

VILNIAUS UNIVERSITETAS

Rasa Lileikytė

ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMAS

Daktaro disertacija
Technologijos mokslai, informatikos inžinerija (07T)

Vilnius, 2012

Disertacija rengta 2007–2012 metais Vilniaus universiteto Matematikos ir informatikos institute.

Mokslinis vadovas

prof. habil. dr. Laimutis Telksnys (Vilniaus universitetas, technologijos mokslai, informatikos inžinerija – 07T).

Reziუმė

Disertacijoje nagrinėjamas šnekos signalų atpažinimo požymių kokybės vertinimas. Pagrindiniai tyrimo objektai yra šnekos signalų atpažinimo požymių kokybės vertinimo metodai ir metrikos. Vertinant atpažinimo sistemų tikslumą tradiciniu būdu, su kiekvienu tiriamos požymių sistemos tipu tenka atlikti atpažinimo tikslumo skaičiavimus. Tokių darbų apimtis galima sumažinti įvertinus pasirenkamų požymių kokybę. Pagrindinis disertacijos tikslas – pateikti naują šnekos signalų atpažinimo požymių kokybės vertinimo metodą, kuris palengvina šnekos signalų atpažinimo uždavinį.

Disertacijoje pateiktas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, grindžiamas metrikų naudojimu. Kokybiška požymių sistema yra klasifikavimo klaidą minimizuojanti požymių sistema. Pateiktame metode klasifikavimo klaidos minimizavimą atitinka metrikų minimizavimas. Minimali metrikos reikšmė nurodo, kad klasės yra gerai atskiriamos, maksimali – klasės blogai atskiriamos. Požymių sistemos kokybės rodiklis apibendrina metrikų rezultatus ir nurodo požymių kokybės lygį. Pateiktas metodas leidžia įvertinti požymių kokybę nevykdant klasifikavimo eksperimentų, tinkamas Euklido erdvės klasifikatoriams.

Pateiktas metrikų rinkinys, tinkantis šnekos signalų atpažinimo požymių kokybei vertinti. Metrikų rinkinys susideda iš trijų metrikų: požymių kiekio klasės ribose, klasių artimiausių kaimynų santykio, klasės ribos peržengimo kiekio.

Sukurta programinė įranga šnekos signalų atpažinimo požymių kokybės vertinimo eksperimentiniams tyrimams atlikti.

Eksperimentinių tyrimų rezultatai patvirtino pateikto šnekos signalų požymių kokybės vertinimo metodo teisingumą: požymių kokybės vertinimo rezultatai naudojant pateiktą metodą sutapo su rezultatais, gautais požymių kokybę vertinant klasifikavimo klaida.

Disertaciją sudaro įvadas, trys skyriai, rezultatų apibendrinimas, naudotos literatūros, autoriaus publikacijų disertacijos tema sąrašai ir priedai.

Disertacijos apimtis 103 puslapiai, neįskaitant priedų, tekste panaudotos 122 numeruotos formulės, 38 paveikslai, 10 lentelių ir 2 priedai. Rašant disertaciją buvo panaudota 118 literatūros šaltinių.

Įvadiniame skyriuje aptariama tiriamoji problema, darbo aktualumas, aprašomas tyrimų objektas, formuluojamas darbo tikslas ir uždaviniai, aprašoma tyrimų metodika, darbo mokslinis naujumas, darbo rezultatų praktinė reikšmė, pateikiami ginamieji teiginiai. Įvado pabaigoje pristatomos disertacijos tema autoriaus paskelbtos publikacijos ir pranešimai konferencijose, seminaruose bei disertacijos struktūra.

Pirmasis skyrius skirtas šnekos signalų atpažinimo požymių kokybės vertinimo analitinei apžvalgai. Atlikta šnekos signalų atpažinimo sistemą sudarančių pagrindinių dalių analizė. Atlikta šnekos signalų atpažinimo požymių kokybės vertinimo metodų ir metrikų analitinė apžvalga. Skyriaus pabaigoje formuluojamos išvados ir disertacijos uždaviniai.

Antrajame skyriuje pateiktas sukurtas šnekos signalų atpažinimo požymių kokybės vertinimo metodas. Metodas yra grindžiamas metrikų naudojimu, pateiktos metrikų charakteristikos. Taip pat pateiktas sukurtos programinės įrangos aprašymas.

Trečiajame skyriuje pateikti eksperimentinių tyrimų rezultatai, kurie patvirtino metodo teisingumą.

Disertacijos tema paskelbti 3 straipsniai leidiniuose, įtrauktuose į ISI Web of Science sąrašą. Disertacijos tema perskaityti 2 pranešimai tarptautinėse konferencijose ir seminaruose.

Abstract

The dissertation attributes to the issue of quality estimation method of speech recognition features. The main objects of the research include quality estimation method of speech recognition features based on metrics. Evaluating the accuracy of speech recognition system in ordinary way, the speech recognition accuracy has to be calculated for each explored feature system. The amount of calculations can be reduced if the quality of explored feature system is estimated. The primary purpose of this dissertation is to propose the new method for quality estimation of speech recognition features, that would make the task of speech recognition easier.

The dissertation proposes the method for quality estimation of speech recognition features that is based on metrics. The quality feature system is defined as feature system that minimizes classification error. In our case, the minimization of classification error is coincided with the minimization of metrics. The minimal metric's value defines that classes are easy separable and the maximal metric's value defines that classes are hardly separable. Therefore, the quality index of feature system summarizes the results of metrics and identifies the level of feature system quality. Accordingly, the proposed method enables to evaluate the quality of feature system without performing classification experiments and it is suitable for Euclidean space classifiers.

The set of metrics is proposed for quality estimation of speech recognition features: feature volume of class boundary, nearest neighbour distances ratio of classes, overstep volume of class boundary.

The software was implemented for experimental researches of the proposed method for quality estimation of speech recognition features.

The results of experimental researches confirmed the correctness of proposed quality estimation method of speech recognition features: the results of the proposed method coincided with the results of classification error.

The dissertation consists of three parts including introduction, 3 chapters, conclusions, references and annexes. The total scope of the dissertation – 103

pages without annexes, 122 formulas, 38 pictures, 10 tables and 2 annexes. Also it was used 118 references.

The introduction reveals the investigated problem, importance of the thesis and the object of research and describes the purpose and tasks of the dissertation, research methodology, scientific novelty, the practical significance of results examined in the dissertation and defended statements. The introduction ends in presenting the author's publications on the subject of the defended dissertation, offering the material of performed presentations in conferences, seminars and defining the structure of the dissertation.

Chapter 1 revises the estimation of speech recognition features. The main parts of the speech recognition system are described. The existing methods and metrics of speech recognition features are analyzed. At the end of the chapter, conclusions are drawn and the tasks for the dissertation are reconsidered.

Chapter 2 proposes the method for quality estimation of speech recognition features. The method is based on metrics, also the characteristics of metrics are defined. The description of created software is provided.

Chapter 3 provides the results of experimental researches. The experimental results confirmed the correctness of the proposed method.

3 articles focusing on the subject of the discussed dissertation are published in the ISI Web of Science register. The presentation on the subject has been performed in 2 conferences and seminars at international level.

Žymėjimai

Simboliai

x_n^i – i -osios klasės n -asis požymių vektorius;

x_{nd}^i – i -osios klasės n -ojo požymių vektoriaus d -asis komponentas;

N_i – i -osios klasės vektorių skaičius;

D – vektoriaus komponentų skaičius;

C – klasių skaičius.

Santrumpos

AK klasifikatorius – vieno artimiausio kaimyno klasifikatorius (angl. *Nearest neighbour classifier*);

DLSK klasifikatorius – dinaminio laiko skalės kraipymo klasifikatorius (angl. *Dynamic Time Warping classifier*);

G1 – požymių kiekio klasės ribose metrika;

G2 – klasių artimiausių kaimynų atstumų santykio metrika;

G3 – klasės ribos peržengimo kiekio metrika;

K -AK klasifikatorius – K -artimiausių kaimynų klasifikatorius (angl. *K-Nearest neighbour classifier*);

KK – kepstro koeficientai (angl. *Real-cepstrum coefficients*);

MJM – minimalus jungiantysis medis (angl. *Minimal spanning tree*);

TSPKK – tiesinės suvokimo prognozės kepstro koeficientai (angl. *Perceptual linear prediction cepstrum coefficients*).

Sąvokos

Euklido erdvės klasifikatorius – klasifikatorius, kuris naudoja Euklido atstumo matą;

Kokybė – matas, kuris nusako tiriamojo objekto charakteristikų atitikimo lygį apibrėžtiems reikalavimams;

Kokybiška požymių sistema – klasifikavimo klaidą minimizuojanti požymių sistema. Pateiktame metode yra naudojamos metrikos, o klasifikavimo klaidos minimizavimą atitinka metrikų minimizavimas. Požymių sistemos kokybės rodiklis apibendrina metrikų rezultatus;

Metrika – klasių atskiriamumo lygį nurodantis matas. Metrikos reikšmės išsidėsčiusios tam tikrame intervale. Metrikai įgijus minimalią reikšmę yra teigiama, kad klasės yra gerai atskiriamos. Metrikai įgijus maksimalią reikšmę yra teigiama, kad klasės blogai atskiriamos;

Požymių sistemos kokybės rodiklis – metrikų rezultatus apibendrinantis matas. Rodiklio reikšmė „0 %“ nurodo žemiausią požymių sistemos kokybę, „100 %“ nurodo aukščiausią požymių sistemos kokybę.

Turinys

ĮVADAS	1
Tiriamoji problema	1
Darbo aktualumas	1
Tyrimų objektas	1
Darbo tikslas	2
Darbo uždaviniai	2
Tyrimų metodika	3
Darbo mokslinis naujumas ir jo reikšmė	3
Darbo rezultatų praktinė reikšmė	4
Ginamieji teiginiai	4
Darbo rezultatų aprobavimas	4
Disertacijos struktūra	5
1. ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMO	
ANALIZĖ.....	7
1.1. Šnekos atpažinimo sistemų analizė.....	7
1.1.1. Šnekos atpažinimo sistemų struktūra	7
1.1.2. Signalų apdorojimas	8
1.1.3. Šnekos atpažinimo požymių sistemos.....	9
1.1.4. Atstumo matas	19

1.1.5. Šnekos atpažinimo metodai	20
1.2. Šnekos atpažinimo požymių kokybės vertinimo analitinė apžvalga	26
1.2.1. Klasifikavimo klaidos apskaičiavimas	27
1.2.2. Geometrinės metrikos	31
1.2.3. Informacijos teorijos metrikos	37
1.2.4. Statistinės metrikos	38
1.3. Pirmojo skyriaus išvados ir disertacijos uždavinių formulavimas	43
2. ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMO METODAS.....	45
2.1. Metodas šnekos atpažinimo požymių kokybei vertinti.....	45
2.2. Metrikos šnekos atpažinimo požymių kokybei vertinti	50
2.2.1. Požymių kiekio klasės ribose metrika.....	51
2.2.2. Klasių artimiausių kaimynų atstumų santykio metrika.....	54
2.2.3. Klasės ribos peržengimo kiekio metrika	56
2.3. Programinė įranga požymių kokybei vertinti.....	57
2.4. Skyriaus išvados	60
3. ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMO EKSPERIMENTINIAI TYRIMAI.....	63
3.1. Eksperimentinių tyrimų duomenys ir sąlygos.....	63
3.2. Požymių kokybės vertinimo eksperimentiniai tyrimai	66
3.3. Skyriaus rezultatai ir išvados.....	82
BENDROSIOS IŠVADOS.....	85
LITERATŪRA IR ŠALTINIAI	89
AUTORIAUS PUBLIKACIJOS DISERTACIJOS TEMA.....	103
PRIEDAI	104
A priedas. Eksperimentinių tyrimų fonemų kombinacijos	104
B priedas. Metrikų vidurkių įverčiai.....	106

Ivadas

Tiriamoji problema

Darbe nagrinėjama šnekos signalų atpažinimo požymių kokybės vertinimo problema, analizuojami požymių kokybės vertinimo metodai ir metrikos.

Darbo aktualumas

Automatinis šnekos atpažinimas yra taikomas įvairiose žmonių veiklose: klientų aptarnavimas didelėse kompanijose – informacijos teikimas, skambučių priėmimų centrų automatizavimas sumažinant laukimo laiką; automatizuoto teksto rinkimas diktuojant – teksto surinkimas nereikalauja surinkimo įgūdžių ir sunaudojama mažiau žmonių darbo resursų; pasinaudodami automatiniu kalbos vertimu, žmonės gali susikalbėti ir nemokėdami reikiamos kalbos.

Šnekos signalų atpažinimo sistemų sudarymas yra sudėtingas uždavinys. Šnekos signalų atpažinimo sistemų tikslumas priklauso nuo šnekos signalus aprašančių požymių ir šiuos požymius naudojančių klasifikatorių savybių. Tradiciškai vertinant atpažinimo sistemų tikslumą, kiekvienai pasirinktai požymių sistemai ir kiekvienam klasifikatoriaus tipui tenka atlikti atpažinimo tikslumo skaičiavimus. Tačiau šis metodas yra neefektyvus, nes atpažinimo sistemos klaidos apskaičiavimas reikalauja daug darbo ir didelių skaičiavimo

resursų. Tokių darbų apimtis galima sumažinti – atpažinimo sistemos konstravimo pradiniam etape atlikus pasirenkamų požymių kokybės vertinimą. Taigi, reikalingas naujas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, kuris palengvintų šnekos signalų atpažinimo uždavinį.

Tyrimų objektas

Tyrimų objektas – šnekos signalų atpažinimo požymių kokybės vertinimo metodai ir metrikos.

Darbo tikslas

Darbo tikslas – pateikti šnekos signalų atpažinimo požymių kokybės vertinimo metodą, suteikiantį galimybę palengvinti šnekos signalų atpažinimo uždavinį.

Darbo uždaviniai

Darbo tikslui pasiekti reikia išspręsti šiuos uždavinius:

1. Pateikti naują šnekos signalų atpažinimo požymių kokybės vertinimo metodą, grindžiamą metrikų naudojimu.
2. Pateikti metrikų rinkinį šnekos signalų atpažinimo požymių kokybei vertinti.
3. Parodyti, kad šnekos signalų požymių kokybę aprašantis metodas Euklido erdvėje suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį, atsisakant klasifikatoriaus įtakos skaičiavimų.
4. Įvertinti šnekos signalų požymių kokybės ir atpažinimo sistemų kokybės vertinimo metodų algoritmų sudėtingumą.
5. Sukurti šnekos signalų atpažinimo požymių kokybės vertinimo tyrimams pritaikytą eksperimentinę bazę.

6. Patvirtinti eksperimentiškai pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą.

Tyrimų metodika

Teorinei analizei, praktinei realizacijai ir tyrimams panaudotos matematinės statistikos, skaitmeninio signalų apdorojimo bei atpažinimo teorijos žinios. Programinė įranga sukurta naudojant PL/SQL kalbą (Oracle SQL Developer 2.1), C++ kalbą (MinGW 0.2) ir Matlab R2007b integruotą programų kūrimo aplinką.

Darbo mokslinis naujumas ir jo reikšmė

Rengiant disertaciją buvo gauti šie mokslui nauji rezultatai:

- Pateiktas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, grindžiamas metrikų naudojimu.
- Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti.
- Pasiūlytas šnekos signalų požymių kokybę aprašantis metodas, suteikiantis galimybę supaprastinti šnekos signalų atpažinimo uždavinį, atsisakant klasifikatoriaus įtakos skaičiavimų.
- Parodyta, kad šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, o dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, R – šnekos signalo objekto vektorių skaičius.
- Sukurta eksperimentinė bazė, leidžianti atlikti eksperimentinius tyrimus šnekos signalų požymių ir šnekos signalų atpažinimo sistemų kokybei vertinti.

Darbo rezultatų praktinė reikšmė

Pasiūlytas šnekos signalų atpažinimo požymių kokybės vertinimo metodas suteikia galimybę kokybiškiau atlikti ir paspartinti šnekos signalų atpažinimo požymių kokybės vertinimo darbus, o taip pat supaprastinti, taupiau spręsti šnekos signalų atpažinimo sistemų kokybės vertinimo uždavinius, atsisakant klasifikatoriaus įtakos skaičiavimų.

Ginamieji teiginiai

- Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti, susidedantis iš trijų metrikų – požymių kiekio klasės ribose, klasių artimiausių kaimynų atstumų santykio, klasės ribos peržengimo kiekio – suteikė galimybę sukurti naudingą šnekos signalų atpažinimo požymių kokybės vertinimo metodą.
- Pateiktas naujas šnekos signalų atpažinimo požymių kokybės vertinimo metodas atspindi šnekos atpažinimo sistemų, veikiančių Euklido erdvėje, savybes.
- Įvesta požymių kokybės rodiklių skalė 0 %–100 %, kur 0 % rodo žemiausią požymių sistemos kokybę, o 100 % – aukščiausią požymių sistemos kokybę, leidžia matuoti šnekos signalų atpažinimo požymių vertingumą.
- Pasiūlytas šnekos signalų požymių kokybę aprašantis metodas suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį, atsisakant klasifikatoriaus įtakos skaičiavimų.
- Šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, o dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, R – šnekos signalo objekto vektorių skaičius.
- Sukurta eksperimentinė bazė padėjo eksperimentiškai parodyti, kad naujasis šnekos signalų atpažinimo požymių kokybės vertinimo metodas

Euklido erdvėje leidžia mažesnėmis sąnaudomis aprašyti atpažinimo sistemų kokybę.

Darbo rezultatų apibavimas

Disertacijos tema atspausdinti 3 moksliniai straipsniai žurnaluose, įtrauktuose į ISI Web of Science sąrašą:

1. Lileikytė, R.; Telksnys, L. 2011. Quality Estimation Methodology of Speech Recognition Features, *Electronics and Electrical Engineering*. 4(110), 113–116. ISSN 1392–1215.
2. Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering*. 2(118). ISSN 1392–1215.
3. Lileikytė, R.; Telksnys, L. Quality estimation of speech recognition features for Dynamic Time Warping classifier, *Information Technology and Control*. ISSN 1392–124X. (Priimta spausdinimui).

Rezultatai paskelbti 2 mokslinėse konferencijose ir seminaruose:

1. *Quality Estimation Methodology of Speech Recognition Features* pristatytas tarptautinėje konferencijoje „Elektronika“ 2011 m., Kaune.
2. *Šnekos atpažinimo požymių kokybės vertinimas* pristatytas tarptautiniame seminare „Institute of Electrical and Electronic Engineers“ 2011 m., Vilniuje.

Disertacijos struktūra

Disertaciją sudaro įvadas, trys skyriai, rezultatų apibendrinimas ir du priedai.

Darbo apimtis yra 103 puslapiai, neįskaitant priedų, tekste panaudotos 122 numeruotos formulės, 38 paveikslai ir 10 lentelių. Rašant disertaciją buvo panaudota 118 literatūros šaltinių.

Šnekos atpažinimo požymių kokybės vertinimo analizė

Skyriuje pateikiama šnekos atpažinimo sistemų analizė, apžvelgiama atpažinimo sistemų struktūra, signalo apdorojimas, šnekos atpažinimo požymių sistemos ir atpažinimo metodai. Pateikiama šnekos atpažinimo požymių kokybės vertinimo analizė, apžvelgiami požymių kokybės vertinimo metodai, nagrinėtos metrikos suskirstytos į tris grupes: geometrines, informacijos teorijos ir statistines.

1.1. Šnekos atpažinimo sistemų analizė

1.1.1. Šnekos atpažinimo sistemų struktūra

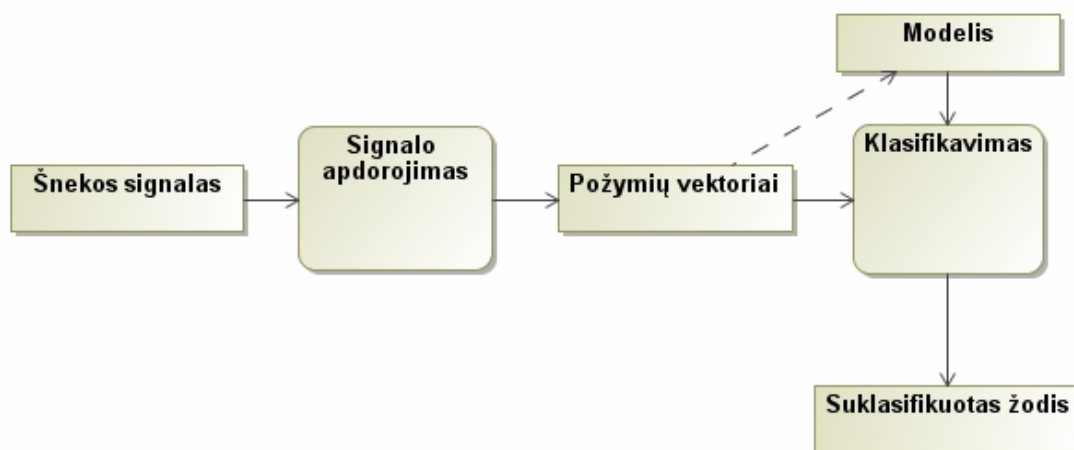
Automatinė šnekos atpažinimo technologija suteikė galimybę kompiuteriui suprasti žmogaus šneką ir vykdyti išstartas komandas (Juang *et al.* 2005; Anusuya *et al.* 2009). Šnekos atpažinimas yra tiriamas apie šešiasdešimt metų, siekiant mechaniškai realizuoti žmogaus šnekos gebėjimą, automatizuoti uždavinius, kuriuos žmogus gali atlikti sąveikaudamas su kompiuteriu. Automatinės šnekos atpažinimo sistemos šiandien yra naudojamos skambučių priėmimų centrų automatizavimui sumažinant laukimo laiką, užklausomis grįstoms sistemoms, kurios teikia atnaujintą informaciją apie keliones, akcijų

kainų pokyčius, orus. Taip pat naudojama tokiems uždaviniams kaip duomenų įvedimas, diktavimas balsu.

