbos emocijų klasifikavimo tikslumo priklausomybę nuo emocijų skaičiaus. Juos nagrinėjant išryškėja dvi tendencijos:

- augant nagrinėjamų emocijų skaičiui, bendras vidutinis klasifikavimo tikslumas mažėja;
- nagrinėjamų pavyzdžių kiekis taip pat turi įtakos bendram tikslumui. Didesnis pavyzdžių kiekis gali lemti mažesnę bendrą klasifikavimo tikslumą. To priežastis – didesnis emocijų požymių išsibarstymas ir persidengimas, gaunamas dėl didesnio pavyzdžių kiekio.

34 pav. ir 35 pav. pateikta požymių rinkinių dydžio priklausomybė nuo nagrinėjamų emocijų skaičiaus.



34 pav. Atrankos metodų grąžinamo vektorių dydžio priklausomybė nuo nagrinėjamų emocijų (vokiečių kalba) skaičiaus



35 pav. Atrankos metodų grąžinamo vektorių dydžio priklausomybė nuo nagrinėjamų emocijų (lietuvių kalba) skaičiaus

Šiose priklausomybėse galima pastebėti tokias tendencijas:

- augant nagrinėjamų emocijų kiekiui, klasifikavimo tikslumui maksimizuoti reikalingų požymių skaičius rinkinyje kai kuriais atvejais gali didėti;
- augant nagrinėjamų pavyzdžių kiekiui, požymių vektoriaus dydis taip pat auga. Tai susiję su didėjančiu klasifikavimo uždavinio sudėtingumu;
- mažiausias požymių skaičiaus rinkinyje pokytis fiksuojamas SFS atrankos metodui.

Vertinant požymių rinkinių sudėties priklausomybę nuo nagrinėjamų emocijų skaičiaus, galima išvelgti nuoseklų rinkinių didėjimą augant emocijų skaičiui. Pavyzdžiui, vokiečių kalbos emocijų atveju taikant SFS atranką, 57 proc. požymių, naudojamų 3 emocijoms klasifikuoti, patenka į rinkinį 4 emocijoms klasifikuoti, o 27 proc. požymių, atrinktų 4 emocijoms klasifikuoti, patenka į 5 emocijų klasifikavimo rinkinį. SFS atranką taikant lietuvių kalbos emocijų požymių analizei, 62 proc. viso 3 emocijoms klasifikuoti sudaryto rinkinio patenka į 4 emocijų klasifikavimo rinkinį, o 36 proc. 4 emocijoms skirtą rinkinio – į 5 emocijoms klasifikuoti suformuotą rinkinį. Šiek tiek mažesni, bet analogiški rezultatai užfiksuoti ir kitų atrankos

metodų atvejais. Taigi net ir naudojant suboptimalius požymių atrankos metodus, rinkinių formavimas yra pakankamai kryptingas ir nėra visiškai atsitiktinis.

#### **4.5.5 Alternatyvios hierarchinės schemas**

Aukščiau eksperimentiškai nagrinėtos hierarchinės schemas efektyvumui įvertinti papildomai atliksime dar vieną eksperimentinį tyrimą, kurio tikslas – palyginti alternatyvias hierarchines emocijų klasifikavimo schemas.