Automatinių šnekos atpažinimo sistemų veikimas susideda iš dviejų etapų (1.1 pav.) (Junqua *et al.* 1996): apmokymo ir atpažinimo.

Apmokymo metu formuojami šnekos vienetų modeliai (frazių, žodžių, fonemų), kurie sudaro atpažinimo žodyną. Priklausomai nuo naudojamo atpažinimo metodo, šie modeliai gali būti grįsti etaloniniais modeliais arba statistiniais parametriniais modeliais.

Atpažinimo metu skaičiuojamas panašumo matas tarp nežinomo šnekos signalo ir turimų modelių. Nežinomas signalas yra priskiriamas klasei, kurios modelius atitinka geriausiai.



1.1 pav. Automatinės šnekos atpažinimo sistemos struktūra

1.1.2. Signalų apdorojimas

Šnekos signalo apdorojimas dažniausiai susideda iš pradinės filtracijos, dalijimo į kadrus, lango funkcijos taikymo, požymių išskyrimo etapų (Junqua *et al.* 1996; Rabiner *et al.* 1993).

Šnekos signalo spektro dažniams didėjant, būdingas signalo intensyvumo silpnėjimas 6 dB/oktavai. Todėl naudojant pradinio apdorojimo filtrą, signalo spektras yra „išlyginamas“. Tam tikslui naudojamas pirmos eilės skaitmeninis filtras:

$$\tilde{x}(n) = x(n) - \alpha x(n-1), \quad (1.1)$$

čia $\tilde{x}(n)$ yra filtruotas šnekos signalas, $x(n)$ – pradinis diskretizuotas šnekos signalas, α – koeficientas, nurodantis spektro išlyginimo laipsnį, parenkamas iš $0,9 \leq \alpha \leq 1,0$ intervalo.

Šnekos signalas suskaidomas į trumpas atkarpas – kadrus, kurių trukmė yra 20–30 ms, daroma prielaida, kad šnekos savybės esančios trumpoje atkarpoje nekinta laike (yra stacionarios). Kadrai persikloja taip, kad gretimas kadras prasideda nuo tam tikros prieš jį einančios kadro dalies.

Siekiant nuglodinti kadro kraštus, kiekvienam kadrai yra taikoma lango funkcija. Naudojamos įvairios lango funkcijos, tokios kaip Hamingo, Haningo:

Hamingo lango funkcija:

$$w_{HM}(n) = 0,54 - 0,46 \cos\left(\frac{2n\pi}{N-1}\right), \quad 1 \leq n \leq N, \quad (1.2)$$

čia N yra signalo kadro ilgis atskaitomis.

Haningo lango funkcija:

$$w_{HN}(n) = 0,5 - 5 \cos\left(\frac{2n\pi}{N-1}\right), \quad 1 \leq n \leq N. \quad (1.3)$$

Kitame skyriuje apžvelgiamos šnekos atpažinimo požymių sistemos, suskirstytos į keturias grupes: laiko analizės, kepstro analizės, tiesinės prognozės analizės, tiesinės suvokimo prognozės analizės.

1.1.3. Šnekos atpažinimo požymių sistemos

1.1.3.1. Laiko analizės požymiai

Paprasčiausios signalo charakteristikos yra priskiriamos laiko analizei.

Energija. Signalo energija yra geras matas balsėms nuo priebalsių atskirti. Priebalsių energija yra daug mažesnė nei balsių. Taip pat, kokybiško signalo atveju, energija naudojama atskirti priebalsius nuo tylos. Trumpalaikio signalo energija išreiškiama formule (Schafer *et al.* 1975):

$$E_n = \sum_{m=0}^{N-1} [x(m)w(n-m)]^2. \quad (1.4)$$

čia $x(m)$ yra diskretinis signalas, $w(\cdot)$ – lango funkcija laiko momentu n , N yra signalo kadro ilgis atskaitomis.

Nulio kirtimų dažnis. Nulio kirtimų dažnis parodo, kiek kartų signalo amplitudė pakeičia ženklą per laiko vieneta. Nulio kirtimų dažnis naudojamas balsiams nuo priebalsių atskirti. Didelė šio mato reikšmė gaunama priebalsių atveju, maža reikšmė – balsių atveju. Trumpalaikio signalo nulio kirtimų dažnis yra apskaičiuojamas:

$$Z_n = \sum_{m=0}^{N-1} |\operatorname{sgn}[x(m)] - \operatorname{sgn}[x(m-1)]| w(n-m), \quad (1.5)$$

$$\operatorname{sgn}[x(n)] = \begin{cases} 1, & x(n) \geq 0, \\ -1, & x(n) < 0. \end{cases} \quad (1.6)$$

Autokoreliacija. Autokoreliacijos funkcija parodo nagrinėjamo signalo periodiškumą. Tarkime, turime trumpalaikį signalą, kurio kadro ilgis yra N :

$$x_l(n) = x(n+1), \quad 0 \leq n \leq N-1, \quad (1.7)$$

čia l žymi kadro pradžią. Trumpalaikė autokoreliacija išreiškiama:

$$r_l(m) = \frac{1}{N} \sum_{n=0}^{N-m-1} x_l(n)x_l(n+m), \quad 0 \leq m \leq M-1, \quad (1.8)$$

čia m yra vėlinimo parametras.

Laiko požymiai nėra efektyvūs, nes nėra atsparūs triukšmui. Automatinėse šnekos atpažinimo sistemose dažnai naudojami tokie požymiai, kaip melų skalės kepstro koeficientai, tiesinės prognozės kepstro koeficientai, tiesinės suvokimo prognozės kepstro koeficientai (Pinto 2010). Šie požymiai aprašomi kituose skyriuose.

1.1.3.2. Kepstro analizės požymiai

Kepstro koeficientai. Signalo kepstros yra apibrėžiamas kaip signalo Furjė transformacijos logaritmo atvirkštinė transformacija (Childers *et al.* 1977; Furu 2001; Bohme 1974). Kepstras dar vadinamas realiuoju kepstru, kai

naudojama realioji kepstro dalis. Kepstro koeficientai (KK) trumpalaikiui šnekos signalui yra apskaičiuojami pagal (1.9)–(1.10) išraiškas:

$$X(k) = \sum_{n=0}^{N-1} x(n)w(m-n)e^{-j2\pi kn/N}, \quad 0 \leq k \leq N-1, \quad (1.9)$$

$$c_n = \frac{1}{N} \sum_{k=0}^{N-1} \log|X(k)|e^{j2\pi kn/N}, \quad 0 \leq n \leq N-1, \quad (1.10)$$

čia c_n yra n -asis kepstro koeficientas, $X(k)$ yra signalo $x(n)$ spektras atlikus diskrečiąją Furjė transformaciją, $w(\cdot)$ – lango funkcija, k – dažnio indeksas, N yra signalo kadro ilgis atskaitomis. Dažnai naudojama Greitoji Furjė transformacija (Cooley *et al.* 1967; Oppenheim 1970; Heideman *et al.* 1984).

Melų skalės kepstro koeficientai. Melų skalė grindžiama žmogaus klausos sistemos dažninėmis savybėmis (Davis *et al.* 1980; Han *et al.* 2006; Hossan *et al.* 2010; Goyani *et al.* 2010). Melų skalės dažnio ir tiesinės skalės dažnio sąryšis išreiškiamas:

$$mel(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right), \quad (1.11)$$

čia f yra tiesinės dažnių skalės dažnis.

Melų dažnių filtrų blokas dažniausiai yra sudarytas iš persidengiančių trikampių filtrų. Filtrai žemų dažnių srityje išsidėstę tiesiškai, aukštų dažnių srityje išsidėstę logaritmiškai. Apskaičiuojant melų skalės kepstro koeficientus, pirmiausia šnekos signalo kadrai pritaikoma diskrečioji Furjė transformacija (DFT) (1.2 pav.):

$$X(k) = \sum_{n=0}^{N-1} x(n)w(m-n)e^{-j2\pi kn/N}, \quad 0 \leq k \leq N-1. \quad (1.12)$$

Po to atliekamas filtravimas ir logaritmuojama:

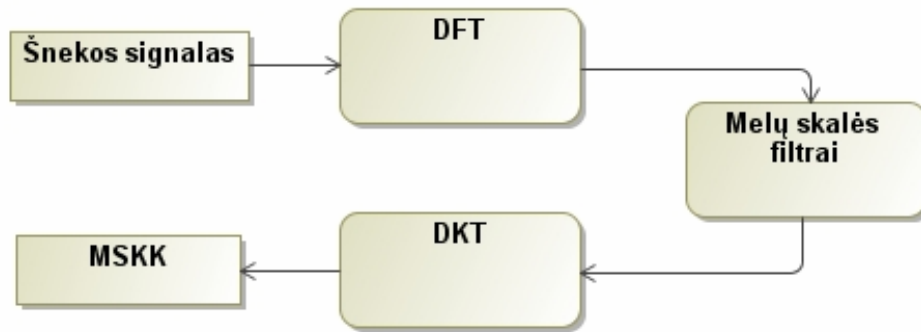
$$F(l) = \log \left(\sum_{k=0}^{N-1} |X(k)|^2 H_l(k) \right), \quad 1 \leq l \leq L, \quad (1.13)$$

čia $H_l(k)$ yra l -ojo trikampio filtro dažninė charakteristika.

Kitame etape yra atliekama diskrečioji kosinuso transformacija (DKT) ir gaunami melų skalės kepstro koeficientai (MSKK):

$$cm_n = \sum_{l=0}^{L-1} F(l) \cos\left(\frac{\pi n}{L} \left(l - \frac{1}{2}\right)\right), \quad 1 \leq n \leq L, \quad (1.14)$$

čia cm_n yra n -asis melų skalės kepstro koeficientas.



1.2 pav. Melų skalės kepstro koeficientų radimo schema

Delta, delta-delta koeficientai. Naudojant kepstro koeficientus, pavyzdžiui, KK, MSKK, apskaičiuojami delta ir delta-delta koeficientai. Delta koeficientai apskaičiuojami kaip skirtumas tarp dviejų kadų kepstro vektorių:

$$\Delta x_i = \frac{1}{2}(x_{i+1} - x_{i-1}), \quad (1.15)$$

čia x_{i+1} , x_{i-1} yra atitinkamai $i+1$ -ojo, $i-1$ -ojo kadų kepstrai.

Delta-delta koeficientai apskaičiuojami kaip skirtumas tarp dviejų delta koeficientų vektorių:

$$\Delta \Delta x_i = \frac{1}{2}(\Delta x_{i+1} - \Delta x_{i-1}). \quad (1.16)$$

1.1.3.3. Tiesinės prognozės modelio analizės požymiai

Tiesinės prognozės modelio koeficientai. Tiesinės prognozės modelyje (TPM) turint signalą laiko momentu n , signalas $x(n)$ gali būti aproksimuotas kaip p ankstesnių signalo reikšmių tiesinė kombinacija (Rabiner *et al.* 1993; Makhoul 1975; Tierney 1980; Van der Merwe *et al.* 1991):

$$x(n) \approx a_1 x(n-1) + a_2 x(n-2) + \dots + a_p x(n-p), \quad (1.17)$$

čia a_1, a_2, \dots, a_p yra prognozės koeficientai.

Įvedus žadinimo signalą, išraiška (1.17) išreiškiama:

$$x(n) = \sum_{k=1}^p a_k x(n-k) + Gu(n), \quad (1.18)$$

čia $u(n)$ yra žadinimo signalas, G – stiprinimo koeficientas.

Išraiška (1.18) išreiškiama z plokštumoje ir naudojamas visų polių filtras:

$$H(z) = \frac{G}{1 - \sum_{k=1}^p a_k z^{-k}} = \frac{G}{A(z)}. \quad (1.19)$$

Prognozės klaida išreiškiama:

$$e(n) = x(n) - \tilde{x}(n) = x(n) - \sum_{k=1}^p a_k x(n-k). \quad (1.20)$$

Tam, kad sudarytume lygtį, kurią išsprendus būtų galima rasti prognozės koeficientus, apibrėžiamas trumpalaikis signalas ir klaida n -uoju laiko momentu:

$$x_n(m) = x(n+m), \quad (1.21)$$

$$e_n(m) = e(n+m). \quad (1.22)$$

Siekama minimizuoti vidutinę kvadratinę klaidą laiko momentu n :

$$E_n = \sum_m \left[x_n(m) - \sum_{k=1}^p a_k x_n(m-k) \right]^2 \quad (1.23)$$

E_n yra diferencijuojamas pagal a_k ir prilyginamas nuliui:

$$\frac{\partial E_n}{\partial a_k} = 0, \quad 1 \leq k \leq p. \quad (1.24)$$

Gaunama:

$$\sum_m x_n(m-i)x_n(m) = \sum_{k=1}^p \tilde{a}_k \sum_m x_n(m-i)x_n(m-k). \quad (1.25)$$

Pažymėkime:

$$\phi_n(i, k) = \sum_m x_n(m-i)x_n(m-k). \quad (1.26)$$

Tuomet (1.26) užrašoma taip:

$$\phi_n(i, 0) = \sum_{k=1}^p \tilde{\alpha}_k \phi_n(i, k). \quad (1.27)$$

Minimali vidutinė kvadratinė klaida \tilde{E}_n gali būti išreiškama

$$\tilde{E}_n = \sum_m x_n^2(m) - \sum_{k=1}^p \tilde{\alpha}_k \sum_m x_n(m)x_n(m-k), \quad (1.28)$$

$$\tilde{E}_n = \phi_n(0, 0) - \sum_{k=1}^p \tilde{\alpha}_k \phi_n(0, k). \quad (1.29)$$

Tam, kad surastume prognozės koeficientus, reikia apskaičiuoti $\phi_n(i, k)$, $1 \leq i \leq p$, $1 \leq k \leq p$ išsprendžiant gautų p lygybių.

Vienas iš metodų prognozės koeficientams rasti yra *autokoreliacijos metodas*. Tarkime, signalas $x_n(m)$ už intervalo $0 \leq m \leq N-1$ ribų lygus nuliui. Tai prilygsta signalo $x(m+n)$ padauginimui iš lango funkcijos $w(m)$. Šnekos signalas išreiškiamas:

$$x_n(m) = \begin{cases} x(m+n) \cdot w(m), & 0 \leq m \leq N-1, \\ 0, & \text{kitu atveju.} \end{cases} \quad (1.30)$$

Tada minimali kvadratinė suma lygi:

$$E_n = \sum_{m=0}^{N-1+p} e_n^2(m), \quad (1.31)$$

ir $\phi_n(i, k)$ gali būti aprašoma:

$$\phi_n(i, k) = \sum_{m=0}^{N-1-(i-k)} x_n(m)x_n(m+i-k), \quad 1 \leq i \leq p, \quad 0 \leq k \leq p. \quad (1.32)$$

Kovariacinė funkcija $\phi_n(i, k)$ išreiškama autokoreliacijos funkcija:

$$\phi_n(i, k) = r_n(i - k) = \sum_{m=0}^{N-1-(i-k)} x_n(m)x_n(m+i-k). \quad (1.33)$$

Kadangi autokoreliacijos funkcija yra simetrinė, tiesinės prognozės modelio lygtys gali būti apskaičiuojamos:

$$\sum_{k=1}^p r_n(|i-k|)\tilde{a}_k = r_n(i), \quad 1 \leq i \leq p, \quad (1.34)$$

ir užrašoma matricos forma:

$$\begin{bmatrix} r_n(0) & r_n(1) & \dots & r_n(p-1) \\ r_n(1) & r_n(0) & \dots & r_n(p-2) \\ r_n(2) & r_n(1) & \dots & r_n(p-3) \\ \dots & \dots & \dots & \dots \\ r_n(p-1) & r_n(p-2) & \dots & r_n(0) \end{bmatrix} \begin{bmatrix} \tilde{a}_1 \\ \tilde{a}_2 \\ \tilde{a}_3 \\ \dots \\ \tilde{a}_p \end{bmatrix} = \begin{bmatrix} r_n(1) \\ r_n(2) \\ r_n(3) \\ \dots \\ r_n(p) \end{bmatrix}. \quad (1.35)$$

Ši autokoreliacijos reikšmių matrica yra Tioplico matrica (angl. *Toeplitz*), vienas iš jos sprendimo būdų – *Durbino algoritmas*:

$$E^{(0)} = r(0), \quad (1.36)$$

$$k_i = \left\{ r(i) - \sum_{j=1}^{L-1} \alpha_j^{(i-1)} r(|i-j|) \right\} / E^{(i-1)}, \quad 1 \leq i \leq p, \quad (1.37)$$

$$\alpha_i^{(i)} = k_i, \quad (1.38)$$

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)}, \quad (1.39)$$

$$E^{(i)} = (1 - k_i^2) E^{(i-1)}, \quad (1.40)$$

čia (1.37) patogumo dėlei $r_l(m)$ žymėjimas l yra praleidžiamas. Lygtys (1.36)–(1.40) yra išsprendžiamos rekursyviai $1 \leq i \leq p$ ir gaunama:

$$a_m = \text{TPM koeficientai} = \alpha_m^{(p)}, \quad 1 \leq m \leq p, \quad (1.41)$$

$$k_m = \text{dalinės koreliacijos koeficientai}, \quad (1.42)$$

$$g_m = \text{logaritminės srities santykio koeficientai}. \quad (1.43)$$

Dalinės koreliacijos koeficientai (angl. *PARCOR*) k_m gali būti apskaičiuojami tiesiogiai iš prognozės koeficientų, naudojant rekursiją:

$$k_m = a_m^{(m)}, \quad (1.44)$$

$$a_j^{(m-1)} = \frac{a_j^{(m)} - a_m^{(m)} a_{m-j}^{(m)}}{1 - k_m^2}, \quad 1 \leq j \leq m-1, \quad (1.45)$$

čia m atitinka reikšmes $p, p-1, \dots, 1$, pradinė sąlyga $a_j^{(p)} = a_j$, $1 \leq j \leq p$.

Dalinės koreliacijos koeficientų reikšmių intervalas $-1 \leq k_m \leq 1$.

Logaritminės srities santykio koeficientai (angl. *log area ratio*) išvedami iš dalinės koreliacijos koeficientų:

$$g_m = \log \left[\frac{1 - k_m}{1 + k_m} \right]. \quad (1.46)$$

Autokoreliacijos koeficientai gali būti išvedami iš a_k – prognozės koeficientų:

$$R_a(m) = \sum_{k=0}^{p-m} a_k a_{k+1}. \quad (1.47)$$

Tiesinės spektrinės poros koeficientai (angl. *line spectrum pair*). Nustatomi atitikmenys P nulių iš $A(z)$ ant vienetinio apskritimo naudojant dvi z transformacijas $P(z)$ ir $Q(z)$:

$$P(z) = A(z) + z^{-(p+1)} A(z^{-1}), \quad (1.48)$$

$$Q(z) = A(z) - z^{-(p+1)} A(z^{-1}). \quad (1.49)$$

Tiesinės spektrinės poros koeficientai atitinka nulių dažnius, esančius ant vienetinio apskritimo.

Tiesinės prognozės modelio keistro koeficientai. Tiesinės prognozės modelio keistro koeficientai išvedami iš tiesinės prognozės modelio koeficientų ir yra apskaičiuojami:

$$c_0 = \ln \sigma^2, \quad (1.50)$$

$$c_m = a_m + \sum_{k=1}^{m-1} \binom{m-1}{k} c_k a_{m-k}, \quad 1 \leq m \leq p, \quad (1.51)$$

$$c_m = \sum_{k=1}^{m-1} \binom{m-1}{k} c_k a_{m-k}, \quad m > p, \quad (1.52)$$

čia σ^2 yra stiprinimo koeficientas.

1.1.3.4. Tiesinės suvokimo prognozės modelio analizės požymiai

Tiesinė suvokimo prognozė (angl. *Perceptual Linear Prediction*) modeliuoja klausos suvokimu grįstą spektrą. Tam naudojama visų polių funkcija bei tiesinės prognozės autokoreliacijos technika (Hermansky *et al.* 1985; Hermansky *et al.* 1990; Gunawan *et al.* 2001; Yuan *et al.* 2010). Tiesinis suvokimo prognozės modelis skiriasi nuo tiesinio prognozės modelio, nes šiuo atveju modelis naudoja: 1) barkų skalės dažnių filtrus; 2) garsumo išlyginimą (angl. *equal-loudness preemphasis*); 3) kubinės šaknies garsumo intensyvumo kompresiją (angl. *cubic-root intensity loudness compression*). Šie papildomi apdorojimo etapai modeliuoja žmogaus klausos sistemos savybes.

Barkų skalės dažnio ir dažnio hercais sąryšis yra išreiškiamas:

$$bark(f) = 6 \log \left[\frac{f}{600} + \sqrt{\left(\frac{f}{600} \right)^2 + 1} \right], \quad (1.53)$$

čia f yra dažnis hercais.