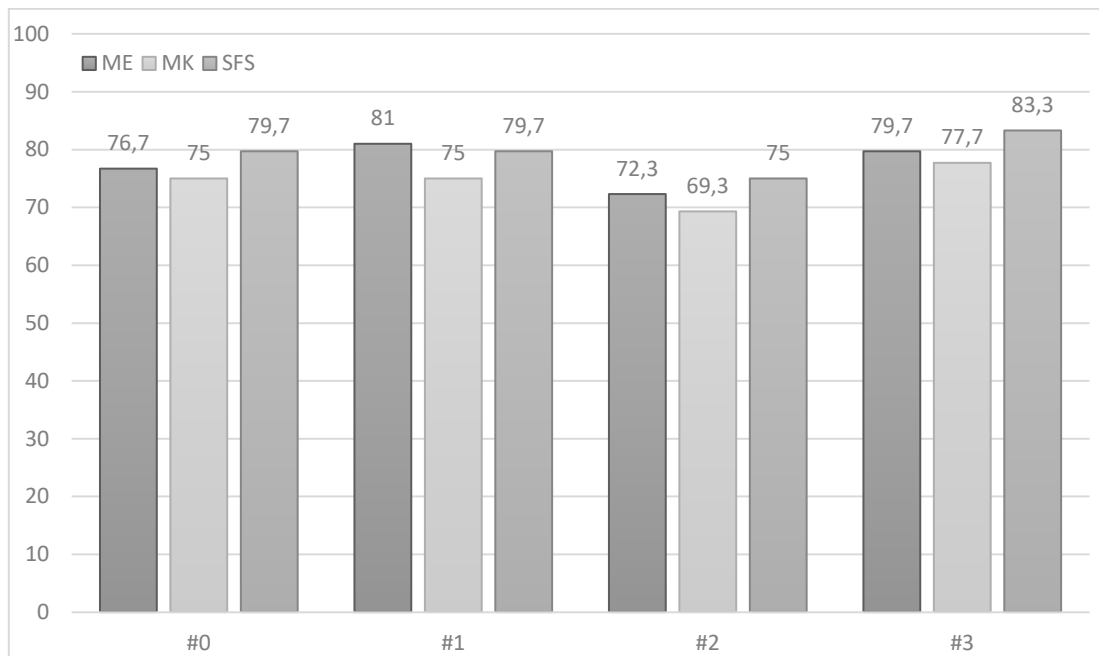
Palyginimui pasirinkome 5 emocijų uždavinį su vokiečių kalbos įrašais (kadangi kiti autoriai taip pat naudoja šios bazės įrašus). Eksperimentiškai palyginome tokias eksperimentines schemas:

- Schema #0. Tai mūsų 4.5.4. skyrelyje nagrinėtoji emocijų klasifikavimo schema.
- Schema #1. Tai mūsų jau tirtosios emocijų klasifikavimo schemas modifikacija. Joje aukšto tono emocijoms priskiriamas pyktis ir džiaugsmas, žemo tono – neutrali būseną, liūdesys ir baimė. Gaunama 2 lygių hierarchinė schema, panaši į nagrinėtą skyrelyje 4.5.4.
- Schema #2. Ši schema sudaryta remiantis darbu (Kotti & Paterno, 2012). Neutrali būseną ir džiaugsmas priskirti teigiamo valentingumo klasei, o pyktis, liūdesys ir baimė – neigiamos valentingumo klasei. Pastaroji antrajame lygmenyje skirstoma į teigiamos aktyvacijos (pyktis, baimė) ir neigiamos aktyvacijos (liūdesys) klases. Vėlgi gauname 2 lygių hierarchinę klasifikavimo schemą.
- Schema #3. Pagal darbą (Lugger, et al., 2009 ) sudaryta schema, kurioje visos 5 emocijos pirmajame lygyje skirstomos į aukšto aktyvumo (pyktis, džiaugsmas, baimė) ir žemo aktyvumo (neutrali būseną ir liūdesys) klases. Antrajame lygyje pirmoji klasė skirstoma į stiprių emocijų (pyktis ir džiaugsmas, kurios atskiriamos trečiajame lygyje) ir silpnų emocijų (baimė) klases. Žemo aktyvumo emocijos irgi

klasifikuojamos į stiprias (neutrali būseną) ir silpnas (liūdesys) emocijas.

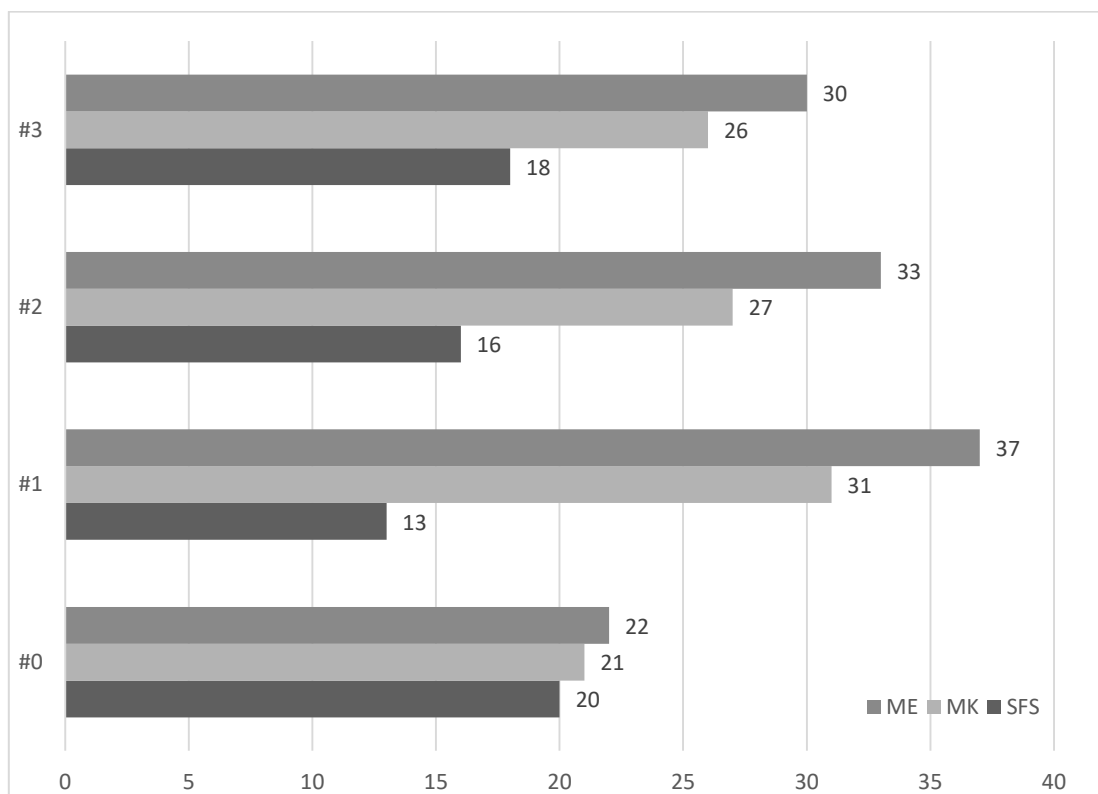
Visos schemas iš esmės skiriasi emocijų grupavimu. Lygindami klasifikavimo, naudojant šias schemas rezultatus, mes galėsime įvertinti pasiūlytos hierarchinės schemas atsparumą emocijų skirstymui į klases.

Klasifikavimo eksperimentas įvykdytas antrajam eksperimentui analogiškomis sąlygomis: kryžminis įvertinimas skirstant duomenis į 3 dalis ir 5 artimiausių kaimynų klasifikatorius. Požymių atrankai panaudoti visi trys metodai.



**36 pav. Alternatyvių hierarchinių schemų vidutiniai emocijų klasifikavimo rezultatai**

Vidutiniai klasifikavimo tikslumo rezultatai pateikti 36 pav. Matome, jog nė viena tirtoji schema nebuvo iš esmės pranašesnė už kitas. Tarkime, schema #3 su SFS atranka buvo 3–8 proc. pranašesnė už kitas, tačiau ta pati su ME atranka nusileido Schemai #2. Taigi galime teigti, jog pasiūlyta hierarchinė schema, nepriklausomai nuo pritaikyto emocijų grupavimo, leido pasiekti panašų vidutinį klasifikavimo tikslumą.



**37 pav. Alternatyvių hierarchinių schemų vidutiniai emocijų klasifikavimo rezultatai**

Taip pat neišsiskyrė ir nė vienas atrankos metodas – visi rodė daugiau mažiau panašius emocijų klasifikavimo rezultatus, skirtumas tarp jų svyravo 5–8 proc.

37 pav. pateiktos gautų požymių rinkinių dydžių reikšmės tirtosioms schemoms.

Matome, jog visų schemų atveju mažiausias požymių rinkinys buvo gaunamas naudojant SFS atranką – požymių skaičius kito nuo 13 iki 20. Didžiausi požymių rinkinio svyravimai gauti ME atveju – jų skaičius kito 22–37 ribose.

Klaidų matricių analizė parodė, jog daugiausia klaidų buvo padaroma analogiškai ankstesniojo eksperimento rezultatams – atskiriant pykčio ir džiaugsmo būsenas. Taigi šias emocijas galime laikyti sunkiausiai atskiriamomis akustiniu požiūriu.