Pirmiausia trumpalaikiam šnekos signalui pritaikoma diskrečioji Furjė transformacija (DFT), gautas spektras praleidžiamas pro barkų skalės dažnių filtrus, naudojant septyniolika filtrų:

$$\tilde{X}(\omega) = \sum_{k=0}^{N-1} |X(k)|^2 H_{\omega}(k), \quad (1.54)$$

čia $X(k)$ žymi šnekos signalą pritaikius Furjė transformaciją, $H_{\omega}(k)$ yra barkų skalės filtrų funkcija, ω – kampinis dažnis. Gautam rezultatui taikoma

garso išlyginimo funkcija, kuri reprezentuoja žmogaus klausos jautrumą skirtingiems dažniams. Išlyginimo funkcija išreiškiama (Junqua *et al.* 1996):

$$E(\omega) = 1,151 \sqrt{\frac{(\omega^2 + 144 \cdot 10^4) \cdot \omega^2}{(\omega^2 + 16 \cdot 10^4)(\omega^2 + 961 \cdot 10^4)}}. \quad (1.55)$$

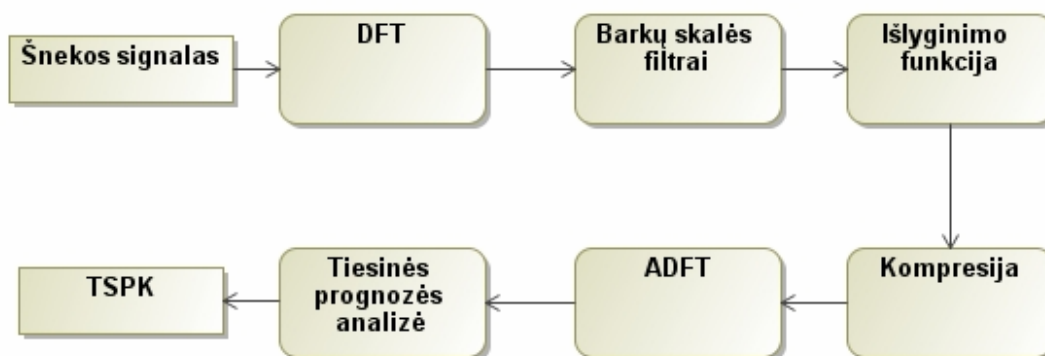
Gautajam signalui pritaikoma išlyginimo funkcija:

$$F(\omega) = E(\omega) \tilde{X}(\omega). \quad (1.56)$$

Kitame etape žmogaus klausos suvokimas realizuojamas nustatant intensyvumo suvokimo garsumą. Tono suvokimo garsumas yra apytiksliai proporcingas tono intensyvumo kubinei šakniai:

$$Q(\omega) = \sqrt[3]{F(\omega)}. \quad (1.57)$$

Po to atliekama atvirkštinė diskrečioji Furjė transformacija (ADFT). Pritaikoma tiesinės prognozės analizė naudojant (1.34)–(1.43) formules ir gaunami *tiesinės suvokimo prognozės koeficientai* (TSPK) (1.3 pav.).



1.3 pav. Tiesinės suvokimo prognozės koeficientų radimo schema

Tiesinės suvokimo prognozės kepstro koeficientai. Tiesinės suvokimo prognozės kepstro koeficientai (TSPKK) apskaičiuojami tiesinės suvokimo prognozės koeficientams pritaikius kepstro apskaičiavimo (1.50)–(1.52) formules.

1.1.4. Atstumo matas

Atstumas naudojamas pavyzdžių palyginimo metoduose kaip pavyzdžių panašumo matas. Spektrinės analizės atveju apskaičiuojama norma (Junqua *et al.* 1996):

$$d_{NORM}(a_n, t_n) = k \sqrt[k]{\sum_{d=1}^D (a_{nd} - t_{nd})^k}, \quad (1.58)$$

čia a_{nd} yra n -ojo etalono požymių vektoriaus d -asis komponentas, t_{nd} – n -ojo nežinomo pavyzdžio požymių vektoriaus d -asis komponentas, D yra komponentų skaičius, k – normos eilė. Dažniausiai yra naudojamos pirmos eilės – Čebyšovo ir antros eilės – Euklido normos.

Tiesinės prognozės modelio atveju siūloma naudoti tokius matus kaip tikėtinumo santykio, logaritminio tikėtinumo santykio matas. Tikėtinumo santykio matas yra apskaičiuojamas:

$$d_{TS}(a_n, t_n) = \frac{a_n R a_n^T}{t_n R t_n^T} - 1, \quad (1.59)$$

čia R yra nežinomo pavyzdžio autokoreliacijos matrica.

Logaritminis tikėtinumo santykio matas apskaičiuojamas:

$$d_{LTS}(a_n, t_n) = \log \left(\frac{a_n R a_n^T}{t_n R t_n^T} \right). \quad (1.60)$$

Kepstro koeficientų atveju siūloma naudoti Euklido atstumą:

$$d_{EU}(a_n, t_n) = \sqrt{\sum_{d=1}^D (a_{nd} - t_{nd})^2}. \quad (1.61)$$

Taip pat Euklido atstumas yra apskaičiuojamas naudojant svorio funkciją:

$$d_{EUS}(a_n, t_n) = \sqrt{\sum_{d=1}^D w(d) (a_{nd} - t_{nd})^2}. \quad (1.62)$$

čia $w(d)$ yra svorio funkcija. Svorio funkcijai apskaičiuoti naudojami kepstro koeficientų indeksai $w(d)=d$, taikomos lango funkcijos.

1.1.5. Šnekos atpažinimo metodai

Šnekos signalo atpažinimo metodai yra skirstomi į tris grupes (Rabiner *et al.* 1993; Anusuya *et al.* 2009; Anusuya *et al.* 2011):

- Akustiniai – fonetiniai metodai.
- Pavyzdžiais grįsti metodai.
- Dirbtinio intelekto atpažinimo metodai.

Akustiniai fonetiniai metodai yra grįsti akustine fonetine teorija ir daro prielaidą, kad šnekamojoje kalboje yra baigtinis fonetinių vienetų skaičius ir kad šie fonetiniai vienetai gali būti aprašomi šnekos signalo laiko arba spektro savybėmis. Pirmajame akustinio fonetinio metodo etape vykdomas segmentavimas, kurio metu šnekos signalas yra suskirstomas į atkarpas, kur signalo akustinės savybės atstovauja vieną (arba kelis) fonetinį vienetą (arba klases). Tada pagal akustines savybes kiekvienam segmentui yra priskiriama fonetinė klasė. Antrajame etape iš gautų fonetinių klasių sekos nustatomas žodis (ar žodžių seka), kuris atitinka šnekos atpažinimo uždavinio apribojimus, pavyzdžiui, žodžiai turi atitikti žodžius, esančius žodyne, žodžių seka atitinka sintaksines taisykles bei semantinę reikšmę.

Pavyzdžiais grįstus metodus sudaro du etapai: apmokymo ir atpažinimo. Apmokymo metu sukuriamas žodynas, kurį sudaro apmokymo pavyzdžių rinkinys. Šie apmokymo pavyzdžiai gali būti šablonai arba modeliai, reprezentuojantys statistines pavyzdžių požymių charakteristikas. Atpažinimo metu nežinomas pavyzdys yra lyginamas su apmokymo pavyzdžiais. Remiantis gautais atstumų rezultatais, nežinomas pavyzdys yra priskiriamas tam tikrai klasei. Pavyzdžiais grįstų metodų konstravime išskiriami du atvejai. Pirmuoju atveju naudojamas atstumo matas – nežinomas pavyzdys priskiriamas tai klasei, kuriai gautas atstumas yra mažiausias. Antruoju atveju išskiriami statistiniai metodai, kurie remiasi tikimybiniais modeliais.

Dirbtinio intelekto atpažinimo metodai grindžiami akustinių fonetinių metodų ir pavyzdžiais grįstų metodų idėjomis. Dirbtinio intelekto atpažinimo metodai remiasi aspektais, pagal kuriuos žmonės pasitelkę savo intelektą vizualizuoja, analizuoja, priima sprendimus apie akustinius požymius.

Pavyzdžiui, šnekos segmentavimui ir klasių priskyrimui naudojami akustiniai fonetiniai metodai, tuo tarpu dirbtiniai neuroniniai tinklai taikomi apmokymui.

Disertacijoje pateiktas metodas šnekos atpažinimo požymių kokybei vertinti, kai klasifikavimui yra naudojami Euklido erdvės klasifikatoriai. Tai klasifikatoriai, naudojantys Euklido atstumą. Eksperimentiniuose tyrimuose naudojami klasifikatoriai – artimiausio kaimyno (AK), dinaminio laiko skalės kraipymo (DLSK) klasifikatoriai, todėl šie metodai aprašomi detaliau. Taip pat pristatomas dirbtinių neuroninių tinklų metodas.

Artimiausio kaimyno metodas. K -artimiausių kaimynų (K -AK) klasifikatorius yra mokymo su mokytoju klasifikatorius, naudojamas sistemoms apmokyti ir duomenims išgauti. K -AK – pavyzdžiu paremtas metodas, tiesiogiai dirba su apmokymo pavyzdžių duomenimis (Cover *et al.* 1967; Hua *et al.* 2010; Chee *et al.* 2009; Golipour *et al.* 2009; Zuo *et al.* 2008).

K -AK klasifikatorius nežinomam vektoriui randa K artimiausių apmokymo aibės vektorių. Nežinomas vektorius priskiriamas klasei, kurios vektorių yra daugiausia tarp artimiausių K kaimynų. Vieno artimiausio kaimyno atveju (AK) randamas vienas artimiausias apmokymo vektorius.

Apmokymo (etalonų) vektorių aibė $A = \{(a_1, f(a_1)), \dots, (a_R, f(a_R))\}$ yra sudaryta iš a_R – apmokymo vektorių, $f(a_R)$ – apmokymo vektoriams priskirtų klasių, kur R yra visų apmokymo vektorių skaičius, $X = \{X^1, X^2, \dots, X^C\}$ – klasių aibė, C – klasių skaičius. K -AK algoritmas skaičiuoja atstumą (mūsų atveju Euklido atstumą) tarp nežinomo vektoriaus t_z ir visų apmokymo vektorių, kur $1 \leq z \leq Z$, Z – nežinomų vektorių skaičius. Randami K -artimiausi apmokymo vektoriai. Nežinomo vektoriaus priskyrimas klasei išreiškiamas:

$$\hat{f}(t_z) = \arg \max_{X^i \in X} \sum_{j=1}^K \delta(X^i, f(a_j)), 1 \leq i \leq C, \quad (1.63)$$

čia $\delta(\cdot)$ funkcija nežinomą vektorių priskiria klasei, kuriai priklauso daugiausiai išrinktųjų artimiausių kaimynų vektorių (AK atveju $K = 1$).

AK atpažinimo sistemos algoritmo sudėtingumas dviejų klasių atveju yra $O(2R)$ ($N = 2R$, kur N yra visų vektorių skaičius). K -AK kaimyno klasifikavimo klaida randama klaidingai suklasifikuotų testuojamų vektorių skaičių E padalinus iš visų testuojamų vektorių skaičiaus Z :

$$AK_{KL} = \frac{E}{Z}. \quad (1.64)$$

Dinaminio laiko skalės kraipymo metodas. Dinaminio laiko skalės kraipymo (DLSK) klasifikatorius priklauso pavyzdžių palyginimo metodų grupei, naudoja dinaminį programavimą – palygina pavyzdžius minimizuodamas atstumą tarp jų (Sakoe *et al.* 1978; Sakoe *et al.* 1979; Myers *et al.* 1980).

Tarkime, turime etaloną $A = \{a_1, a_2, \dots, a_R\}$ ir nežinomą pavyzdį $T = \{t_1, t_2, \dots, t_Z\}$, kur a_i yra i -asis etalono požymių vektorius, t_j – j -asis nežinomo pavyzdžio vektorius. Geometrinio interpretavimo atžvilgiu, etalono ir pavyzdžio vektorių indeksai išdėstomi taip, kad suformuotų tinklelį, kurio dydis yra $R \times Z$. Kiekvienas tinklelio taškas (i, j) nusako atstumą $d(i, j)$ tarp i -ojo etalono vektoriaus, ir j -ojo nežinomo pavyzdžio vektoriaus (mūsų atveju naudojamas Euklido atstumas). Naudojant suformuotą tinklelį yra surandama kraipymo trajektorija $W = \{w_1, w_2, \dots, w_K\}$, kur $w_k = (i_k, j_k)$ ir $\max(R, Z) \leq K \leq R + Z - 1$. Tam, kad būtų surandama optimali kraipymo trajektorija (dinaminio laiko skalės kraipymo atstumas tarp nagrinėjamų pavyzdžių), DLSK algoritmas sprendžia minimizavimo uždavinį – nustato trajektoriją, kuri minimizuoja atstumą tarp nagrinėjamų pavyzdžių (Rabiner *et al.* 1978):

$$d_{DLSK}(A, T) = \min \frac{1}{K} \sum_{k=1}^K w_k. \quad (1.65)$$

Didėjant nagrinėjamų pavyzdžių dydžiams, eksponentiškai didėja ir galimų trajektorijų skaičius. Dinaminis programavimas suranda optimalią trajektoriją, sumuodamas atstumą tarp nagrinėjamų pavyzdžių bei minimalų atstumą trajektorijos, kurios pradžia yra taške $(1, 1)$ o pabaiga – (i, j) :

$$D(i, j) = d(i, j) + \min \left\{ \begin{array}{l} D(i-1, j-1) \\ D(i-1, j) \\ D(i, j-1) \end{array} \right\}, \quad (1.66)$$

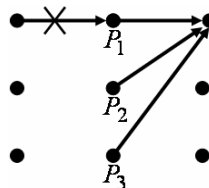
čia $D(i, j)$ – suminis atstumas, apskaičiuotas trajektorijai, kurios pradžia yra taške $(1, 1)$, o pabaiga (i, j) , $d(i, j)$ yra atstumas tarp nagrinėjamų pavyzdžių i -ojo ir j -ojo vektorių. Galimų trajektorijų skaičius yra didelis, todėl kraipymo trajektorijai taikomi apribojimai (Berndt *et al.* 1994):

1. *Monotoniškumas ir nuoseklumas.* Trajektorijos taškai turi būti išdėstyti monotoniškai, galimi perėjimai tik į tolimesnius taškus:

$$0 \leq i_k - i_{k-1} \leq 1, \quad 0 \leq j_k - j_{k-1} \leq 1. \quad (1.67)$$

2. *Galo taškų apribojimai.* Nurodoma kraipymo trajektorijos pradžia ir pabaiga. Paprasčiausiu atveju pradžia yra $(1, 1)$, pabaiga (R, Z) , t.y. trajektorija prasideda pirmajame tinklelio taške, baigiasi paskutiniajame.

3. *Lokalus krypties apribojimas.* Kraipymo trajektorijų paieškos skaičius gali būti apribojamas nustatant galimų ėjimų skaičių viena kryptimi. Darbe naudojamas lokalus Itakuros krypties apribojimas (1.4 pav.), kur P yra galimų perėjimų seka, nusakoma koordinatėmis $P \rightarrow (p_1, q_1)(p_2, q_2) \dots (p_R, q_R)$.



1.4 pav. Lokalus Itakuros krypties apribojimas, $P_1 \rightarrow (1, 0)$, $P_2 \rightarrow (1, 1)$, $P_3 \rightarrow (1, 2)$ iš eilės einantys perėjimai yra negalimi

4. *Globalus krypties apribojimas.* Taip pat kraipymo trajektorijų paieškos skaičius gali būti apribojamas apibrėžiant paieškos sritį. Apribojimai iš kairės ir dešinės išreiškiami (1.5 pav.):

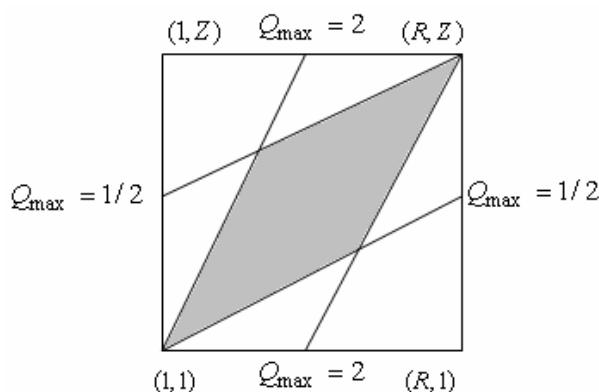
$$1 + (D(i) - 1) / Q_{\max} \leq D(j) \leq 1 + Q_{\max} (D(i) - 1), \quad (1.68)$$

$$Z + Q_{\max} (D(i) - R) \leq D(j) \leq Z + (D(i) - R) / Q_{\max}, \quad (1.69)$$

čia Q_{\max} yra maksimalus nuokrypio koeficientas. Dažnai naudojamas Itakuros lygiagretainio apribojimas su $Q_{\max} = 2$, 1.5 paveikslas. Išraiškos (1.63) kairiąją pusę sulyginus su (1.64) dešiniąją, o (1.63) dešinę su (1.64) kairiąją, gaunamos ribinės pavyzdžių ilgių išraiškos:

$$R = Q_{\max} (Z - 1) + 1, \quad (1.70)$$

$$Z = Q_{\max} (R - 1) + 1. \quad (1.71)$$



1.5 pav. Globalus Itakuros lygiagretainio krypties apribojimas, $Q_{\max} = 2$

DLSK atpažinimo sistemos algoritmo sudėtingumas yra $O(R^2)$, ($N = 2R$, kur N yra visų vektorių skaičius). DLSK klasifikavimo klaida randama klaidingai suklasifikuotų testuojamų pavyzdžių skaičių E padalinus iš visų testuojamų pavyzdžių skaičiaus Z :

$$DLSK_{KL} = \frac{E}{Z}. \quad (1.72)$$

Dirbtinių neuroninių tinklų metodas. Neuronų modelis susideda iš daugelio įėjimų, kurie yra sumuojami padauginus juos iš tam tikrų, svoriais vadinamų koeficientų, ir nukreipiami aktyvavimo funkcijai. Neuronų modelis

yra išreiškiamas šia formule (Rosenblatt 1958; Dzemyda *et al.* 2008; Chung *et al.* 2009; Alsmadi *et al.* 2009):

$$y = f(h) = f\left(\sum_{d=1}^D w_d x_d + w_0\right), \quad (1.73)$$

čia y – neurono išėjimo reikšmė, x_d – įėjimo vektoriaus elementai, w_d yra svorių vektoriaus elementai, w_0 – neurono slenkstis, $f(\cdot)$ – aktyvavimo funkcija. Įvedamas nulinis įėjimas x_0 , kuris dažniausiai yra pastovus, $x_0 = 1$. Naudojamos tokios aktyvavimo funkcijos kaip slenkstinė, sigmoidinė, hiperbolinio tangento.

Paprasčiausias neuroninis tinklas yra perceptronas, sudarytas iš vieno sluoksnio K neuronų, sujungtų su D įėjimų. Kiekvienas perceptrono išėjimas y_k yra įėjimų x_1, x_2, \dots, x_D funkcija, kuri apskaičiuojama pagal formulę:

$$y_k = f(h_k) = f\left(\sum_{d=0}^D w_{kd} x_d\right), \quad 1 \leq k \leq K. \quad (1.74)$$

Perceptrono mokymo procese svoriai keičiami taip, kad tinklo išėjimo vektorius būtų kuo artimesnis norimų reikšmių vektoriui t_1, t_2, \dots, t_K . Paklaidos funkcija išreiškiama formule:

$$PF = \frac{1}{2} \sum_{k=1}^K (t_k - y_k)^2. \quad (1.75)$$

Paprasčiausias metodas perceptronui mokyti yra delta taisyklė, kuri siekia minimizuoti perceptrono išėjimo klaidą: po kiekvienos svorių pataisos iteracijos daroma kita pataisa, proporcinga nuostolių funkcijos išvestinei pagal visus svorių vektoriaus komponentus (Widrow *et al.* 1960; Hui *et al.* 1994; Sandon 1988; Liu *et al.* 2010):

$$w_{kd}(t+1) = w_{kd}(t) + \Delta w_{kd}(t), \quad (1.76)$$

$$\Delta w_{kd}(t) = \eta(t_k - y_k)x_d, \quad (1.77)$$

čia η yra mokymo greičio parametras, kuriuo reguliuojamas perceptrono mokymo greitis. Klasifikavimo klaida randama klaidingai suklasifikuotų testuojamų vektorių skaičių E padalinus iš testuojamų vektorių skaičiaus Z :

$$TK_{KL} = \frac{E}{Z}. \quad (1.78)$$

1.2. Šnekos atpažinimo požymių kokybės vertinimo analitinė apžvalga

Kokybė tai matas, kuris nusako tiriamojo objekto charakteristikų atitikimo lygį apibrėžtiems reikalavimams. Jeigu objekto charakteristikos atitinka apibrėžtus reikalavimus, yra pasiekama aukšta objekto kokybė (Bareisa *et al.* 2007; Targamadžė *et al.* 2010; Eidukas *et al.* 2010). Kokybiška požymių sistema yra klasifikavimo klaidą minimizuojanti požymių sistema. Pateiktame metode šnekos signalų atpažinimo požymių kokybei vertinti naudojamos metrikos, tuo tarpu klasifikavimo klaidos minimizavimą atitinka metrikų minimizavimas. Metrikai įgijus minimalią reikšmę yra laikoma, kad klasės yra gerai atskiriamos, metrikai įgijus maksimalią reikšmę – klasės yra blogai atskiriamos. Požymių sistemos kokybės rodiklis yra metrikų rezultatus apibendrinantis matas. Požymių kokybės rodiklio skalė yra 0 %–100 %, kur 0 % rodo žemiausią požymių sistemos kokybę, o 100 % – aukščiausią požymių sistemos kokybę.

Pirmiausia apžvelgsime klasifikavimo sudėtingumo priežastis: klasių dviprasmiškumą, klasių ribų sudėtingumą (Basu *et al.* 2006). Klasių dviprasmiškumas galimas, kai klasių sąvokos yra prastai apibrėžtos ir dėl šios priežasties klasifikatorius neatskiria klasių. Pavyzdžiui, raidė ‚1‘ ir skaitmuo ‚1‘ yra vaizduojamos tokiu pat būdu, turėtų priklausyti tai pačiai klasei. Kitas galimas klasių dviprasmiškumo atvejis – kai klasių sąvokos yra gerai apibrėžtos, tačiau naudojamos nekokybiškos požymių sistemos, dėl kurių klasės požymiai persidengia ir klasės tampa blogai atskiriamos. Klasės, turinčios sudėtingas ribas, yra blogiau atskiriamos nei klasės, turinčios

nesudėtingas ribas. Klasifikatoriaus apmokymo metu turint mažai duomenų, padidėja galimybė neteisingai suklasifikuoti duomenis. Tai aktualu esant dideliame vektorių komponentų skaičiui.