## 4.6 Skyriaus apibendrinimas

Atlikę hierarchinės 3, 4 ir 5 emocijų klasifikavimo schemų tyrimus, jų rezultatus apibendriname:

- Hierarchinis šnekos emocijų klasifikavimas, pagrįstas požymių atranka, leidžia pasiekti gerokai aukštesnį klasifikavimo tikslumą nei tiesioginis klasifikavimas naudojant pilnus požymių rinkinius.
- Pasiūlytoji hierarchinio šnekos emocijų klasifikavimo schema pasižymi mažu požymių persidengimu tarp atskirų klasifikavimo lygmenų ir pakankamai dideliu požymių persidengimu tarp skirtingo emocijų skaičiaus uždavinių.
- Akivaizdaus pranašumo suformuluotų požymių atrankos metodų nepastebėta. Tiek maksimalaus efektyvumo, tiek minimalios koreliacijos, tiek nuoseklaus aibės didinimo atrankos metodai lėmė panašius rezultatus.
- Didėjant nagrinėjamų emocijų skaičiui, vidutinis emocijų klasifikavimo tikslumas mažėja, o klasifikavimo tikslumui maksimizuoti reikalingas požymių kiekis didėja.
- Pasiūlyta hierarchinė emocijų klasifikavimo schema pademonstravo nepriklausomybę nuo pasirinkto emocijų grupavimo. Eksperimentiškai ištirtų 3 alternatyvių klasifikavimo schemų vidutiniai rezultatai buvo panašūs.

## Bendros išvados

Disertacijoje išnagrinėtas emocijų atpažinimo šnekoje uždavinys, išanalizuotos hierarchinio emocijų klasifikavimo schemas. Analitinėje dalyje emocijoms klasifikuoti pritaikyta sprendimų medžio struktūra ir pasiūlyta hierarchinio klasifikavimo schema.

Pagrindiniai rezultatai:

- Remiantis sprendimų medžių idėja, suformuota hierarchinio klasifikavimo schema, leidžianti nagrinėti pageidaujamą šnekos emocijų kiekį. Pasiūlytos schemas struktūra priklauso nuo pasirinktojo emocijų grupavimo.
- Požymių rinkiniams sudaryti suformuluoti maksimalaus efektyvumo ir minimalios koreliacijos atrankos metodai, pritaikytas nuoseklus aibės didinimo metodas. Visi atrankos metodai pritaikyti hierarchinei klasifikavimo schemai.
- Atliktas pasiūlytos klasifikavimo schemas eksperimentinis tyrimas su emocijų vokiečių kalbos įrašais. Atlikti didelės apimties eksperimentai su suvaidintų emocijų lietuvių kalba įrašais.

Atlikus eksperimentinius tyrimus, suformuluotos tokios išvados bei jas patvirtinantys eksperimentiniai rezultatai:

- Hierarchinė emocionalios šnekos klasifikavimo schema yra iš esmės efektyvesnė už tiesioginį (vieno etapo) klasifikavimą tikslumo požiūriu.
  - Eksperimentų metu gautas hierarchinės schemas vidutinis klasifikavimo tikslumas iki 40 proc. didesnis nei tiesioginės vieno etapo klasifikavimo schemas.
  - Požymių persidengimas tarp atskirų hierarchinių lygmenų kito nuo 0 proc. iki 3,4 proc.
- Požymių atrankos taikymas leidžia reikšmingai sumažinti nagrinėjamų duomenų kiekį ir kartu padidinti klasifikavimo tikslumą, palyginti su pilnais požymių rinkiniais.

- Pasiūlytieji ir pritaikytieji atrankos metodai leido sumažinti požymių skaičių nuo 6 552 (pilno požymių rinkinio atveju) iki 9–38 požymių tiesioginės schemos atveju bei 7–110 požymių hierarchinės schemos atveju. Požymių rinkinių apimties požiūriu efektyviausias – nuoseklaus aibės didinimo metodas, kuris leido gauti iki 4 kartų mažesnės apimties požymių rinkinius, palyginti su kitais dviem atrankos metodais.
- Hierarchinėje klasifikavimo schemoje skirtingi požymių atrankos metodai neturi esminės įtakos visos schemos tikslumui.
  - Tyrime skirtingi atrankos metodai lėmė 1,5–5 proc. tikslumo skirtumus vokiečių kalbos atveju ir 2–8 proc. skirtumus lietuvių kalbos atveju.
- Didėjant nagrinėjamų emocijų skaičiui, vidutinis emocijų klasifikavimo tikslumas mažėja, o klasifikavimo tikslumui maksimizuoti reikalingas požymių rinkinio sudėtingumas bei dydis auga.
  - Nagrinėjamų emocijų skaičiui pakitus nuo 3 iki 5, požymių skaičius rinkiniuose padidėjo 1,5–2 kartus. Nagrinėjamam vienos emocijos pavyzdžių skaičiui padidėjus 5 kartus, požymių skaičius rinkiniuose eksperimentų metu padidėjo 2–6 kartus.
  - Nagrinėjamų emocijų skaičiui pakitus nuo 3 iki 4, gautųjų požymių rinkinių persidengimas siekė 62 proc., emocijų skaičiui pakitus nuo 4 iki 5 – 36 proc.