Automatinės šnekos atpažinimo sistemos sudarymas yra sudėtingas uždavinys. Kokybiškų požymių parinkimas yra svarbus etapas šnekos atpažinimo sistemos projektavimo procese, nes naudojant kokybiškus požymius gaunama maža šnekos atpažinimo sistemos klaida. Egzistuojant šnekos atpažinimo požymių sistemų įvairovei (pvz. kepstro koeficientai, melų skalės kepstro koeficientai, tiesinės prognozės koeficientai), iškyla klausimas, kaip parinkti kokybišką požymių sistemą.

1.2.1. Klasifikavimo klaidos apskaičiavimas

Literatūroje išskiriami požymių kokybės vertinimo metodai:

- Požymių sistemos kokybės vertinimas yra pagrįstas eksperimento vykdytojo intuicija.
- Klasifikavimo klaidos apskaičiavimas.

Kai požymių sistemos kokybė yra vertinama pagal eksperimento vykdytojo intuiciją, eksperimento vykdytojas parenka požymių sistemą remdamasis savo žiniomis apie požymių sistemas. Tačiau šis metodas yra pagrįstas eksperimento vykdytojo žiniomis bei spėjimu.

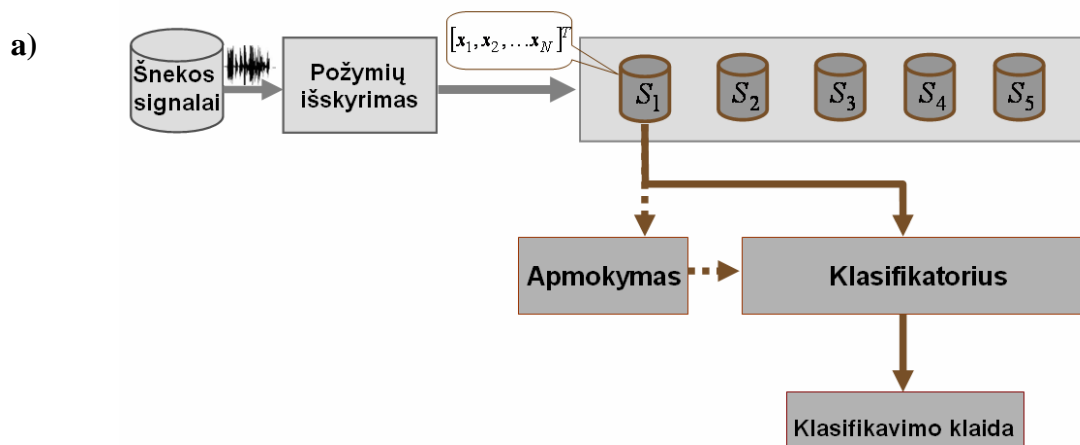
Požymių sistemos kokybę vertinant klasifikavimo klaida, klasifikavimas yra vykdomas su kiekviena tiriama požymių sistema ir kokybiška požymių sistema išrenkama ta požymių sistema, su kuria gaunama mažiausia klasifikavimo klaida (Voitovetsky *et al.* 1997; Webb 2002; Bernado–Mansilla *et al.* 2005; Chen *et al.* 2004; Soryani *et al.* 2005; Gavrilis *et al.* 2008; Krishna *et al.* 2011; Wittenburg *et al.* 2010). Požymių kokybės vertinimo klasifikavimo klaida trūkumai:

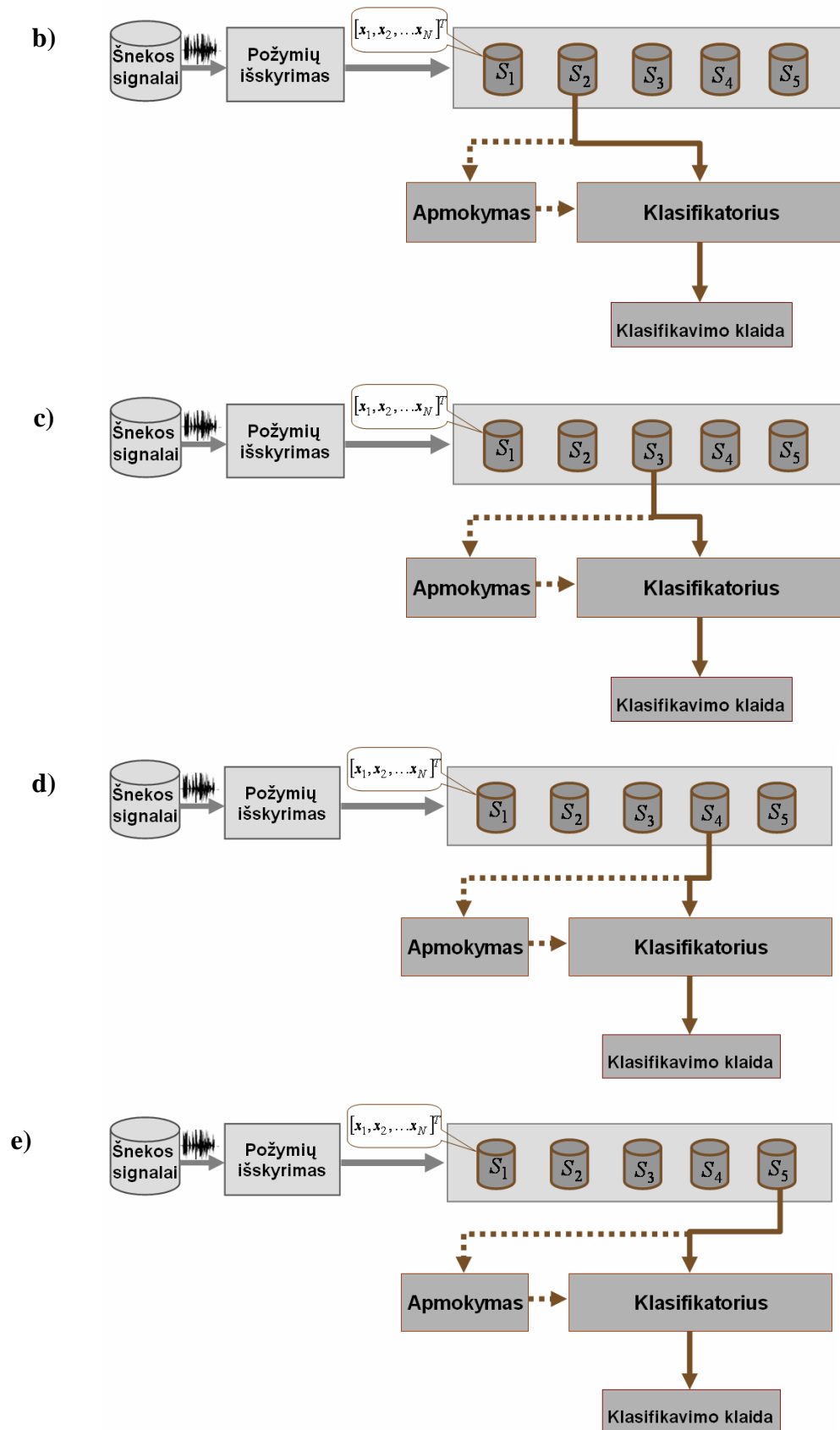
- Kokybiškai požymių sistemai nustatyti naudojant klasifikavimo klaidą, su kiekviena tiriama požymių sistema reikia vykdyti klasifikavimo eksperimentus. Tarkime, turime penkias tiriamas požymių sistemas S_1 , S_2 , S_3 , S_4 , S_5 . Norint nustatyti kokybišką požymių sistemą naudojant

klasifikavimo klaidą, reikia penkis kartus vykdyti klasifikavimo eksperimentus (1.6 pav.).

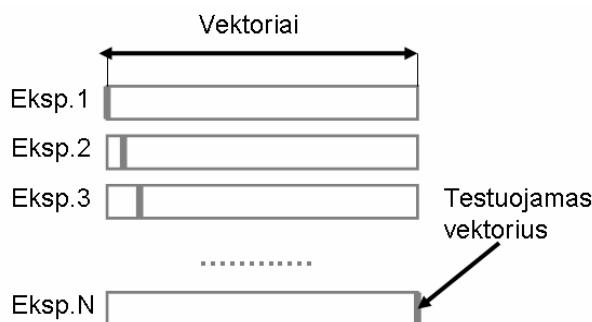
- Klasifikavimo klaidos apskaičiavimas reikalauja daug darbo atliekant apmokymą, sudarant modelius.
- Klasifikavimo klaidos apskaičiavimas reikalauja didelių skaičiavimo resursų, ypač naudojant tokias testavimo technikas kaip slenkančiojo vektoriaus (angl. *leave-one-out*). Naudojant slenkančio vektoriaus testavimo techniką, vienas vektorius yra išimamas iš duomenų aibės, apmokymas vyksta naudojant visus vektorius išskyrus išimtąjį, testuojama su išimtuoju vektoriumi (Larranaga *et al.* 2002; Duda *et al.* 2001). Kitoje iteracijoje išimtas vektorius grąžinamas į duomenų aibę, išimamas kitas vektorius (1.7 pav.). Procedūra kartojama su visais vektoriais.

Reikalingas naujas požymių kokybės vertinimo metodas. Pagrindiniai šios disertacijos uždaviniai – pateikti naują šnekos signalų atpažinimo požymių kokybės vertinimo metodą, grindžiamą metrikų naudojimu. Parodyti, kad pateiktas šnekos signalų požymių vertinimo kokybės metodas Euklido erdvėje suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį, nereikalaujant klasifikavimo eksperimentų vykdymo. Įvertinti pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo algoritmo sudėtingumą. Sukurti programinę įrangą ir eksperimentiškai patvirtinti pateikto metodo teisingumą.





1.6 pav. Požymių kokybės nustatymas naudojant klasifikavimo klaidą, penkių tiriamų požymių sistemų atveju: a) pirmoji; b) antroji; c) trečioji; d) ketvirtoji; e) penktoji



1.7 pav. Slenkančio vektoriaus testavimo technika

Tam, kad pateiktume metodą požymių kokybei vertinti, atlikome metrikų, naudojamų panašioms literatūroje uždaviniams spręsti, analitinę apžvalgą. Šios metrikos skirstomos į tris grupes: geometrinės metrikas, informacijos teorijos metrikas, statistines metrikas (1.1 lentelė). Metrikos taikomos šiems uždaviniams spręsti (detalūs metrikų aprašymai pateikti kituose skyriuose):

- Duomenų struktūros tyrimas. Remiantis metrikomis nustatomas duomenų struktūros sudėtingumas. Duomenų struktūros tyrimo uždaviniui naudojamos geometrinės grupės metrikos (Ho *et al.* 2002; Elizondo *et al.* 2009; Ho 2001; Ho 2002; Bernado–Mansilla *et al.* 2006; Bernado–Mansilla *et al.* 2004; Sotoca *et al.* 2005; Mollineda *et al.* 2005; Lebourgeois *et al.* 1996; Ho 2000).
- Dirbtinių duomenų generavimas. Taikant metrikas generuojami norimo sudėtingumo dirbtiniai duomenys. Naudojamos geometrinės grupės metrikos (Macia *et al.* 2008).
- Požymių išrinkimas. Požymių išrinkimo metu yra išrenkami informatyviausi požymiai (Guyon *et al.* 2003; Guo *et al.* 2008; Miche *et al.* 2007). Vienas iš požymių išrinkimo etapų yra metrikos parinkimas, kuri tampa optimizavimo kriterijumi išrenkant požymius. Šiam tikslui naudojamos informacijos teorijos (Kerroum *et al.* 2008; Arauzo–Azofra *et al.* 2009; Peng *et al.* 2005; Balagani *et al.* 2010; Bao *et al.* 2006; Chen 1971; Doan *et al.* 2004; Battiti 1994), statistinės metrikos (Webb 2002; Theodoridis *et al.* 1999; Costa *et al.* 1999; Malina 1981; Miyamoto *et al.* 2003).

1.1 lentelė. Metrikų grupės ir metrikos

Metrikų grupė		Metrika
Geometrinės metrikos	Požymių atskiriamumas	Požymių efektyvumas
		Maksimalus Fišerio diskriminantinis santykis
		Požymių persidengimo sritis
	Klasių atskiriamumas	Požymių kiekis klasės ribose
		Klasių artimiausių kaimynų atstumų santykis
		Klasės ribos peržengimo kiekis
		ϵ -kaimynystės klasterių skaičius
		Tiesinio klasifikatoriaus klaida
		Artimiausio kaimyno klasifikavimo klaida
		Tiesinio klasifikatoriaus netiesiškumas
Artimiausio kaimyno klasifikatoriaus netiesiškumas		
Informacijos teorijos metrikos	Entropija	
	Sąlyginė entropija	
	Jungtinė entropija	
	Informacijos kiekis	
Statistinės metrikos	Hipotezės tikrinimas	
	Atstumo matu paremti kriterijai	
	Kovariacijos matrica paremti kriterijai	
	Koreliacija	
	Skirstinio kreivės formos kriterijai	
	Vidutinis vektorių skaičius komponentų skaičiaus atžvilgiu	

1.2.2. Geometrinės metrikos

Geometrinės metrikos pagal savo charakteristikas yra skirstomos į dvi grupes (1.1 lentelė):

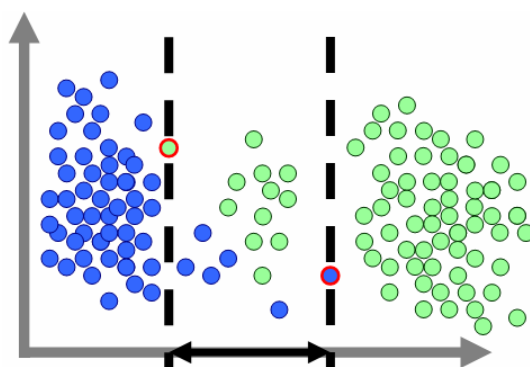
- Požymių atskiriamumo grupė – atsižvelgia į atskiro požymio komponento informatyvumą atskiriant klases.
- Klasių atskiriamumo grupė – atsižvelgia į klasių ribos formą, jos sudėtingumą, klases sudarančių klasterių formą ir tarpusavio padėtį.

Geometrinės metrikos yra glaudžiai susijusios su klasifikavimo sudėtingumo priežastimis, pavyzdžiui, klasių persidengimu, klasių sudėtingų ribų identifikavimu.

Požymių efektyvumo metrika. Klasių vektoriai yra išsidėstę tam tikrame intervale. Klasės yra gerai atskiriamos, jeigu jų intervalai yra nepersidengiantys bent vieno požymių komponento atžvilgiu (Ho *et al.* 2002; Basu *et al.* 2006).

Požymių efektyvumo metrika yra apskaičiuojama:

- Pašalinami vektoriai, patenkantys į klasių persidengiančius intervalus kiekvieno požymių komponento atžvilgiu – gaunamos klasės be persidengiančių intervalų (1.8 pav.).
- Apskaičiuojamas nepersidengiančių vektorių ir visų vektorių skaičiaus santykis kiekvieno požymių komponento atžvilgiu.
- Požymių efektyvumas nustatomas išrenkant gautą didžiausią santykio reikšmę.



1.8 pav. Klasių vektoriai patenkantys į persidengiantį intervalą

Metrika, nurodyma klasių atskiriamumą, atsižvelgia į plokštumą, statmeną požymių koordinatinių ašims, todėl geriausiai identifikuoja klasių atskiriamumą, kai klasės atskiriamos plokštuma, statmena ašims. Kai klasės yra gerai atskiriamos pasvira plokštuma, metrika gali nurodyti blogą klasių atskiriamumą.

Maksimalus Fišerio diskriminantinio santykio metrika. Maksimalus Fišerio diskriminantinis santykis yra apskaičiuojamas (Ho *et al.* 2002; Bernado–Mansilla *et al.* 2005; Basu *et al.* 2006):

- Randamas Fišerio diskriminantinis santykis kiekvienam požymių komponentui.
- Išrenkama maksimali Fišerio diskriminantinio santykio reikšmė.

Maksimalaus Fišerio diskriminantinio santykio metrika išreiškiama formule:

$$FM_d = \frac{\sum_{i=1, j=1, i \neq j}^C p^i p^j (\mu_d^i - \mu_d^j)^2}{\sum_{i=1}^C p^i \sigma_d^i}, \quad (1.79)$$

$$F2 = \max_d FM_d, \quad (1.80)$$

čia FM_d yra požymių d -ojo komponento maksimalus Fišerio diskriminantinis santykis, kur $1 \leq d \leq D$, D – požymių komponentų skaičius, μ_d^i , μ_d^j – i -osios ir j -osios klasių, d -ojo komponento vidurkiai, $1 \leq i \leq C$, $1 \leq j \leq C$, $i \neq j$, kur C yra klasių skaičius, σ_d^i yra i -osios klasės d -ojo komponento dispersija, p^i , p^j yra i -osios ir j -osios klasių proporcijos, apskaičiuojamos:

$$p^i = N_i / N, \quad (1.81)$$

$$p^j = N_j / N, \quad (1.82)$$

čia N_i , N_j yra i -osios ir j -osios klasės vektorių skaičius, N – visų klasių vektorių skaičius. Metrika tinkamiausia naudoti duomenims, pasiskirsčiusiems pagal normalųjį dėsnį.

Požymių persidengimo srities metrika. Požymių persidengimo srities metrika identifikuoja klasių atskiriamumą randant klasių persidengimo sritį (Ho *et al.* 2002; Basu *et al.* 2006).

Požymių persidengimo srities metrika apskaičiuojama:

- Randami kiekvienos klasės mažiausią ir didžiausią reikšmę turintys vektoriai kiekvieno požymių komponento atžvilgiu.
- Nustatomas klasių persidengimo atstumas bei klasių užimamas atstumas.

- Apskaičiuojamas persidengimo atstumo ir klasių užimamo atstumo santykis.
- Santykiai, apskaičiuoti kiekvienam požymių komponentui, yra sudauginami.

Požymių persidengimo srities metrika išreiškiama formule:

$$F3 = \sum_{i=1, j=1, i \neq j}^C \prod_{d=1}^D \frac{\min(\max x_d^i, \max x_d^j) - \max(\min x_d^i, \min x_d^j)}{\max(\max x_d^i, \max x_d^j) - \min(\min x_d^i, \min x_d^j)}, \quad (1.83)$$

čia $\max x_d^i$ ir $\max x_d^j$ yra i -osios bei j -osios klasių, d -ojo komponento maksimalios reikšmės, $\min x_d^i$, $\min x_d^j$ – atitinkamai minimalios reikšmės.

Metrika įgyja nulį, kai sutampa klasių minimalios ir maksimalios reikšmės. Tuo atveju metrikos reikšmė nenurodo tikrojo klasių persidengimo lygio. Taip pat operuojant su neigiamomis reikšmėmis dėl sandaugos metrika įgyja teigiamą reikšmę ir ši reikšmė klaidinančiai nurodo, kad klasės yra gerai atskiriamos. Taigi metrika netinkama šnekos signalams taikyti, nes šnekos signalų požymiai gali įgyti ir neigiamas reikšmes.

Požymių kiekio klasės ribose metrika. Požymių kiekio klasės ribose metrika apskaičiuojama randant minimalaus jungiančiojo medžio (MJM) briaunomis sujungtas skirtingų klasių viršūnes (Ho *et al.* 2002). Metrika naudojama pateiktame metode požymių kokybei vertinti ir yra detaliau aprašoma 2.2 skyriuje.

Klasių artimiausių kaimynų atstumų santykio metrika. Klasių artimiausių kaimynų atstumų santykio metrika palygina atstumą tarp vektorių klasės viduje su atstumu tarp klasių (Bernado–Mansilla *et al.* 2005). Metrika naudojama pateiktame metode požymių kokybei vertinti, detaliau aprašoma 2.2 skyriuje.

Klasės ribos peržengimo kiekio metrika. Klasės ribos peržengimas fiksuojamas, kai klasės vektorius patenka į kitos klasės ribas (Li *et al.* 1998). Metrika naudojama pateiktame metode požymių kokybei vertinti, detaliau aprašoma 2.2 skyriuje.

ε-kaimynystės klasterių skaičiaus metrika. *ε*-kaimynystės klasterių skaičiaus metrika atsižvelgia į klases sudarančių klasterių skaičių (Ho *et al.* 2002; Lebourgeois *et al.* 1996). Ši metrika nusako klasių vidinę struktūrą.

ε-kaimynystės klasterių skaičiaus metrika yra apskaičiuojama:

- Tos pačios klasės požymių vektoriai apjungiami į apskritimo formos klasterius. Klasteris „auga“ kol vektoriai nuo apskritimo centro yra nutolę ne toliau nei dydis *ε* ir kol nėra sutinkamas vektorius priklausantis kitai klasei. Apskritimo centras yra atsitiktinai pasirinktas klasės vektorius. Dydis *ε* yra minimalus Euklido atstumas tarp priešingų klasių vektorių.
- Klasterių skaičius padalinamas iš visų vektorių skaičiaus.

Dydis *ε* išreiškiamas formule:

$$\varepsilon = \min_k d(x_n^i, x_k^j), \quad (1.84)$$

čia $\min d(x_n^i, x_k^j)$ – minimalus Euklido atstumas tarp *i*-osios klasės *n*-ojo vektoriaus ir *j*-osios klasės *k*-ojo vektoriaus, $1 \leq n \leq N_i$, N_i – *i*-osios klasės vektorių skaičius, $1 \leq k \leq N_j$, N_j – *j*-osios klasės vektorių skaičius.

Metrika aprašoma išraiška:

$$G4 = \frac{B}{N}, \quad (1.85)$$

čia *B* – gautųjų klasterių skaičius, *N* – visų vektorių skaičius.

Tiesinio klasifikatoriaus klaidos metrika. Metrika apskaičiuojama kaip tiesinio klasifikatoriaus klaida (Ho *et al.* 2002; Bernado–Mansilla *et al.* 2005; Basu *et al.* 2006). Metrika randama klaidingai suklasifikuotų testuojamų vektorių skaičių padalinus iš visų vektorių skaičiaus. Tiesinio klasifikavimo klaidos apskaičiavimo principas yra pateikiamas 1.1.3 skyriuje.

Artimiausio kaimyno klasifikatoriaus klaidos metrika. Artimiausio kaimyno klasifikatoriaus klaidos metrika apskaičiuojama kaip AK klasifikavimo klaida

(Ho *et al.* 2002; Bernado–Mansilla *et al.* 2005; Basu *et al.* 2006). Metrika apskaičiuojama klaidingai suklasifikuotų testuojamų vektorių skaičių padalinus iš visų vektorių skaičiaus. Detalus AK klasifikavimo klaidos apskaičiavimo algoritmas yra pateikiamas 1.1.3 skyriuje.

Tiesinio klasifikatoriaus netiesiškumo metrika. Tiesinio klasifikatoriaus netiesiškumo metrika apskaičiuojama randant tiesinio klasifikatoriaus klaidą, kai testavimui yra naudojami interpoliuoti duomenys (Ho *et al.* 2002; Hoekstra *et al.* 1996). Metrika randama:

- Sukuriami testavimo duomenys, pritaikius tiesinę interpoliaciją apmokymo duomenims.
- Tiesinis klasifikatorius apmokomas naudojant apmokymo duomenis.
- Tiesinio klasifikatoriaus klaidai apskaičiuoti naudojami interpoliuotieji testavimo duomenys.

Tiesinio klasifikatoriaus netiesiškumo metrika apskaičiuojama klaidingai suklasifikuotų testuojamų vektorių skaičių (testavimui naudojami interpoliuoti duomenys) padalinus iš visų vektorių skaičiaus.

Artimiausio kaimyno klasifikatoriaus netiesiškumo metrika. Kaip ir tiesinio klasifikatoriaus netiesiškumo metrika, AK klasifikatoriaus netiesiškumo metrika apskaičiuojama randant AK klasifikatoriaus klaidą, testavimui naudojant interpoliuotus duomenis (Ho *et al.* 2002; Hoekstra *et al.* 1996). Metrika apskaičiuojama:

- Sukuriami testavimo duomenys, pritaikius tiesinę interpoliaciją apmokymo duomenims.
- AK klasifikatorius apmokomas naudojant apmokymo duomenis.
- AK klasifikavimo klaidai apskaičiuoti naudojami interpoliuotieji testavimo duomenys.

AK klasifikatoriaus netiesiškumo metrika randama klaidingai suklasifikuotų testuojamų vektorių skaičių (testavimui naudojami interpoliuoti duomenys) padalinus iš visų vektorių skaičiaus.

1.2.3. Informacijos teorijos metrikos

Informacijos teorijos metrikos charakterizuoja atsitiktinių dydžių elgseną, suteikia galimybę apskaičiuoti jų tikimybę (Arndt 2001; Cover *et al.* 2006). Požymių išrinkimo uždaviniams naudojama entropija, sąlyginė entropija, jungtinė entropija ir informacijos kiekis (Sohn 1999; Kerroum *et al.* 2008; Battiti 1994).

Entropija nusako atsitiktinio dydžio atsitiktinumo lygį. Klasių entropija išreiškiama:

$$H(K) = -\sum_{i=1}^C p(k^i) \log_2 p(k^i), \quad (1.86)$$

čia k^i yra i -oji klasė, $p(k^i)$ – apriorinė i -osios klasės tikimybė, $1 \leq i \leq C$, C yra klasių skaičius.

Sąlyginė entropija (angl. *conditional entropy*) nurodo atsitiktinio dydžio atsitiktinumo lygį, kai kito atsitiktinio dydžio reikšmė yra žinoma. Klasių ir požymių sąlyginė entropija apskaičiuojama:

$$H(K | X) = \sum_i^C \sum_j^N p(k^i, x_j) \log_2 \frac{p(x_j)}{p(k^i, x_j)}, \quad (1.87)$$

čia $X = \{x_1, x_2, \dots, x_N\}$ yra požymių vektorius, N – vektorių skaičius, $p(x_j)$ yra apriorinė j -ojo požymių elemento tikimybė, $1 \leq j \leq N$.

Jungtinė entropija (angl. *joint entropy*) parodo priklausomybės lygį tarp atsitiktinių dydžių. Klasių ir požymių jungtinė entropija apskaičiuojama:

$$H(K, X) = -\sum_{i=1}^C \sum_{j=1}^N p(k^i, x_j) \log_2 p(k^i, x_j). \quad (1.88)$$

Informacijos kiekis (angl. *mutual information*) parodo informacijos kiekį, kurį vienas atsitiktinis dydis turi apie kitą atsitiktinį dydį. Informacijos kiekis tarp klasių ir požymių yra apskaičiuojamas:

$$I(K, X) = H(K) - H(K | X) = \sum_{i=1}^C \sum_{j=1}^N p(k^i, x_j) \log_2 \frac{p(k^i, x_j)}{p(k^i)p(x_j)}. \quad (1.89)$$

Požymių išrinkimo metu išrenkami požymiai, kurie turi didžiausią informacijos kiekį nagrinėjamų klasių atžvilgiu (Peng *et al.* 2005).

1.2.4. Statistinės metrikos

Šiame skyriuje apžvelgiamos statistinės metrikos, naudojamos požymiams išrinkti (Theodoridis *et al.* 1999; Webb 2002; Costa *et al.* 1999).

Hipotezės tikrinimas. Požymių išrinkimo metu hipotezės tikrinimas gali būti naudojamas kaip optimizavimo kriterijus (Theodoridis *et al.* 1999). Nustatoma, ar skirtingų klasių atskiro požymio reikšmės skiriasi statistiškai reikšmingai. Iškeliama nulinė H_0 ir alternatyvi H_1 hipotezės:

H_1 : klasių požymio reikšmės nesiskiria statistiškai reikšmingai.

H_0 : klasių požymio reikšmės skiriasi statistiškai reikšmingai.

Jeigu požymio reikšmės skiriasi statistiškai reikšmingai, požymis yra pakankamai informatyvus ir dalyvauja kitame požymių išrinkimo etape.

Hipotezės tikrinimas nustato, ar skiriasi nagrinėjamų klasių požymiai, atsižvelgiant į jų vidurkius. Tačiau šios informacijos nepakanka, kad būtų užtikrinamas klasių atskiriamumas. Klasių požymio vidurkiai gali ženkliai skirtis, tačiau dispersija gali būti didelė ir sąlygoti klasių persidengimą.

Atstumo matu paremti kriterijai. Tikimybiniai atstumai (angl. *probabilistic distance*) taip pat naudojami kaip optimizavimo kriterijus požymių išrinkimo uždavinyje (Theodoridis *et al.* 1999, Webb 2002; Puzicha *et al.* 1999; Wachter *et al.* 2007). Naudojant tikimybinius atstumus yra matuojamas atstumas tarp dviejų sąlyginių tikimybinio tankio funkcijų $p(x|k^i)$ ir $p(x|k^j)$, čia k^i , k^j yra i -oji ir j -oji klasės, x – požymio reikšmė. Pateiksime keletą tikimybinių atstumų.

Bratačaro atstumas (angl. *Brattacharyya distance*):

$$J_B = -\log \int \left(p(x|k^i) p(x|k^j) \right)^{1/2} dx. \quad (1.90)$$

Divergencijos atstumas (angl. *Divergence distance*):

$$J_D = \int [p(x|k^i) - p(x|k^j)] \log \left(\frac{p(x|k^i)}{p(x|k^j)} \right) dx. \quad (1.91)$$

Tikimybinių atstumų vertinimas yra sudėtingas dėl naudojamų tankių ir integralų apskaičiavimo. Priėmę prielaidą, kad duomenys yra pasiskirstę pagal normaliuosius dėsnius $N(\mu^i, \Sigma^i)$ ir $N(\mu^j, \Sigma^j)$, pateiksime formules šiems atstumo matams apskaičiuoti, kur μ^i , μ^j – i -osios ir j -osios klasių vidurkiai, Σ^i , Σ^j yra i -osios ir j -osios klasių kovariacinės matricos.

Bratačaro atstumas:

$$d_B = \frac{1}{4} (\mu^i - \mu^j)^T (\Sigma^i + \Sigma^j)^{-1} (\mu^i - \mu^j) + \frac{1}{2} \log \left(\frac{|\Sigma^i + \Sigma^j|}{2 \left(|\Sigma^i| |\Sigma^j| \right)^{1/2}} \right). \quad (1.92)$$

Divergencijos atstumas:

$$d_D = \text{Tr} \left\{ \Sigma^{i-1} \Sigma^j + \Sigma^{j-1} \Sigma^i - 2I \right\} + \frac{1}{2} (\mu^i - \mu^j)^T (\Sigma^{i-1} + \Sigma^{j-1}) (\mu^i - \mu^j), \quad (1.93)$$

čia funkcija $\text{Tr}\{A\}$ apskaičiuoja matricos diagonalinių elementų sumą, kur I žymi vienetinę matricą.

Mahalanobio atstumas (angl. *Mahalanobis distance*):

$$d_M = (\mu^i - \mu^j)^T \Sigma^{-1} (\mu^i - \mu^j), \quad (1.94)$$

čia $\Sigma^i = \Sigma^j = \Sigma$.

Metrikos vertina požymių vektorių efektyvumą, atsižvelgia į požymių vidurkį ir kovariaciją. Tačiau norint apskaičiuoti šias metrikas, reikalinga prielaida, kad duomenys yra pasiskirstę pagal normalųjį dėsnį.

Kovariacijos matrica paremti kriterijai. Šie kriterijai taip pat naudojami požymių išrinkimo procese (Webb 2002; Theodoridis *et al.* 1999; Costa *et al.* 1999; Fontaine *et al.* 1997). Kovariacijos matricomis paremti kriterijai yra grindžiami klasių ir tarp-klasių sklaidos matricomis.

Klasės sklaidos matrica (angl. *within-class scatter matrix*) randama:

$$S_W = \sum_{i=1}^C \frac{N_i}{N} S^i, \quad (1.95)$$

čia S^i yra i -osios klasės kovariacinė matrica, apskaičiuojama:

$$S^i = \frac{1}{N_i} \sum_{j=1}^N z_{ij} (x_j - \mu^i)(x_j - \mu^i)^T, \quad (1.96)$$

z_{ij} išreiškiamas:

$$z_{ij} = \begin{cases} 1 & \text{jei } x_j \in X^i, \\ 0 & \text{kitu atveju,} \end{cases} \quad (1.97)$$

čia X^i – i -oji klasė, N_i – i -osios klasės vektorių skaičius, randamas:

$$N_i = \sum_{j=1}^N z_{ij}, \quad (1.98)$$

čia N – visų vektorių skaičius.

Tarp-klasių sklaidos matrica (angl. *between-class scatter matrix*) apskaičiuojama:

$$S_B = \sum_{i=1}^C \frac{N_i}{N} (\mu^i - \mu)(\mu^i - \mu)^T, \quad (1.99)$$

$$\mu = \sum_{i=1}^C \frac{N_i}{N} \mu^i. \quad (1.100)$$

Apžvelgsime keletą kovariacijos matrica paremtų kriterijų. Kriterijus J_1 :

$$J_1 = Tr\{S_W + S_B\} = Tr\{S_M\}, \quad (1.101)$$

čia $S_M = S_W + S_B$.

Kriterijus J_1 charakterizuoja bendrą klasių sklaidą, kuri nepriklauso nuo atskiros klasės informacijos, todėl nėra geras kriterijus požymių išrinkime.

Kovariacijos matricomis paremtų kriterijų tikslas yra surasti požymius, kurių sklaida klasės viduje yra maža, o tarp klasių – didelė. Sklaidą klasės viduje ir tarp klasių vertina kriterijai J_2 , J_3 , J_4 . Kriterijus J_2 apskaičiuojamas:

$$J_2 = \frac{Tr\{S_M\}}{Tr\{S_W\}}. \quad (1.102)$$

Kriterijus J_2 įgyja didelę reikšmę, kai klasei priklausantys vektoriai yra susitelkę į klasterius, o atskirų klasių klasteriai yra nutolę vieni nuo kitų.

Kriterijus J_3 naudojamas kaip alternatyva kriterijui J_2 :

$$J_3 = \frac{Tr |S_M|}{Tr |S_W|} = |S_W^{-1} S_M|. \quad (1.103)$$

Taip pat naudojamas kriterijus J_4 :

$$J_4 = Tr\{S_W^{-1} S_M\}. \quad (1.104)$$

Tam, kad būtų galima naudoti šias metrikas, reikia, jog klasių skirstiniai būtų vienmodaliniai (angl. *unimodal*) ir turėtų skirtingus vidurkius.

Koreliacija (Dowdy *et al.* 2004). Koreliacija rodo statistinį ryšį tarp kintamųjų. Dažniausiai naudojama Pirsono (angl. *Pearson*) koreliacija, kuri nurodo dviejų kintamųjų tiesinę priklausomybę. Pirsono koreliacija išreiškiama:

$$r_{xy} = \frac{\sum_{i=1}^N (x_i - \mu^x)(y_i - \mu^y)}{\sqrt{\sum_{i=1}^N (x_i - \mu^x)^2 \sum_{i=1}^N (y_i - \mu^y)^2}}, \quad (1.105)$$

čia x_i yra i -asis elementas priklausantis klasei X , μ^x – klasės X vidurkis, y_i yra i -asis elementas priklausantis klasei Y , μ^y – klasės Y vidurkis. Siekiama sudaryti požymių rinkinį, kuris įgytų mažą koreliaciją tarp atskirų klasių. Tačiau šis kriterijus nenurodo klasių atskiriamumo lygio.

Skirstinio kreivės formos kriterijai (Joanes *et al.* 1998). Šiai grupei priklauso asimetrijos ir eksceso koeficientai. Asimetrijos koeficientas yra statistinė charakteristika, nurodanti skirstinio dažnumų kreivės formos persišliėjimą.

Asimetrijos koeficientas apskaičiuojamas:

$$A_s = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \right)^3}, \quad (1.106)$$

čia μ yra klasės vidurkis, N – klasės elementų skaičius.

Tuo atveju, kai $A_s = 0$, skirstinys yra pasiskirstęs simetriškai, kai $A_s > 0$ yra skirstinio viršūnė koncentruojasi kairėje pusėje, $A_s < 0$ – viršūnė koncentruojasi dešinėje pusėje.

Ekscesas yra statistinė charakteristika, nurodanti skirstinio dažnumo kreivės viršūnės smailumo laipsnį, lyginant su normaliuoju skirstiniu.

Eksceso koeficientas randamas:

$$K_s = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\left(\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2 \right)^2}. \quad (1.107)$$

Esant $K_s > 0$ – skirstinys turi smailią viršūnę, $K_s < 0$ – skirstinio viršūnė nėra smaili. Skirstinio kreivės formos metrikos nenurodo klasių atskiriamumo lygio.

Vidutinis vektorių skaičius komponentų skaičiaus atžvilgiu. Metrika apskaičiuojama požymių vektorių skaičių padalinus iš vektorių komponentų skaičiaus (Ho et al. 2002):

$$VD = \frac{N}{D}, \quad (1.108)$$

čia N yra visų klasių vektorių skaičius, D – vektorių komponentų skaičius.

Ši metrika nenurodo klasių atskiriamumo lygio, pavyzdžiui – klasės gali būti sunkiai atskiriamos, net jeigu vektorių komponentų skaičius yra mažas, o vektorių skaičius yra didelis.

1.3. Pirmojo skyriaus išvados ir disertacijos uždavinių formulavimas

- Šnekos atpažinimo sistemas sudaro šios pagrindinės dalys: šnekos signalo apdorojimas, kurį atlikus išskiriami požymių vektoriai, bei atpažinimas – vykdomas nežinomo signalo atpažinimas.
- Apžvelgtos požymių sistemos suskirstytos į grupes: laiko analizės, kepstro analizės, tiesinės prognozės analizės, tiesinės suvokimo prognozės analizės. Laiko analizės požymiai yra neefektyvūs dėl neatsparumo triukšmui. Šnekos atpažinimo sistemose dažniausiai naudojami tokie požymiai kaip melų skalės kepstro koeficientai, tiesinės prognozės kepstro koeficientai, tiesinės suvokimo prognozės kepstro koeficientai.
- Kokybiškos požymių sistemos parinkimui naudojamas klasifikavimo klaidos apskaičiavimo metodas. Tačiau šis metodas turi trūkumų:
 1. Su kiekviena tiriama požymių sistema reikia vykdyti klasifikavimą.
 2. Klasifikavimo klaidos apskaičiavimas reikalauja daug darbų atliekant apmokymą, sudarant modelius.
 3. Klasifikavimo klaidos apskaičiavimas reikalauja didelių skaičiavimo resursų.

- Tam, kad pateiktume metodą požymių kokybei vertinti, apžvelgėme metrikas, naudojamas panašioms uždaviniams spręsti: duomenų struktūros tyrimui, dirbtinių duomenų generavimui, požymių išrinkimui. Metrikos skirstomos į tris grupes: geometrinės metrikas, informacijos teorijos metrikas, statistines metrikas.
- Darbo tikslui pasiekti suformuluoti šie uždaviniai:
 1. Pateikti naują šnekos signalų atpažinimo požymių kokybės vertinimo metodą, kuris yra grindžiamas metrikų naudojimu.
 2. Pateikti metrikų rinkinį šnekos signalų atpažinimo požymių kokybei vertinti.
 3. Parodyti, kad pateiktas šnekos signalų požymių vertinimo kokybės metodas Euklido erdvėje suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį, nereikalaujant klasifikavimo eksperimentų vykdymo.
 4. Įvertinti pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo algoritmo sudėtingumą.
 5. Sukūrus programinę įrangą, eksperimentiškai patvirtinti pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą.

Šnekos atpažinimo požymių kokybės vertinimo metodas

Šiame skyriuje pristatomas pateiktas šnekos signalų atpažinimo požymių kokybės vertinimo metodas ir metrikų rinkinys, kuriuo yra grindžiamas sukurtas metodas. Pateikiamas sukurtos požymių kokybės vertinimo programinės įrangos, realizuojančios pateiktą požymių kokybės vertinimo metodą, aprašymas.

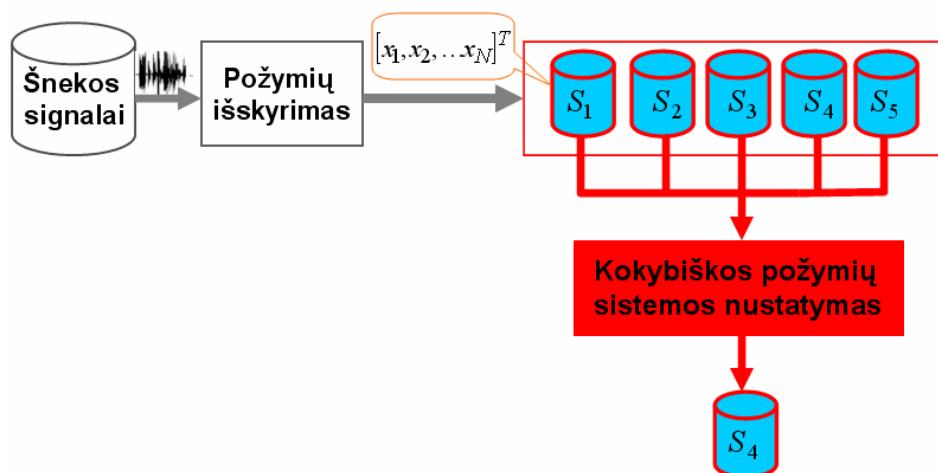
2.1. Metodas šnekos atpažinimo požymių kokybei vertinti

Literatūroje nurodomas metodas požymių kokybei vertinti – klasifikatoriaus naudojimas, kai klasifikavimo klaida apskaičiuojama kiekvienai tiriamai požymių sistemai. Tarkime, yra tiriamos penkios požymių sistemos. Naudojant šį metodą, klasifikavimo eksperimentai turi būti atliekami su visomis penkiomis požymių sistemomis (1.6 pav.). Požymių sistema, kuriai gauta mažiausia klasifikavimo klaida, yra nustatoma kokybiškąją požymių sistema. Kitame etape išrinktoji kokybiška požymių sistema yra naudojama klasifikavimo uždaviniui. Tačiau šis metodas dėl klasifikavimo eksperimentų vykdymo sąlygoja didelius skaičiavimų resursus.

Šiame skyriuje pateiktas naujas metodas šnekos signalų atpažinimo požymių kokybei vertinti, turintis šias savybes:

- Taikant metodą, nereikia atlikti klasifikavimo eksperimentų.
- Metodas tinkamas taikyti Euklido erdvės klasifikatoriams.

Siūlomo metodo schema pateikta 2.1 paveiksle. Tarkime, sprendžiamam uždaviniui duotos penkios požymių sistemos S_1, S_2, S_3, S_4, S_5 . Iškyla klausimas – kurią požymių sistemą naudoti? Užuot vykdant klasifikavimą su kiekviena požymių sistema, naudojant pateiktą metodą, pasitelkiant metrikas, yra išrenkama kokybiška požymių sistema. Kitame etape klasifikavimas vykdomas tik su išrinktąja požymių sistema. Tarkime, šiuo atveju naudojant pateiktą metodą S_4 yra išrenkama kokybiška požymių sistema. Klasifikavimas yra vykdomas tik su požymių sistema S_4 .