## Literatūros sąrašas

- [1] Adell, J., Bonafonte Cavez, A., Escudero Mancebo, D. (2005). Analysis of prosodic features towards modelling of emotional and pragmatic attributes of speech. *Procesamiento del lenguaje natural*, pp. 277–283.
- [2] Albornoz, E. M., Milone, D. H., Rufiner, H. L. (2011). Spoken emotion recognition using hierarchical classifiers. *Computer Speech & Language*, pp. 556–570.
- [3] Anagnostopoulos, C.-N., Iliou, T., Giannoukos, I. (2012). Features and classifiers for emotion recognition from speech: a survey from 2000 to 2011. *Artificial Intelligence Review*.
- [4] Arias, J., Busso, C., Yoma, N. (2014). Shape-based modeling of the fundamental frequency contour for emotion detection in speech. *Computer Speech & Language*, pp. 278–294.
- [5] Asad, K., Ahmed, T., Rahman, M. (2012). Movie popularity classification based on inherent movie attributes using C4.5, PART and correlation coefficient. *International Conference on Informatics, Electronics & Vision (ICIEV)*.
- [6] Ayadi, M., Kamel, M., Karray, F. (2011). Survey on speech emotion recognition: Features, classification schemes, and databases. *Pattern Recognition*, pp. 572–587.
- [7] Azhagusundari, B., Selvadoss Thanamani, A. (2013). Feature Selection based on Information Gain. *International Journal of Innovative Technology and Exploring Engineering*, pp. 18–21.
- [8] Bhargava, M., Polzehl, T. (2012). Improving Automatic Emotion Recognition from speech using Rhythm and Temporal feature. *In: ICECIT*, pp. 133–147.
- [9] Bitouk, D., Verma, R., Nenkova, A. (2010). Class-level spectral features for emotion recognition. *Speech Communication*, pp. 613–625.
- [10] Bramer, M. (2007). *Principles of Data Mining*. Springer London.
- [11] Breiman, L., Friedman, J., Stone, C., Olshen, R. (1984). *Classification and Regression Trees*. Chapman and Hall/CRC.

- [12] Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W., Weiss, B. (2005). A database of German emotional speech. *In: Proceedings of Interspeech*, pp. 1517–1520.
- [13] Busso, C., Bulut, M., Narayanan, S. (2012). Toward effective automatic recognition systems of emotion in speech. *Social emotions in nature and artifact: emotions in human and human-computer interaction*, pp. 1–23.
- [14] Casale, S., Russo, A. (2007). Multistyle classification of speech under stress using feature subset selection based on genetic algorithms. *Speech Communication*, pp. 801–810.
- [15] Chen, L., Mao, X., Wei, P., Xue, Y., Ishizuka, M. (2012). Mandarin emotion recognition combining acoustic and emotional point information. *Applied Intelligence*, pp. 602–612.
- [16] Chen, L., Mao, X., Xue, Y., Cheng, L. (2012). Speech emotion recognition: Features and classification models. *Digital Signal Processing*, pp. 1154–1160.
- [17] Chiou, B.-C., Chen, C.-P. (2013). Feature space dimension reduction in speech emotion recognition using support vector machine. *Signal and Information Processing Association Annual Summit and Conference*, pp. 1–6.
- [18] Cichosz, J., Slot, K. (2007). *Emotion recognition in speech signal using emotion-extracting binary decision trees*.
- [19] Cornelius, R. (1996). *The science of emotion : research and tradition in the psychology of emotions*. London: Prentice Hall.
- [20] Costa, E. P., Lorena, A. C., Carvalho, A. C., Freitas, A. A. (2007). A Review of Performance Evaluation Measures for Hierarchical Classifiers. *Evaluation Methods for Machine Learning II*, pp. 182–196.
- [21] Cowie, R., Cornelius, R. (2003). Describing the emotional states that are expressed in speech. *Speech Communication*, pp. 5–32.
- [22] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Taylor, J. (2001). Emotion recognition in human-computer interaction. *IEEE Signal Processing Magazine*, pp. 47–57.
- [23] Dash, M., Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis*, pp. 131–156.
- [24] Dellaert, F., Polzin, T., Waibel, A. (1996). Recognizing emotion in speech. *Spoken Language*, pp. 1970–1973.