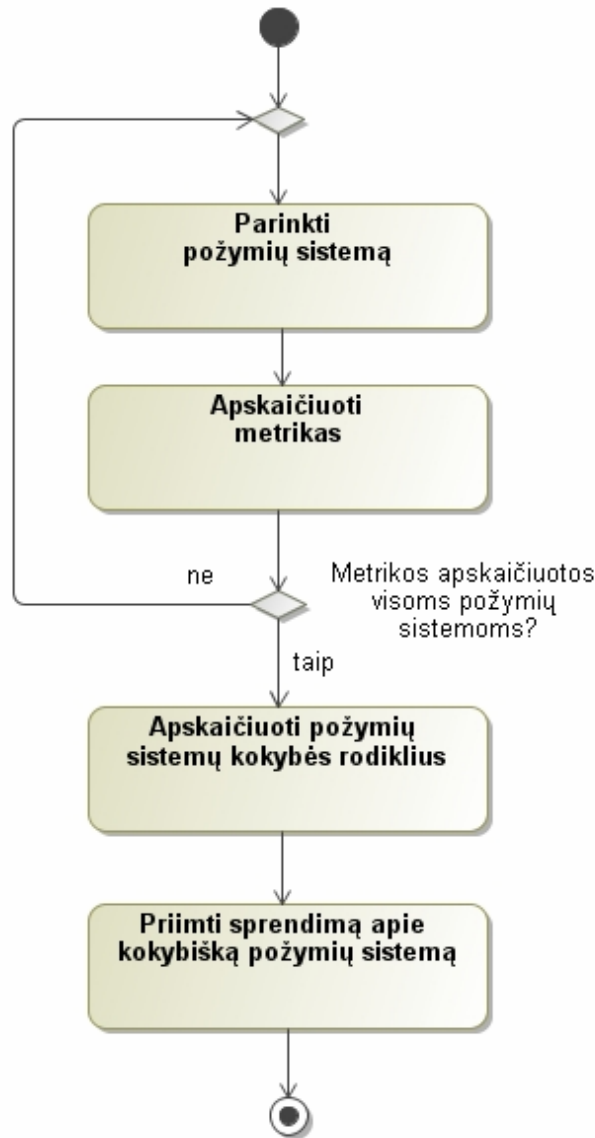


2.1 pav. Požymių kokybės nustatymas naudojant pateiktą metodą

Pateiktas šnekos signalų atpažinimo požymių kokybės nustatymo metodas yra grindžiamas trijų metrikų naudojimu: požymių kiekio klasės ribose metrika, klasių artimiausių kaimynų atstumų santykio metrika, klasės ribos peržengimo kiekio metrika. Šių metrikų aprašymai yra pateikiami 2.2 skyriuje. Pateiktas požymių kokybės nustatymo metodas susideda iš šių pagrindinių etapų (2.2 pav.) (Lileikytė *et al.* 2012):

- Požymių sistemos parinkimas.
- Metrikų apskaičiavimas.
- Požymių sistemų kokybės rodiklių apskaičiavimas.
- Sprendimo priėmimas apie kokybišką požymių sistemą.

Taip pat šiame darbe atlikti eksperimentiniai tyrimai pateikto metodo teisingumui patikrinti – klasifikavimo klaidos vertinimas visoms tiriamoms požymių sistemoms.



2.2 pav. Kokybiškos požymių sistemos nustatymo metodo detali schema

Požymių sistemos parinkimas. Atlikus požymių sistemų išskyrimą, metrikos yra apskaičiuojamos kiekvienai nagrinėjamai požymių sistemai.

Metrikų apskaičiavimas. Metrikos apskaičiuojamos visoms nagrinėjamų klasių pavyzdžių kombinacijoms:

$$V_{nm}^{(p,k)}(i, j) = \sigma^{(p,k)}(h_n^i, h_m^j), \quad (2.1)$$

čia $\sigma^{(p,k)}(h_n^i, h_m^j)$ yra p -osios požymių sistemos atveju apskaičiuota k -oji metrika h_n^i i -osios klasės n -ajam pavyzdžiui ir h_m^j j -osios klasės m -ajam pavyzdžiui, $1 \leq k \leq K$, K – metrikų skaičius, $1 \leq p \leq P$, P – požymių sistemų skaičius, $1 \leq n \leq H_i$, H_i yra i -osios klasės pavyzdžių skaičius, $1 \leq m \leq H_j$, H_j yra j -osios klasės pavyzdžių skaičius, $1 \leq i \leq C$, $1 \leq j \leq C$, $i \neq j$, C yra klasių skaičius.

Nagrinėjamos požymių sistemoms apskaičiuojami metrikų vidurkiai klasių kombinacijų atžvilgiu:

$$VG^{(p,k)}(i, j) = \frac{\sum_{n=1}^{H_i} \sum_{m=1}^{H_j} V_{nm}^{(p,k)}(i, j)}{PBsk}, \quad (2.2)$$

čia $PBsk$ – nagrinėjamų klasių pavyzdžių kombinacijų skaičius. Mažas metrikos vidurkis identifikuoja gerą klasių atskiriamumą.

Požymių sistemų kokybės rodiklių apskaičiavimas. Apskaičiavus metrikų vidurkius atskiroms klasių kombinacijoms, priskiriamas balas tai požymių sistemai, kuriai gautas metrikos vidurkis yra mažiausias:

$$BG^{(p,k)}(i, j) = f(\arg \min_k VG^{(p,k)}(i, j)), \quad (2.3)$$

čia funkcija $f(\cdot)$ gražina 1, jeigu p -ajai požymių sistemai, k -osios metrikos vidurkio reikšmė yra mažiausia. Kitu atveju gražina 0.

Apskaičiuojamas požymių sistemos balų koeficientas nagrinėjamų klasių kombinacijų atveju:

$$FG^P(i, j) = o(\arg \max_p \sum_{k=1}^K BG^{(p,k)}(i, j)), \quad (2.4)$$

čia funkcija $o(\cdot)$ gražina 1, jei daugiausia balų surinko p -oji požymių sistema. Tuo atveju, jei daugiausiai balų surinko keletas požymių sistemų, funkcija gražina $1/Psk$, kur Psk – požymių sistemų skaičius, surinkusių vienodai balų, ir 0 gražina kitu atveju.

Požymių sistemos kokybės rodiklis apskaičiuojamas kaip balų koeficientų vidurkis kiekvienos požymių sistemos atžvilgiu:

$$RFG^p = \frac{\sum_{i=1}^C \sum_{j=1}^C FG^p(i, j)}{KBsk} \cdot 100, \quad (2.5)$$

čia $KBsk$ – klasių kombinacijų skaičius.

Požymių sistemos kokybės rodiklis nurodo požymių sistemos kokybę, kur 0 % – identifikuoja žemiausią požymių sistemos kokybę, 100 % – aukščiausią požymių sistemos kokybę.

Sprendimo priėmimas apie kokybišką požymių sistemą. Kokybiška požymių sistema išrenkama sistema, turinti didžiausią požymių sistemos kokybės rodiklį:

$$SFG = \arg \max_p RFG^p. \quad (2.6)$$

Pateikto metodo teisingumo patikrinimas. Darbe atliktas pateikto metodo teisingumo patikrinimas – ar kokybiškos požymių sistemos nustatymo rezultatai, gauti naudojant pateiktą metodą, sutampa su rezultatais, gautais naudojant klasifikavimo klaidą. Klasifikavimo klaidos apskaičiavimas susideda iš dviejų etapų: pirmajame etape apskaičiuojama klasifikatoriaus klaida kiekvienai tiriamai požymių sistemai, visoms klasių kombinacijoms ir žymėsime $EP^p(i, j)$; antrajame etape apskaičiuojamas klasifikavimo klaidos vidurkis, klasių kombinacijų atžvilgiu:

$$EK^p = \frac{\sum_{i=1}^C \sum_{j=1}^C EP^p(i, j)}{KBsk} \cdot 100. \quad (2.7)$$

Identifikuojama aukščiausią kokybę turinti požymių sistema, kurios klasifikavimo klaida yra mažiausia:

$$PEK = \arg \min_p EK^p. \quad (2.8)$$

Pateikto metodo teisingumas yra patvirtinamas, jeigu naudojant pateiktą metodą identifikuota kokybiška požymių sistema sutampa su kokybiška požymių sistema, nustatyta naudojant klasifikavimo klaidą.

Pateiktas metodas yra grindžiamas trijų metrikų naudojimu. Metrikos požymių kiekio klasės ribose algoritmo apskaičiavimo sudėtingumas yra $O(2R \log 2R)$, metrikos klasių artimiausių kaimynų atstumų santykis yra $O(2R \log 2R)$ ir metrikos klasės ribos peržengimo kiekis yra $O(2R)$, kur R yra šnekos signalo objekto vektorių skaičius ($N = 2R$, kur N – klasių vektorių skaičius). Metrikų detalūs aprašymai suformuluoti 2.2 skyriuje. Pateikto šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, kai dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$.

2.2. Metrikos šnekos atpažinimo požymių kokybei vertinti

Pateiktas metodas požymių kokybei vertinti yra grindžiamas metrikų naudojimu. Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti sudarytas iš 3 metrikų (2.1 lentelė) (Lileikytė *et al.* 2011; Lileikytė *et al.* 2012).

Metrikos atitinka metodui iškeltus reikalavimus: nereikalauja klasifikavimo eksperimentų vykdymo, tinkamas Euklido erdvės klasifikatoriams. Metrikų rinkinys priklauso geometrinių metrikų grupei (1.2.2 skyrius), yra glaudžiai susijęs su klasifikavimo sudėtingumo priežastimis – klasių persidengimu, klasių ribų sudėtingumo identifikavimu.

2.1 lentelė. Metrikos šnekos signalų atpažinimo požymių kokybei vertinti

Metrika	Žymėjimas
Požymių kiekis klasės ribose	G1
Klasių artimiausių kaimynų atstumų santykis	G2
Klasės ribos peržengimo kiekis	G3

Kitos geometrinės grupės metrikos netenkina iškeltų metrikų savybių bei turi trūkumų, dėl kurių yra netinkamos požymių kokybei vertinti. Išskiriamos metrikos, kurios remiasi klasifikatoriaus vykdymu: tiesinio klasifikatoriaus klaidos, artimiausio kaimyno klasifikatoriaus klaidos, tiesinio klasifikatoriaus netiesiškumo, artimiausio kaimyno klasifikatoriaus netiesiškumo metrikos. Maksimalaus Fišerio diskriminantinio santykio metrika taikoma vienmodaliniams skirstiniams, pasiskirsčiusiems pagal normalųjį dėsnį. Metrika, nurodanti požymių persidengimo sritį, netinkama naudoti šnekos signalų atpažinimo požymiams, kadangi ji nepritaikyta operuoti su neigiamomis požymių vektorių reikšmėmis. Požymių efektyvumo metrika, nustatydamą klasių atskiriamumą, atsižvelgia tik į plokštumą, statmeną požymių ašims. Štai ϵ -kaimynystės klasterių skaičiaus metrika tinkamesnė vidinei klasių struktūrai nusakyti, nei klasių atskiriamumui vertinti.

2.2.1. Požymių kiekio klasės ribose metrika

Požymių kiekio klasės ribose metrika apskaičiuojama randant Minimalaus jungiančiojo medžio (MJM) briaunomis sujungtas skirtingų klasių viršūnes (Ho *et al.* 2002).

Požymių kiekio klasės ribose metrikos apskaičiavimas susideda iš šių etapų:

- Klasėms apskaičiuojamas MJM. Randamos briaunomis sujungtos skirtingų klasių viršūnės. Mūsų atveju viršūnės yra klasių vektoriai, briaunos – Euklido atstumai tarp vektorių.
- Šių viršūnių skaičius padalinamas iš visų viršūnių skaičiaus.

MJM yra jungiantysis medis, kurio briaunų svorių suma yra mažiausia tarp visų galimų jungiančiųjų medžių. Tegul $G = (V, E)$ yra jungusis, neorientuotas grafas, čia $V = \{v_1, v_2, \dots, v_N\}$ yra viršūnės, kur N – viršūnių skaičius, $E = \{e_1, e_2, \dots, e_M\}$ yra briaunos, kur M yra briaunų skaičius. Tegul $W = \{w_1, w_2, \dots, w_M\}$ yra briaunoms priskirti svoriai (skaitinės reikšmės)

(Bock 1971; Cheriton *et al.* 1976; Graham *et al.* 1985; Xiao-hua *et al.* 2009; Guo *et al.* 2009).

Kiekvienas grafo G pografis gali būti išreiškiamas naudojant vektorių $P = \{p_1, p_2, \dots, p_M\}$, kur p_i apibrėžiamas (Ai-bing *et al.* 2007):

$$p_i = \begin{cases} 1, & \text{jei briauna } e_i \text{ priklauso pografui,} \\ 0, & \text{kitu atveju.} \end{cases} \quad (2.9)$$

Tegul G_p yra grafo G pografis. G_p yra laikomas grafo G jungiančiuoju medžiu, jeigu:

1. Pografis G_p turi visas grafo G viršūnes.
2. Pografis yra jungusis ir neturi ciklą.

Tegul GM yra visų jungiančiųjų medžių aibė grafe G . MJM yra išreiškiamas:

$$z(P) = \arg \min \sum_{i=1}^M w_i p_i, \quad P \in GM. \quad (2.10)$$

MJM pasižymi šiomis savybėmis:

1. MJM neturi ciklą, jis apjungia visas viršūnes, su $N - 1$ briaunų skaičiumi.
2. Tarp bet kurių dviejų viršūnių yra tik vienas kelias.
3. Pridėjus briauną, nepriklausančią medžiui, yra sukuriamas ciklas.

Tuo tarpu pašalinus šią briauną – gaunamas MJM.

MJM rasti dažniausiai yra naudojami Kruskalo (angl. *Kruskal*) ir Primo (angl. *Prim*) godūs algoritmai. Kruskalo algoritmas randa mažiausią svorį turinčią grafo briauną ir įtraukia į medį, jei briauna nesudaro ciklo su anksčiau įtrauktomis medžio briaunomis (Kruskal 1956; Sudhakar *et al.* 2011; Alpert *et al.* 1993). Kruskalo algoritmo sudėtingumas yra $O(M \log N)$. Tegul U yra aibės V poaibis ir $(u, v) \in E$. Kruskalo algoritmas yra pateiktas 2.3 paveiksle:

1. E_{KR} yra tuščia briaunų aibė.
2. Briaunos $(u, v) \in E$ yra surūšiuojamos svorių didėjimo tvarka.
3. Vykdoma kol E_{KR} turi mažiau nei $N - 1$ briauną:
4. jei $E_{KR} \cup (u, v)$ nesudaro ciklą
5. $E_{KR} \leftarrow E_{KR} \cup (u, v)$.

2.3 pav. Kruskalo algoritmas

Tuo tarpu Prim'o algoritmas randa mažiausią svorį turinčią grafo briauną, kurios viena gretima viršūnė jau yra įtraukta į medį, o kita viršūnė – nėra (Prim 1957; Aissa *et al.* 2009; Vandervalk *et al.* 2009; Leung *et al.* 1991; Huang *et al.* 2009). Prim'o algoritmo sudėtingumas yra $O(M + N \log N)$. Prim'o algoritmas yra pateiktas 2.4 paveiksle:

1. $V_{PR} = \{v\}$ yra viršūnių aibė su atsitiktinai pasirinkta viršūne iš viršūnių aibės V , $E_{PR} = \{\}$ yra tuščia briaunų aibė.
2. Briaunos $(u, v) \in E$ yra surūšiuojamos svorių didėjimo tvarka.
3. Vykdoma kol $V_{PR} = V$:
4. pasirenkama briauna (u, v) turinti mažiausią svorį, kai $u \in V_{PR}$ ir $v \notin V_{PR}$.
5. $V_{PR} \leftarrow V_{PR} \cup \{v\}$.
6. $E_{PR} \leftarrow E_{PR} \cup (u, v)$.

2.4 pav. Prim'o algoritmas

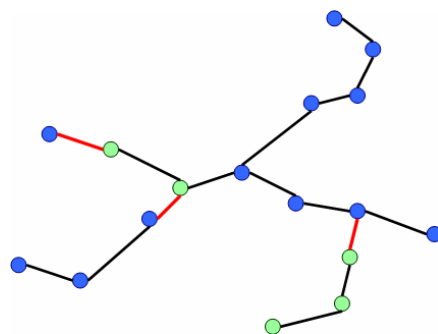
Kai briaunos yra Euklido atstumai tarp vektorių, gali būti naudojamas Euklido minimalus jungiantysis medis (EMJM), kurio algoritmo sudėtingumas yra $O(N \log N)$ (Buchin *et al.* 2009; Narasimhan *et al.* 2000; Yao 1989; Agarwal *et al.* 1991; Jaromczyk *et al.* 1992; Callahan *et al.* 1993).

Apskaičiuojant metriką, pirmiausia randamas MJM. Tegul L yra viršūnių aibė, $L \in V$, $L = l_1, l_2, \dots, l_K$, kurios viršūnės MJM sujungtos briaunomis ir priklauso priešingoms klasėms (2.5 pav.), K – šių viršūnių skaičius. Požymių kiekio klasės ribose metrika apskaičiuojama šių viršūnių skaičių padalinus iš visų viršūnių skaičiaus:

$$G1 = \frac{K}{N}, \quad (2.11)$$

čia N yra visų viršūnių skaičius.

Metrika yra jautri vektoriams, esantiems prie klasių ribų. Minimali metrikos reikšmė yra teigiama reikšmė arti nulio ir rodo, kad klasės yra gerai atskiriamos. Minimali metrikos reikšmė neįgyja nulio, nes klasėms suradus MJM – nors viena briauna jungia priešingų klasių vektorius. Maksimali metrikos reikšmė yra vienetas ir rodo, kad klasės yra blogai atskiriamos.



2.5 pav. Vektoriai, MJM sujungti briaunomis ir priklausantys priešingoms klasėms

2.2.2. Klasių artimiausių kaimynų atstumų santykio metrika

Klasių artimiausių kaimynų atstumų santykio metrika yra randama atsižvelgiant į arčiausiai esančius vektorius, priklausančius tai pačiai klasei, ir arčiausiai esančius vektorius, priklausančius kitai klasei (Bernado–Mansilla *et al.* 2005).

Klasių artimiausių kaimynų atstumų santykio metrika yra apskaičiuojama:

- Klasių vektoriams randami Euklido atstumai iki artimiausio vektoriaus priklausančio tai pačiai klasei ir iki artimiausio vektoriaus priklausančio kitai klasei (2.6 pav.).

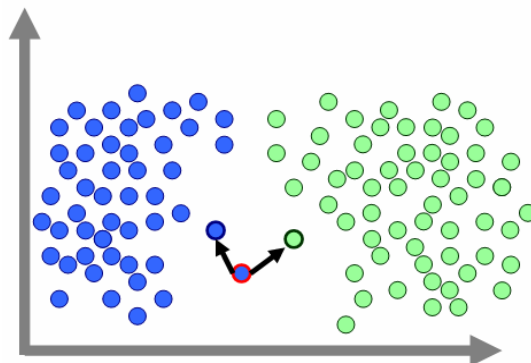
- Randamas šių atstumų santykis.

Metrika yra išreiškiama formule:

$$G2 = \frac{\sum_{i=1}^C \sum_{n=1}^{N_i} \min_k d(x_n^i, x_k^i)}{\sum_{i=1, j=1, i \neq j}^C \sum_{n=1}^{N_i} \sum_{m=1}^{N_j} \min d(x_n^i, x_m^j)}, \quad (2.12)$$

čia $\min_k d(x_n^i, x_k^i)$ yra Euklido atstumas tarp i -osios klasės n -ojo vektoriaus ir arčiausiai esančio bei i -ajai klasei priklausančio k -ojo vektoriaus, $1 \leq n \leq N_i$, $1 \leq k \leq N_i$, N_i yra i -osios klasės vektorių skaičius, $1 \leq i \leq C$, $1 \leq j \leq C$, $i \neq j$, C – klasių skaičius, $\min_m d(x_n^i, x_m^j)$ tai – Euklido atstumas tarp i -osios klasės n -ojo vektoriaus ir arčiausiai esančio j -ajai klasei priklausančio m -ojo vektoriaus, $1 \leq m \leq N_j$, N_j yra j -osios klasės vektorių skaičius.

Metrika yra jautri vektorių atstumui klasės viduje ir tarp klasių. Metrikos algoritmo apskaičiavimo sudėtingumas yra $O(N \log N)$, kur N – visų vektorių skaičius. Minimali metrikos reikšmė yra nulis ir rodo, kad klasės yra gerai atskiriamos. Vienetas yra didžiausia metrikos reikšmė ir gaunama tuo atveju, jeigu atstumas klasės viduje yra mažesnis nei atstumas tarp klasių. Atstumui klasės viduje esant didesniam nei atstumui tarp klasių, metrikos reikšmė yra didesnė už vienetą ir tuo atveju maksimali metrikos reikšmė nėra apibrėžta.



2.6 pav. Atstumai iki tos pačios klasės ir priešingos klasės artimiausių vektorių

2.2.3. Klasės ribos peržengimo kiekio metrika

Klasės ribos peržengimo kiekis fiksuojamas, kai klasės vektorius patenka į kitos klasės ribas (Li *et al.* 1998). Tegul kiekviena klasė yra apskritimas. Jos centras yra klasės vektorių vidurkis, spindulys – atstumas nuo klasės centro iki tolimiausio tos klasės vektoriaus.

Klasės ribos peržengimo kiekio metrika yra apskaičiuojama:

- Randamas vektorių, patenkančių į kitos klasės apskritimo ribas, skaičius (2.7 pav.).
- Randama $(C - 1)$ klasių ir visų vektorių sandauga.
- Rastų vektorių skaičius padalinamas iš gautos sandaugos.

Analizuojamos i -osios klasės spindulys apibrėžiamas:

$$r^i = \max_n d(\mu^i, x_n^i), \quad (2.13)$$

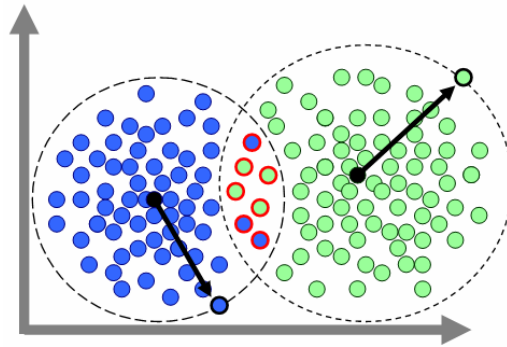
čia $\max_n d(\mu^i, x_n^i)$ yra maksimalus Euklido atstumas tarp μ^i – i -osios klasės centro ir x_n^i – i -osios klasės n -ojo požymių vektoriaus, $1 \leq i \leq C$, C yra klasių skaičius, $1 \leq n \leq N_i$, N_i – i -osios klasės vektorių skaičius.