- [25] Duda, R., Hart, P., Stark, D. (2000). *Pattern Classification*. Wiley-Interscience.
- [26] Dudani, S. (1976). The Distance-Weighted k-Nearest-Neighbor Rule. *IEEE Transactions on Systems, Man, and Cybernetics*, pp. 325–327.
- [27] Eyben, F., Weninger, F., Gross, F., Schuller, B. (2013). Recent developments in openSMILE, the munich open-source multimedia feature extractor. *Proceedings of the 21st ACM international conference on Multimedia*, pp. 835–838.
- [28] Eyben, F., Wollmer, M., Schuller, B. (2009). OpenEAR – Introducing the Munich Open-Source Emotion and Affect Recognition Toolkit. *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–6.
- [29] Gadhe, R., Deshmukh, R., Waghmare, V. (2015). KNN based emotion recognition system for isolated Marathi speech. *International Journal of Computer Science Engineering*, 4, pp. 173–177.
- [30] Garg, V., Kumar, H., Sinha, R. (2013). *Speech based Emotion Recognition based on hierarchical decision tree with SVM, BLG and SVR classifiers*.
- [31] Gharavian, D., Sheikhan, M., Nazerieh, A., Garoucy, S. (2012). Speech emotion recognition using FCBF feature selection method and GA-optimized fuzzy ARTMAP neural network. *Neural Computing and Applications*, pp. 2115–2126.
- [32] Giannoulis, P., Potamianos, G. (2012). A hierarchical approach with feature selection for emotion recognition from speech. *In: Proceedings of the Eighth International Conference on Language Resources and Evaluation*, pp. 1203–1206.
- [33] Gjoreski, M., Gjoreski, H. (2014). Machine Learning Approach for Emotion Recognition in Speech. *Informatica*, pp. 377–384.
- [34] Guyon, I., Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research*, pp. 1157–1182.
- [35] Hansen, J., Patil, S. (2007). Speech Under Stress: Analysis, Modeling and Recognition. *Speaker Classification I*, pp. 108–137.
- [36] Hernandez, J., Enrique Sucar, L., Morales, E. (2014). Multidimensional hierarchical classification. *Expert Systems with Applications*, pp. 7671–7677.

- [37] Hyun, K. H., Kim, H. E., Kwak, K. Y. (2007). Emotional Feature Extraction Based On Phoneme Information for Speech Emotion Recognition. *The 16th IEEE International Symposium on Robot and Human interactive Communication*, pp. 802–806.
- [38] Jang, K.-D., Kwon, O.-W. (2006). Speech Emotion Recognition for Affective Human-Robot Interaction. *SPECOM*, pp. 419–422.
- [39] Khan, M., Goskula, T., Nasiruddin, M., Quazi, R. (2011). Comparison between k-NN and SVM method for speech emotion recognition. *International Journal on Computer Science and Engineering*, 3, pp. 607–611.
- [40] Kohavi, R. (1995). A Study of CrossValidation and Bootstrap for Accuracy Estimation and Model Selection. *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pp. 1137–1143.
- [41] Koller, D., Sahami, M. (1997). Hierarchically Classifying Documents Using Very Few Words. *ICML '97 Proceedings of the Fourteenth International Conference on Machine Learning*, pp. 170–178.
- [42] Koolagudi, S., Rao, K. (2012). Emotion recognition from speech using source, system, and prosodic features. *International Journal of Speech Technology*, pp. 265–289.
- [43] Koolagudi, S., Rao, S. (2012). Emotion recognition from speech: a review. *International Journal of Speech Technology*, 15(2), pp. 99–117.
- [44] Koolagudi, S., Reddy, R., Rao, K. (2010). Emotion recognition from speech signal using epoch parameters. *International Conference on Signal Processing and Communications*, pp. 1–5.
- [45] Kotti, M., Paterno, F. (2012). Speaker-independent emotion recognition exploiting a psychologically-inspired binary cascade classification schema. *International Journal of Speech Technology*, pp. 131–150.
- [46] Kumar, V., Minz, S. (2014). Feature selection: A literature review. *Smart Computing Review*, 4(3).
- [47] Lanjewar, R., Mathurkar, S., Patel, N. (2015). Implementation and Comparison of Speech Emotion Recognition System Using Gaussian Mixture Model (GMM) and K- Nearest Neighbor (K-NN) Techniques. *Procedia Computer Science*, pp. 50–57.