Klasės ribos peržengimo kiekio metrika išreiškiama formule:

$$G3 = F[d(\mu^i, x_m^j) \leq r^i] \cdot \frac{1}{C-1} \cdot \frac{1}{N}, \quad (2.14)$$

čia $d(\mu^i, x_m^j)$ yra Euklido atstumas tarp i -osios klasės centro ir j -osios klasės m -ojo požymių vektoriaus, $1 \leq m \leq N_j$, N_j – j -osios klasės vektorių skaičius, $i \neq j$, N – visų klasių vektorių skaičius. Funkcija $F[\cdot]$ gražina atstumų skaičių, tenkinančių klasės ribos peržengimo sąlygą $d(\mu^i, x_m^j) \leq r^i$.

Minimali metrikos reikšmė yra nulis ir rodo, kad klasės yra gerai atskiriamos. Tuo tarpu maksimali metrikos reikšmė yra vienetasis ir rodo, kad klasės yra blogai atskiriamos. Metrikos algoritmo apskaičiavimo sudėtingumas yra $O(N)$.



2.7 pav. Klasių vektoriai patenkantys į kitų klasių apskritimų ribas

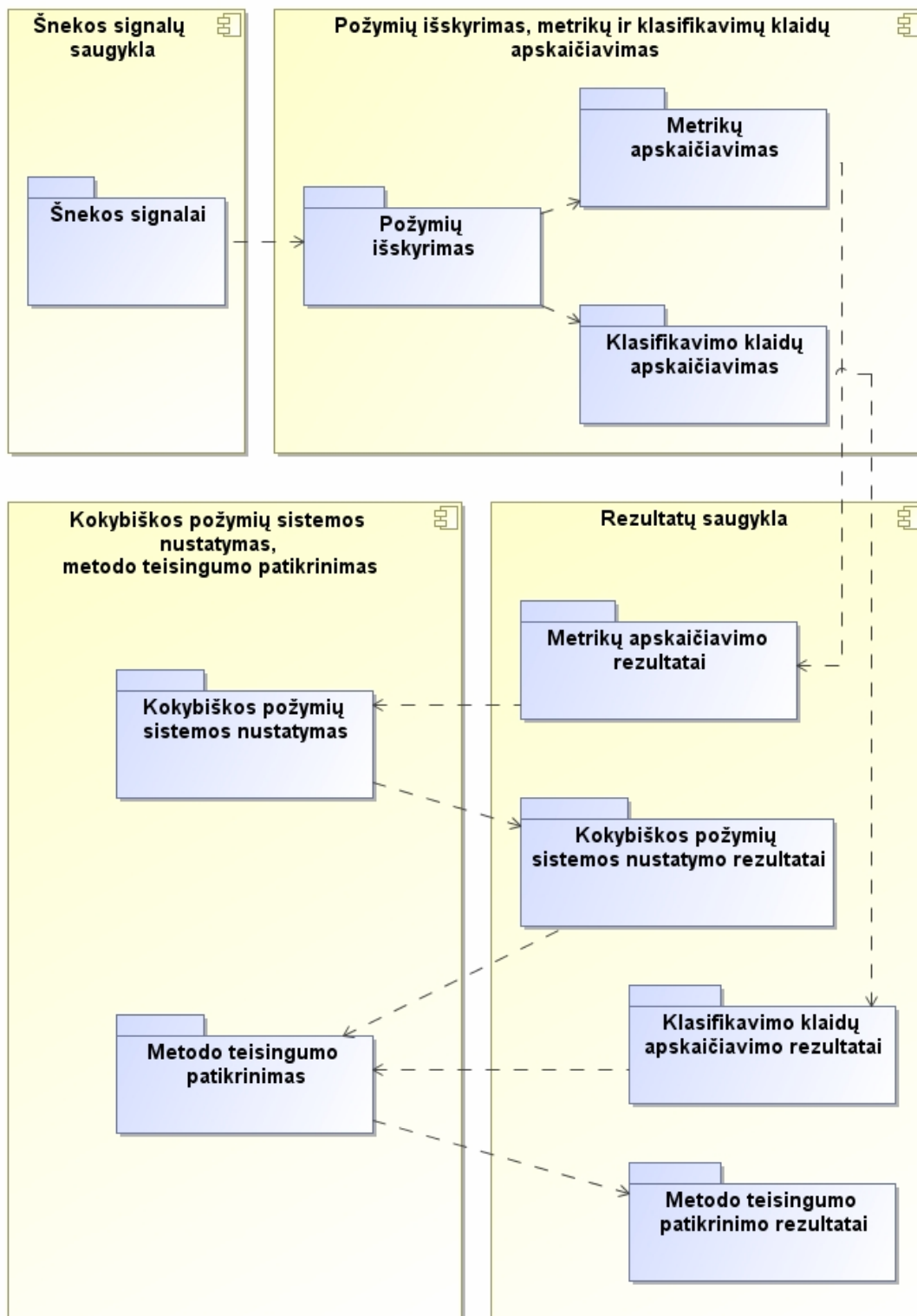
2.3. Programinė įranga požymių kokybei vertinti

Darbo metu sukurta programinė įranga šnekos signalų atpažinimo požymių kokybės vertinimo eksperimentiniams tyrimams atlikti. Šiame skyriuje detaliau apžvelgiami programinę įrangą sudarantys komponentai ir jų funkcijos (2.8 pav.).

Sukurta programinė įranga sudaryta iš keturių komponentų:

- Šnekos signalų saugykla. Šnekos signalų saugykloje yra saugomi šnekos signalų įrašai.
- Požymių išskyrimas, metrikų bei klasifikavimo klaidų apskaičiavimas. Duomenų saugykloje esantys šnekos signalų įrašai yra apdorojami, išskiriami požymių vektoriai. Tada apskaičiuojamos metrikos ir klasifikavimo klaidos. Tai realizuota naudojant Matlab R2007b programinį paketą.
- Rezultatų saugykla. Metrikų ir klasifikavimo klaidų apskaičiavimo rezultatai patalpinami rezultatų saugykloje.
- Naudojant rezultatų saugykloje esančius metrikų rezultatus, nustatoma aukščiausią kokybę turinti požymių sistema. Taip pat aukščiausią kokybę turinti požymių sistema yra nustatoma naudojant rezultatų saugykloje esančius klasifikavimo klaidų rezultatus. Tikrinama, ar kokybiškos požymių sistemos nustatymo rezultatai, gauti naudojant pateiktą metodą, sutampa su rezultatais, gautais naudojant klasifikavimo klaidą. Gauti rezultatai patalpinami rezultatų saugykloje. Šis

funktionalumas realizuotas naudojant PL/SQL programavimo kalbą (Oracle SQL Developer 2.1 programinės įrangos paketas) ir C++ programavimo kalbą (MinGW 0.2 programinės įrangos paketas).



2.8 pav. Šnekos atpažinimo požymių kokybės vertinimo programinės įrangos komponentų struktūra

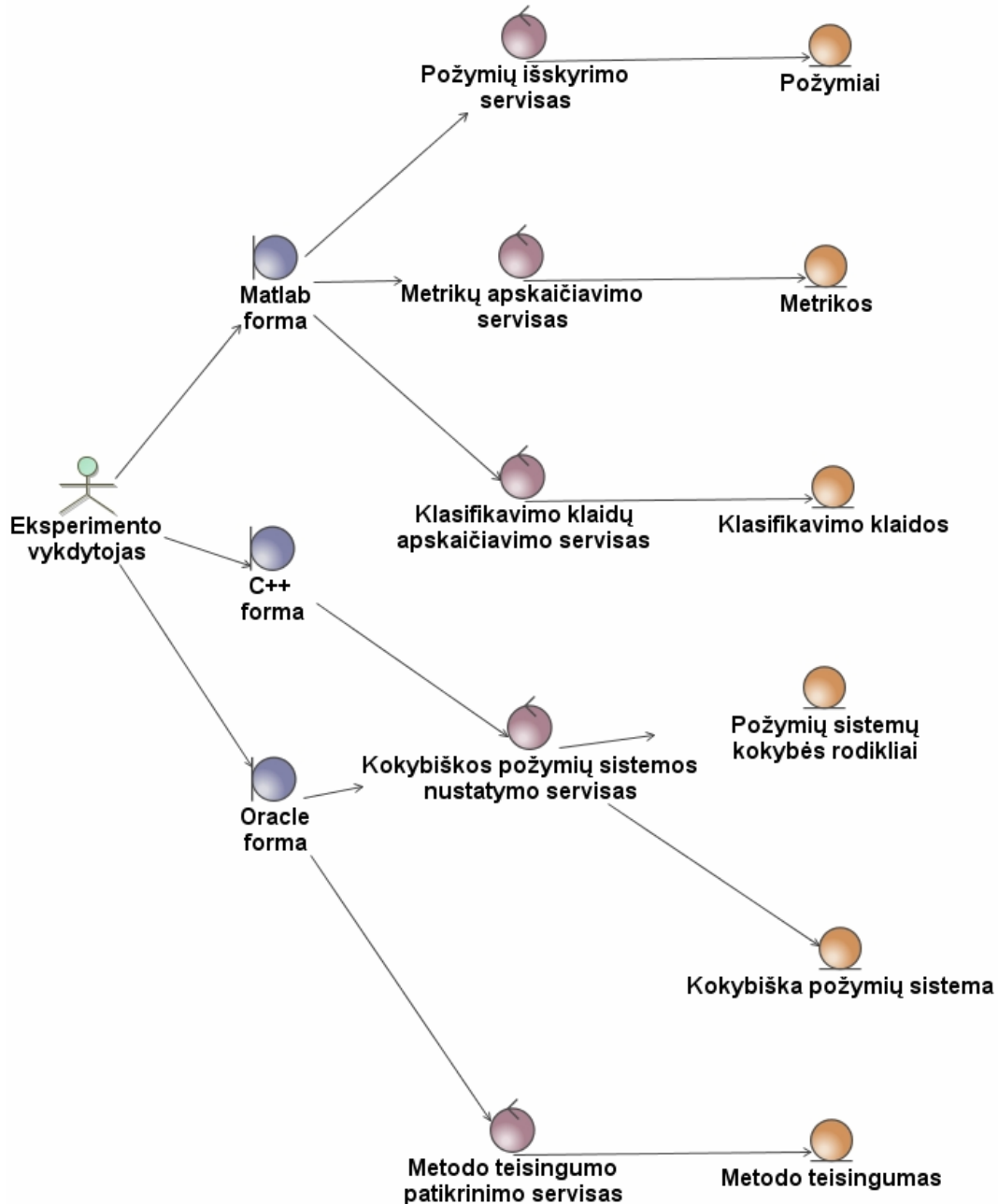
Šnekos signalų apdorojimas, požymių išskyrimas. Šnekos signalo apdorojimo metu šnekos signalas yra dalijamas į 23,22 ms kadrus, kadrai persidengia 11,61 ms. Naudojama Haningo lango funkcija (1.3 formulė). Sistema apskaičiuoja 12 eilės TSPKK požymių sistemą, naudojant pateiktas (1.34)–(1.43), (1.50)–(1.52) formules, bei 12 eilės KK požymių sistemą, naudojant (1.9)–(1.10) formules. Taip pat sistema gali apskaičiuoti pasirinktos eilės TSPKK ir KK požymių sistemas. Be to, galima naudoti norimas požymių sistemas, jas pateikus .mat failų formatu.

Metrikų apskaičiavimas. Metrikas realizuojančios funkcijos apskaičiuoja tris metrikas. Požymių kiekio klasės ribose metrikos (G1) apskaičiavimui naudojama (2.11) išraiška ir algoritmas pateiktas 2.4 paveiksle. Klasių artimiausių kaimynų atstumų santykio metrikai (G2) apskaičiuoti naudojama (2.12) formulė. Klasės ribos peržengimo kiekio metrikai (G3) apskaičiuoti naudojamos (2.13)–(2.14) išraiškos.

Kokybiškos požymių sistemos nustatymas. Kokybiškai požymių sistemai nustatyti naudojamos (2.1)–(2.6) formulės.

Pateikto metodo teisingumo patikrinimas. Pateikto metodo teisingumui patikrinti naudojamos (2.7)–(2.8) išraiškos. AK klasifikavimo klaidos apskaičiavimas atliekamas pagal (1.63)–(1.64) formules. DLSK klasifikavimo klaida apskaičiuojama pagal (1.65)–(1.72) formules, naudojamas lokalus Itakuros krypties apribojimas – pateiktas 1.4 paveiksle, bei globalus Itakuros lygiagretinio krypties apribojimas – pavaizduotas 1.5 paveiksle. DLSK klasifikavimo klaida yra apskaičiuojama nežinomam pavyzdžiui randant apmokymo pavyzdį su mažiausiu dinaminio laiko skalės kraipymo atstumu ir fiksuojant klaidą, jeigu nežinomo pavyzdžio ir apmokymo pavyzdžio klasės nesutampa.

Taip pat 2.9 paveiksle vaizduojama šnekos signalų atpažinimo požymių kokybės vertinimo programinės įrangos servisų struktūra. Šioje servisų struktūroje pateikiami programinės įrangos ryšiai tarp sistemos vartotojo, vartotojo sąsajos (kuri naudojama vartotojui sąveikauti su sistema), sistemos tiekiamų servisų ir duomenų.



2.9 pav. Šnekos atpažinimo požymių kokybės vertinimo programinės įrangos servisų struktūra

2.4. Skyriaus išvados

- Pateiktas metodas šnekos signalų atpažinimo požymių kokybei vertinti yra grindžiamas metrikų rinkinio naudojimu.
- Pateiktą metrikų rinkinį sudaro trys metrikos: požymių kiekio klasės ribose (G1), klasių artimiausių kaimynų atstumų santykio (G2) ir klasės ribos peržengimo kiekio (G3) metrika.

- Pateiktas šnekos signalų požymių kokybę aprašantis metodas Euklido erdvėje suteikia galimybę supaprastinti šnekos signalų atpažinimo uždavinį – nereikalauja klasifikavimo eksperimentų vykdymo.
- Įvesta požymių kokybės rodiklio skalė 0 %–100 %, kur 0 % rodo žemiausią požymių sistemos kokybę, o 100 % – aukščiausią požymių sistemos kokybę.
- Pateikto šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, kai dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, kur R – šnekos signalo objekto vektorių skaičius.
- Sukurta programinė įranga šnekos signalų atpažinimo požymių kokybės vertinimo eksperimentiniams tyrimams atlikti.

Šnekos atpažinimo požymių kokybės vertinimo eksperimentiniai tyrimai

Šiame skyriuje pateikti eksperimentinių tyrimų rezultatai, kurių tikslas yra eksperimentiškai patvirtinti pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą. Skyrius sudarytas iš šių dalių: eksperimentinių tyrimų sąlygų ir duomenų aprašymo, šnekos signalų atpažinimo požymių kokybės vertinimo eksperimentinių tyrimų rezultatų pateikimo.

3.1. Eksperimentinių tyrimų duomenys ir sąlygos

Eksperimentiniai tyrimai atlikti naudojant sukurtą šnekos signalų atpažinimo požymių kokybės vertinimo programinę įrangą. Tiriamos atpažinimo sistemos:

- Atpažinimo sistema, naudojanti 12 eilės TSPKK požymių sistemą ir AK klasifikatorių.
- Atpažinimo sistema, naudojanti 12 eilės KK požymių sistemą ir AK klasifikatorių.
- Atpažinimo sistema, naudojanti 12 eilės TSPKK požymių sistemą ir DLSK klasifikatorių.

- Atpažinimo sistema, naudojanti 12 eilės KK požymių sistemą ir DLSK klasifikatorių.

Požymių sistemų kokybė nustatoma naudojant pateiktą metodą, kuris yra grindžiamas trijų metrikų naudojimu: požymių kiekio klasės ribose (G1), klasių artimiausių kaimynų atstumų santykio (G2) ir klasės ribos peržengimo kiekio (G3) metrika.

Eksperimentiniuose tyrimuose naudojama 14 skirtingų fonemų, reprezentuojančių skirtingas klases. Naudojama 14 dažniausių lietuvių kalbos nekontekstinių fonemų (Raškinis *et al.* 2009): [i], [a], [e], [o:], [s], [k], [t], [t'], [m], [n'], [r], [r'], [j], [s'] (čia „ ’ “ žymi minkštą priebalsį, „ : “ – ilgą balsį). Eksperimentiniuose tyrimuose naudojamas VDU paruoštas VDU-TRI4 garsynas (Raškinis *et al.* 2004). Garsyną sudaro rišlūs lietuvių kalbos sakiniai, garsyno trukmė yra 107,1 min. Garsynas įrašytas 4 kalbėtojų – dviejų moterų ir dviejų vyrų. Kalbėtojų amžiaus vidurkis yra 37 metai. Vienas kalbėtojas kilęs iš Kauno, du kalbėtojai – vakarų aukštaičiai, vienas – rytų aukštaitis iš Utenos. Įrašai daryti tylioje aplinkoje (tačiau ne profesionalioje įrašų studijoje). Iš ištartų frazių iškirptos nekontekstinės fonemos naudojant nemokamai platinamą programinę įrangą PRAAT (Boersma *et al.* 2000) ir VDU paruoštą programinę įrangą Tescal. Naudojamų įrašų diskretizavimo dažnis 11 025 Hz, 16 bitų mono-formatas.

Eksperimentiniai tyrimai atlikti naudojant tris duomenų rinkinius:

- DM1 – vieno kalbėtojo duomenų rinkinys, kalbėtojas – vyras.
- DM2 – vieno kalbėtojo duomenų rinkinys, kalbėtojas – moteris.
- DM3 – keturių kalbėtojų duomenų rinkinys, kalbėtojai – 2 vyrai ir dvi moterys.

Eksperimentinių tyrimų metrikų apskaičiavimų apimtys vienam duomenų rinkiniui, kiekvienai nagrinėjamai požymių sistemai:

- Klasių skaičius – 14.
- Klasių kombinacijų skaičius – 91, t.y. visos galimos klasių porų kombinacijos tarp nagrinėjamų 14 klasių. Klasių kombinacijos pateiktos 3.1 lentelėje.

- Klasę sudarančių pavyzdžių skaičius – 100.
- Pavyzdžių kombinacijų skaičius vienai klasių porai – 10 000.
- Pavyzdžių kombinacijų skaičius visoms klasių poroms – 910 000.
- Metrikų skaičius – 3. Naudojamas metrikų rinkinys: požymių kiekis klasės ribose (G1), klasių artimiausių kaimynų atstumų santykis (G2) ir klasės ribos peržengimo kiekis (G3).
- Metrikų apskaičiavimų skaičius visoms pavyzdžių kombinacijoms yra 2 730 000.

3.1 lentelė. Eksperimentiniuose tyrimuose naudojamos klasių kombinacijos

Klasės	i	a	e	o:	s	k	t	t'	m	n'	r	r'	j	s'
i		+	+	+	+	+	+	+	+	+	+	+	+	+
a			+	+	+	+	+	+	+	+	+	+	+	+
e				+	+	+	+	+	+	+	+	+	+	+
o:					+	+	+	+	+	+	+	+	+	+
s						+	+	+	+	+	+	+	+	+
k							+	+	+	+	+	+	+	+
t								+	+	+	+	+	+	+
t'									+	+	+	+	+	+
m										+	+	+	+	+
n'											+	+	+	+
r												+	+	+
r'													+	+
j														+
s'														

Apibendrintos eksperimentinių tyrimų metrikų apskaičiavimų apimtys kiekvienai nagrinėjamai požymių sistemai, vienam duomenų rinkiniui pateiktos 3.2 lentelėje.

Pastebime, kad eksperimentiniuose tyrimuose vienam duomenų rinkiniui buvo atlikta 5 460 000 metrikų apskaičiavimų. Darbe naudojami 3 duomenų rinkiniai, taigi šiems trims duomenų rinkiniams buvo atlikta 16 380 000 metrikų apskaičiavimų.

Vykdam eksperimentinius tyrimus metodo teisingumui patikrinti, nagrinėjamiems trims duomenų rinkiniams (91 klasių kombinacijai) taip pat apskaičiuotos AK ir DLSK klasifikatorių klasifikavimo klaidos 12 eilės

TSPKK, 12 eilės KK požymių sistemų atvejais. Darbe 70 % duomenų naudojama apmokymui, 30 % – testavimui. Pateikto metodo teisingumas yra patvirtinamas, kai naudojant pateiktą metodą nustatyta kokybiška požymių sistema sutampa su požymių sistema, kokybės nustatymui naudojant klasifikavimo klaidą.

3.2 lentelė. Eksperimentinių tyrimų metrikų apskaičiavimų apimtys vienam duomenų rinkiniui

Vertinimo aspektas	Požymių sistema	Eksperimentų skaičius	Bendras eksperimentų skaičius
Klasių skaičius	12 TSPKK	14	14
	12 KK	14	
Klasių kombinacijų skaičius	12 TSPKK	91	182
	12 KK	91	
Klasę sudarančių pavyzdžių skaičius	12 TSPKK	100	200
	12 KK	100	
Pavyzdžių kombinacijų skaičius vienai klasių porai	12 TSPKK	10 000	20 000
	12 KK	10 000	
Pavyzdžių kombinacijų skaičius visoms klasių poroms	12 TSPKK	910 000	1 820 000
	12 KK	910 000	
Metrikų skaičius	12 TSPKK	3	3
	12 KK	3	
Metrikų apskaičiavimų skaičius visoms pavyzdžių kombinacijoms	12 TSPKK	2 730 000	5 460 000
	12 KK	2 730 000	

3.2. Požymių kokybės vertinimo eksperimentiniai tyrimai

Šiame skyriuje pristatomi šnekos signalų atpažinimo požymių kokybės vertinimo eksperimentinių tyrimų rezultatai naudojant pateiktą metodą. Taip pat pristatomi eksperimentinių tyrimų rezultatai pateikto metodo teisingumui patvirtinti – požymių kokybė vertinama naudojant klasifikavimo klaidą.