- [48] Lee, C.-C., Mower, E., Busso, C., Lee, S., Narayanan, S. (2011). Emotion Recognition Using a Hierarchical Binary Decision Tree Approach. *Speech Communication*, pp. 1162–1171.
- [49] Liu, H., Motoda, H. (1998). *Feature Selection for Knowledge Discovery and Data Mining*. Springer US.
- [50] Liu, J., Chen, C., Bu, J., You, M., Tao, J. (2007). Speech Emotion Recognition using an Enhanced Co-Training Algorithm. *2007 IEEE International Conference on Multimedia and Expo*, pp. 999–1002.
- [51] Lugger, M., Yang, B. (2007). The relevance of voice quality features in speaker independent emotion recognition. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 17–20.
- [52] Lugger, M., Yang, B. W. (2006). Robust estimation of voice quality parameters under real world disturbances. *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1097–1100.
- [53] Lugger, M., Janoir, M.-E., Yang, B. (2009 ). Combining classifiers with diverse feature sets for robust speaker independent emotion recognition. *17th European Signal Processing Conference*, pp. 1225–1229.
- [54] Marill, T., Green, D. M. (1963). On the effectiveness of receptors in recognition systems. *IEEE Transactions on Information Theory*, pp. 11–17.
- [55] Matuzas J., Tišina T., Drabavičius G., Markevičiūtė L. (2015). *Lithuanian Spoken Language Emotions Database*. (Baltic Institute of Advanced Language) Nuskaityta iš <http://datasets.bpti.lt/lithuanian-spoken-language-emotions-database/>.
- [56] Mencattini, A., Martinelli, E., Costantini, G., Todisco, M., Basile, B., Bozzali, M.(2014). Speech emotion recognition using amplitude modulation parameters and a combined feature selection procedure. *Knowledge-Based Systems*, pp. 68–81.
- [57] Oflazoglu, C., Yildirim, S. (2013). Recognizing emotion from Turkish speech using acoustic features. *EURASIP Journal on Audio, Speech, and Music Processing*.
- [58] Origlia, A., Galata, V., Ludusan, B. (2010). Automatic classification of emotions via global and local prosodic features on a multilingual emotional database. *In: Proceedings of Speech Prosody*.



- [59] Park, J.-S., Kim, J.-H., Oh, Y.-H. (2009). Feature vector classification based speech emotion recognition for service robots. *IEEE Transactions on Consumer Electronics*, pp. 1590–1596.
- [60] Peng, H., Long, F., Ding, C. (2005). Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1226–1238.
- [61] Planet, S., Iriondo, I. (2012). Comparative Study on Feature Selection and Fusion Schemes for Emotion Recognition from Speech. *International Journal of Interactive Multimedia and Artificial Intelligence*, 1(6), pp. 44–51.
- [62] Pudil, P., Novovicova, J., Kittler, J. (1994). Floating search methods in feature selection. *Pattern Recognition Letters*, pp. 1119–1125.
- [63] Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning*, pp. 81–106.
- [64] Quinlan, J. R. (1992). *C4.5: Programs for Machine Learning*. Morgan Kaufmann.
- [65] Quinlan, J. R. (1996). Learning decision tree classifiers. *ACM Computing Surveys*, pp. 71–72.
- [66] Ramakrishnan, S., El Emary, I. (2013). Speech emotion recognition approaches in human computer interaction. *Telecommunication Systems*, pp. 1467–1478.
- [67] Rao, K., Koolagudi, S., Vempada, R. (2013). Emotion recognition from speech using global and local prosodic features. *International Journal of Speech Technology*, pp. 143–160.
- [68] Rokach, L., Maimon, O. (2014). *Data Mining with Decision Trees*. World Scientific Publishing Company.
- [69] Rong, J., Li, G., Chen, Y.-P. P. (2009). Acoustic feature selection for automatic emotion recognition from speech. *Information Processing and Management*, pp. 315–328.
- [70] Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information*, pp. 695–729.

- [71] Schuller, B., Reite, S., Rigoll, G. (2006). Evolutionary feature generation in speech emotion recognition. *IEEE International Conference on Multimedia and Expo*, pp. 5–8.
- [72] Schuller, B., Vlasenko, B., Eyben, F., Rigoll, G., Wendemuth, A. (2009). Acoustic Emotion Recognition: A Benchmark Comparison of Performances. *IEEE Workshop on Automatic Speech Recognition & Understanding*, pp. 552–557.
- [73] Siedlecki, W., Sklansky, J. (1989). A note on genetic algorithms for large-scale feature selection. *Pattern Recognition Letters*, pp. 335–347.
- [74] Silla Jr., C., Freitas, A. (2011). A survey of hierarchical classification across different application domains. *Data Mining and Knowledge Discovery*, pp. 31–72.
- [75] Tato, R., Santos, R., Kompe, R., Pardo, J. (2002). Emotional space improves emotion recognition. *7th International Conference on Spoken Language Processing*.
- [76] Theodoridis, S., Koutroumbas, K. (2008). *Pattern Recognition*. Academic Press.
- [77] Ververidis, D., Kotropoulos, C. (2003). A State of the Art Review on Emotional Speech Databases. *Proc. of 1st Richmedia Conf., Laussane*, pp. 1–11.
- [78] Ververidis, D., Kotropoulos, C. (2006). Emotional speech recognition: Resources, features, and methods. *Speech Communication*, pp. 1162–1181.
- [79] Ververidis, D., Kotropoulos, C., Pitas, I. (2004). Automatic emotional speech classification. *Acoustics, Speech, and Signal Processing*.
- [80] Vidrascu, L., Devillers, L., 2007. Five emotion classes detection in real-world call center data: The use of various types of paralinguistic features. *International Workshop on Paralinguistic Speech*, pp. 11–16.
- [81] Vogt, T., Andre, E. (2005). Comparing feature sets for acted and spontaneous speech in view of automatic emotion recognition. *IEEE International Conference on Multimedia and Expo*, pp. 474–477.
- [82] Vogt, T., Andre, E., Wagner, J. (2008). Automatic Recognition of Emotions from Speech: A Review of the Literature and Recommendations for Practical Realisation. *Affect and Emotion in Human-Computer Interaction*, pp. 75–91.

- [83] Whitney, A. (1971). A Direct Method of Nonparametric Measurement Selection. *IEEE Transactions on Computers*, pp. 1100–1103.
- [84] Xiao, Z., Centrale, E., Chen, L., Dou, W. (2009). Recognition of emotions in speech by a hierarchical approach. *3rd International Conference on Affective Computing and Intelligent Interaction and Workshops*, pp. 1–8.
- [85] Xiao, Z., Dellandrea, E., Dou, W., Chen, L. (2007). *Hierarchical Classification of Emotional Speech*. LIRIS UMR 5205 CNRS/INSA de Lyon/Université Claude Bernard.
- [86] Yoon, W.-J., Park, K.-S. (2011). Building robust emotion recognition system on heterogeneous speech databases. *2011 IEEE International Conference on Consumer Electronics*, pp. 825–826.
- [87] You, M., Chen, C., Bu, J., Liu, J., Tao, J. (2006). Emotion recognition from noisy speech. *IEEE International Conference on Multimedia and Expo*, pp. 1653–1656.
- [88] You, M., Chen, C., Bu, J., Liu, J., Tao, J. (2007). Manifolds based emotion recognition in speech. *International Journal of Computational Linguistics and Chinese Language Processing*, pp. 49–64.
- [89] Zhang, S. (2008). Emotion recognition in Chinese natural speech by combining prosody and voice quality features. *5th International Symposium on Neural Networks*, pp. 457–464.
- [90] Zhang, S., Lei, B., Chen, A., Chen, C., Chen, Y. (2010). Spoken emotion recognition using local Fisher discriminant analysis. *IEEE 10th International Conference on Signal Processing*, pp. 538–540.

Tatjana Liogienė

## HIERARCHINIS ŠNEKOS EMOCIJŲ KLASIFIKAVIMAS

Daktaro disertacija

Fiziniai mokslai (P000)

Informatika (09P)

Redaktorė Inesa Didikaitė