Nustatant kokybišką požymių sistemą naudojant pateiktą metodą, pristatyti šie tarpiniai eksperimentinių tyrimų rezultatai:

- Metrikų vidurkiai nagrinėjamų klasių kombinacijų atžvilgiu.
- Požymių sistemų kokybės rodikliai.

Apskaičiavus metrikų vidurkius atskiroms klasių kombinacijoms (maža metrikos reikšmė nurodo, kad klasės yra gerai atskiriamos, didelė metrikos reikšmė nurodo, kad klasės – blogai atskiriamos), apskaičiuojami požymių sistemų kokybės rodikliai. Požymių sistemos kokybės rodiklis nurodo metrikų vertinimo apibendrintus rezultatus. Požymių sistemos kokybės rodiklio skalė yra 0 %–100 %, kur:

- 0 % identifikuoja žemiausią požymių kokybę.
- 100 % – aukščiausią požymių kokybę.

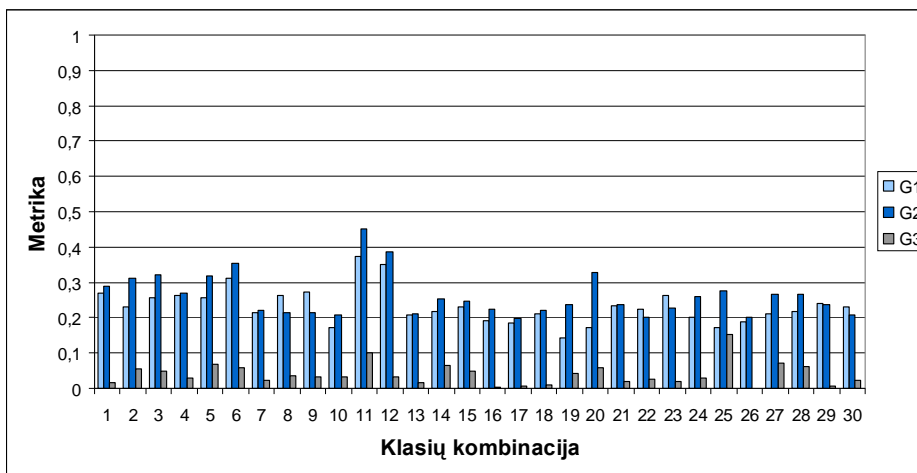
Požymių sistema, kuri turi aukščiausią kokybės rodiklį, yra nustatoma aukščiausią kokybę turinčia požymių sistema.

Metodo teisingumui patikrinti apskaičiuojamos AK ir DLSK klasifikavimo klaidos (su 80 % pasikliautinumo intervalais). Patikrinama, ar kokybiškos požymių sistemos nustatymo rezultatai, gauti naudojant pateiktą metodą, sutampa su rezultatais, gautais naudojant klasifikavimo klaidą.

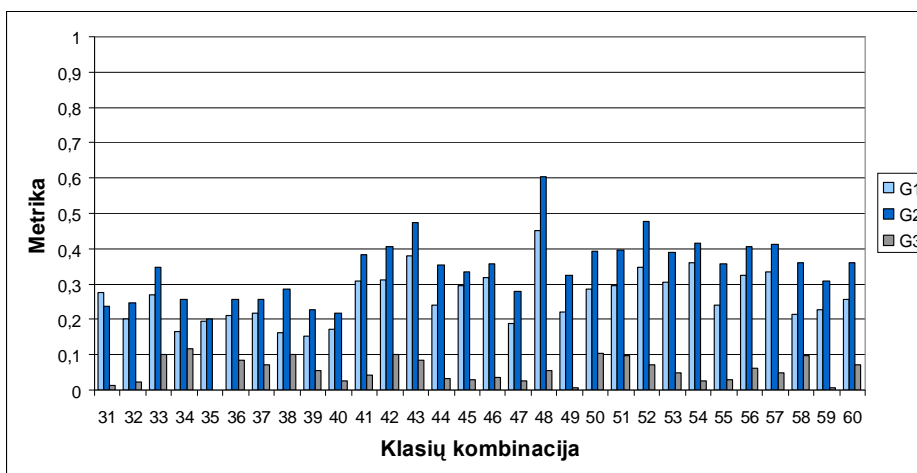
Taip pat priede A pristatomos klasių kombinacijų numeraciją atitinkančios klasės, priede B aprašomi metrikų vidurkiai nagrinėjamų klasių kombinacijų atžvilgiu su 80 % pasikliautinumo intervalais.

Pateikiami *eksperimentinių tyrimų rezultatai naudojant duomenų rinkinį DMI*. Pristatomi metrikų G1, G2, G3 vidurkių įverčiai nagrinėjamų klasių kombinacijų atžvilgiu:

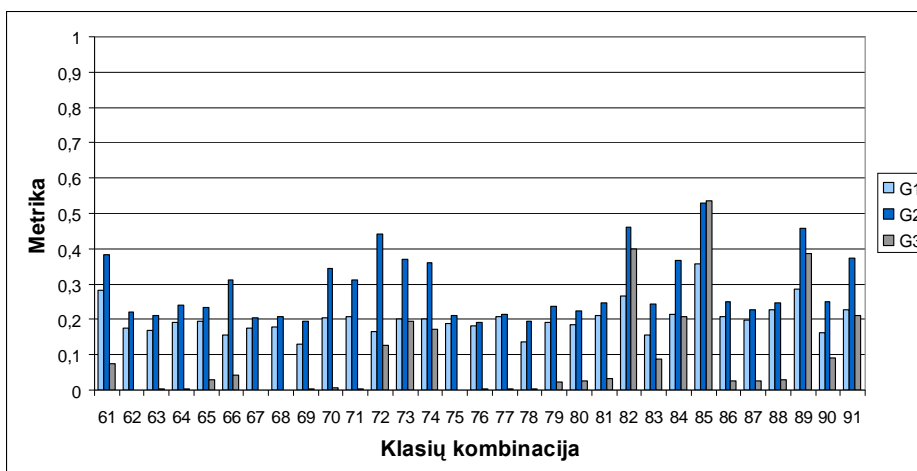
- 12 eilės TSPKK požymių sistemos atveju metrikų G1, G2, G3 vidurkių įverčių diagramos pateikiamos atitinkamai klasių kombinacijoms: 3.1 paveiksle 1–30 klasių kombinacijoms, 3.2 paveiksle 31–60 ir 3.3 paveiksle 61–91 (klasių kombinacijų numeraciją atitinkančios klasės pateiktos priede A).
- 12 eilės KK požymių sistemos atveju metrikų G1, G2, G3 vidurkių įverčiai pateikiami: 3.4 paveiksle 1–30 klasių kombinacijoms, 3.5 paveiksle 31–60 ir 3.6 paveiksle 61–91 (klasių kombinacijų numeraciją atitinkančios klasės aprašytos priede A).



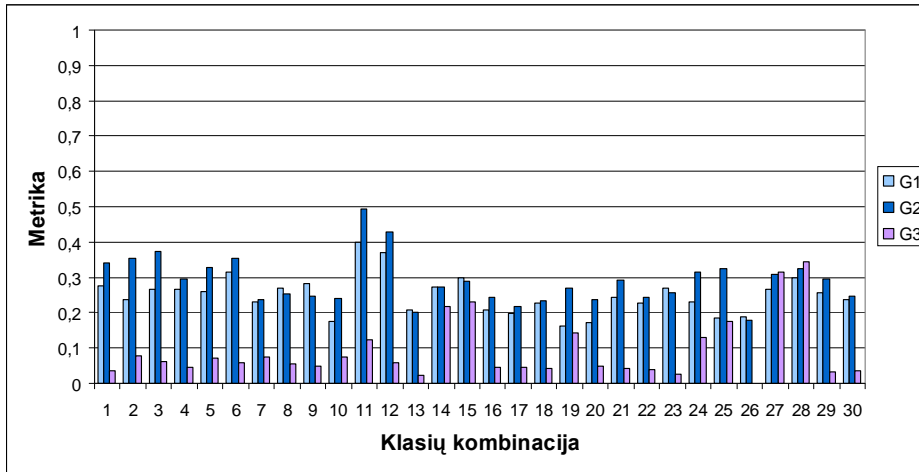
3.1 pav. Metrikų G1, G2, G3 vidurkių įverčiai 1–30 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM1 duomenys)



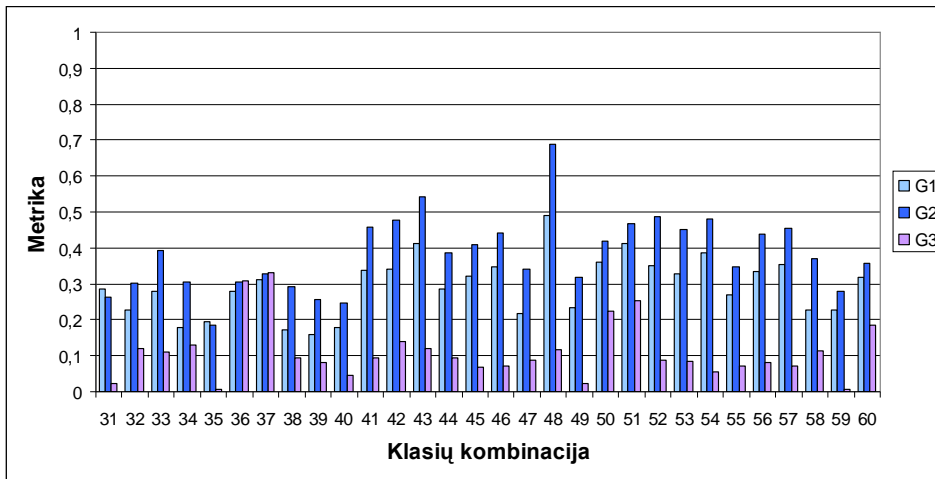
3.2 pav. Metrikų G1, G2, G3 vidurkių įverčiai 31–60 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM1 duomenys)



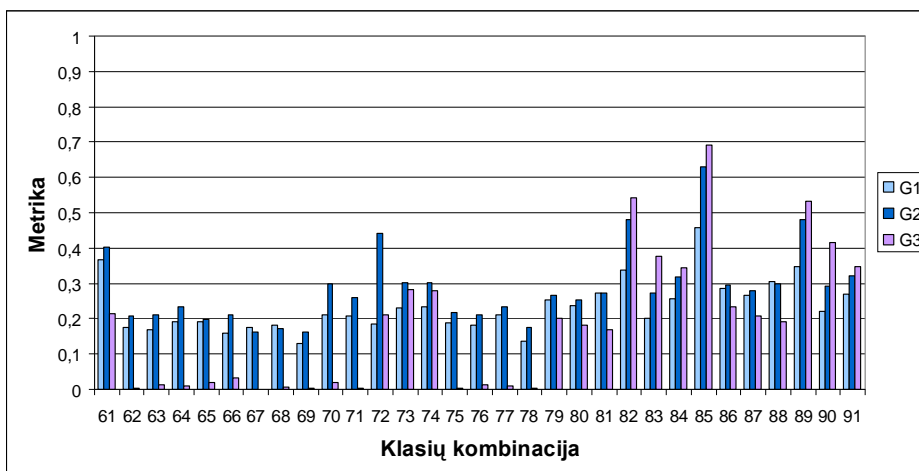
3.3 pav. Metrikų G1, G2, G3 vidurkių įverčiai 61–91 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM1 duomenys)



3.4 pav. Metrikų G1, G2, G3 vidurkių įverčiai 1–30 klasių kombinacijoms, KK požymių sistemos atveju (DM1 duomenys)



3.5 pav. Metrikų G1, G2, G3 vidurkių įverčiai 31–60 klasių kombinacijoms, KK požymių sistemos atveju (DM1 duomenys)



3.6 pav. Metrikų G1, G2, G3 vidurkių įverčiai 61–91 klasių kombinacijoms, KK požymių sistemos atveju (DM1 duomenys)

Naudojant pirmąjį duomenų rinkinį DM1, iš eksperimentinių tyrimų rezultatų, pateiktų (3.1)–(3.6) paveiksluose, kurie gauti apskaičiuavus metrikų G1, G2, G3 vidurkių įverčius – 12 eilės TSPKK ir 12 eilės KK požymių sistemoms, matyti:

- 12 eilės TSPKK (3.1–3.3 pav.) ir 12 eilės KK (3.4–3.6 pav.) požymių sistemų metrikų vidurkių įverčių tendencija yra: dauguma atvejų metrikos G3 įgyjamos reikšmės ženkliai mažesnės nei metrikų G1, G2 reikšmės. Tuo tarpu metrikos G1 reikšmės ne taip ženkliai mažesnės nei G2. Taigi klases vertinant skirtingomis metrikomis, gaunama skirtinga metrikų įverčių reikšmių tendencija.
- Atsižvelgiant į kiekvieno tipo metrikos vidurkių įverčius atskirų požymių sistemų aspektu, galime pastebėti tendenciją, kad dauguma atvejų 12 eilės TSPKK požymių sistemos metrikų vidurkių įgyjami įverčiai mažesni nei KK sistemos, kur maža metrikos reikšmė nurodo gerą klasių atskiriamumą, didelė – blogą klasių atskiriamumą.

3.3 lentelėje pateikti 12 eilės TSPKK ir 12 eilės KK požymių sistemoms gauti metrikų vidurkių įverčių intervalai.

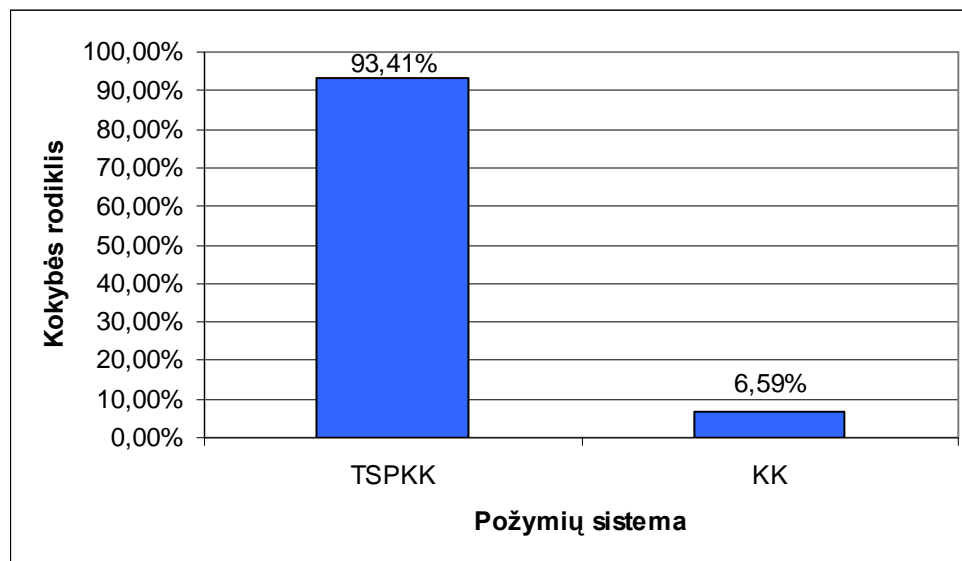
3.3 lentelė. Metrikų vidurkių įverčių intervalai (DM1 duomenų rinkinys)

Požymių sistema	Metrika	Metrikos vidurkio įverčių reikšmių intervalas	Klasės įgyjančios minimalią metrikos vidurkio reikšmę	Klasės įgyjančios maksimalią metrikos vidurkio reikšmę
TSPKK	G1	[0,13; 0,45]	s'-o:	r'-r
	G2	[0,19; 0,60]	s-e	r'-r
	G3	[0,01; 0,57]	s-a	t'-t
KK	G1	[0,13; 0,49]	s'-o:	r'-r
	G2	[0,16; 0,69]	s'-o:	r'-r
	G3	[0,01; 0,69]	m-s	t'-t

Iš 3.3 lentelėje pateiktų metrikų vidurkių įverčių intervalų rezultatų yra matyti, kad:

- Minimalios metrikų reikšmės gaunamos skirtingoms klasių kombinacijoms, tarp jų trimis atvejais minimali reikšmė gaunama s'-o: klasių kombinacijai. Tai rodo gerą šių klasių atskiriamumą.
- Maksimalios metrikų reikšmės gaunamos r'-r ir t'-t klasių kombinacijoms. Rezultatai rodo blogą šių klasių atskiriamumą.

Apskaičiavus metrikų G1, G2, G3 vidurkius atskiroms klasių kombinacijoms, randami požymių sistemų kokybės rodikliai. 12 eilės TSPKK ir 12 eilės KK požymių sistemų kokybės rodiklių rezultatai yra pateikti 3.7 paveiksle.



3.7 pav. Požymių sistemos kokybės rodiklio įvertis gautas naudojant pateiktą metodą (DM1 duomenų rinkinys)

Iš eksperimentinių tyrimų rezultatų, gautų apskaičiavus kokybės rodiklio įverčius 12 eilės TSPKK ir 12 eilės KK požymių sistemoms, yra matyti, kad 12 eilės TSPKK požymių sistemos atveju kokybės rodiklis yra aukštesnis nei 12 eilės KK požymių sistemos atveju. Rezultatai rodo, kad naudojant pateiktą šnekos signalų atpažinimo požymių kokybės vertinimo metodą, DM1 duomenų rinkinio atveju, 12 eilės TSPKK yra nustatyta kokybiškąją požymių sistema.

Sukurto metodo teisingumui patikrinti, apskaičiuoti AK ir DLSK klasifikavimo klaidų įverčiai. Rezultatai pateikti 3.4 lentelėje.

3.4 lentelė. AK, DLSK klasifikavimo klaidų įverčiai (DM1 duomenų rinkinys)

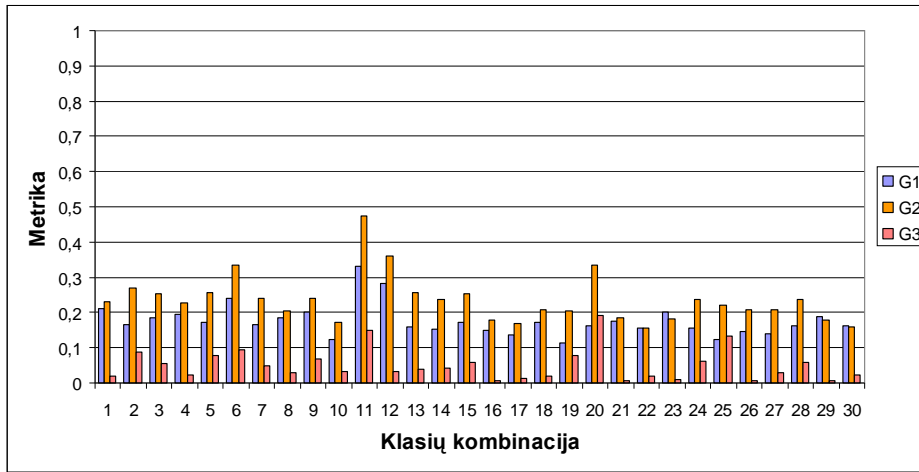
Klasifikatorius	Požymių sistema	Klaidos įvertis
AK	TSPKK	9,14 ± 1,11 %
	KK	13,48 ± 1,51 %
DLSK	TSPKK	3,63 ± 0,76 %
	KK	6,17 ± 1,06 %

Iš eksperimentinių tyrimų rezultatų, gautų apskaičiavus AK ir DLSK klasifikavimo klaidų įverčius, yra matyti:

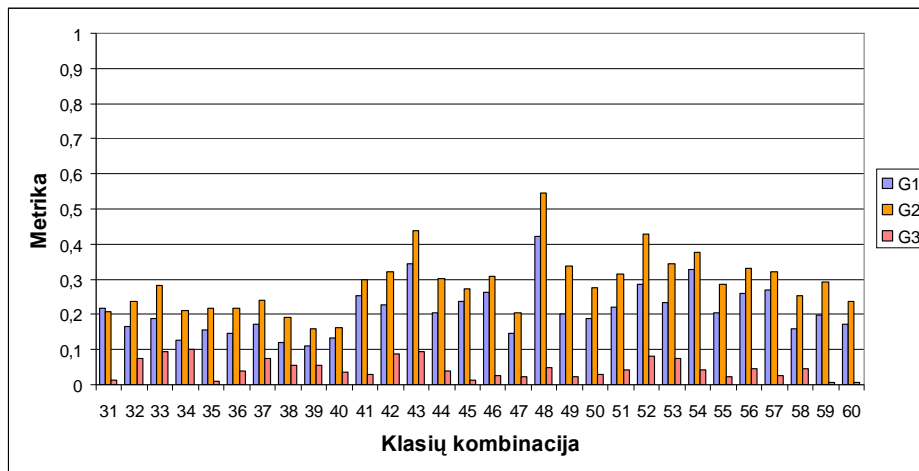
- Atpažinimo sistemai, naudojančiai TSPKK požymius ir AK klasifikatorių, gautas mažesnis klaidų įvertis nei atpažinimo sistemai, naudojančiai KK požymius ir AK klasifikatorių. Rezultatai rodo, kad naudojant AK klasifikatorių, TSPKK yra nustatyta kokybiškąją požymių sistema.
- Atpažinimo sistemai, naudojančiai TSPKK požymius ir DLSK klasifikatorių, gautas mažesnis klaidų įvertis nei atpažinimo sistemai, naudojančiai KK požymius ir DLSK klasifikatorių. Rezultatai rodo, kad naudojant DLSK klasifikatorių, TSPKK yra nustatyta kokybiškąją požymių sistema.

Pateikiami *eksperimentinių tyrimų rezultatai naudojant duomenų rinkinį DM2*. Aprašomi gauti metrikų G1, G2, G3 vidurkių įverčiai nagrinėjamų klasių kombinacijų atžvilgiu:

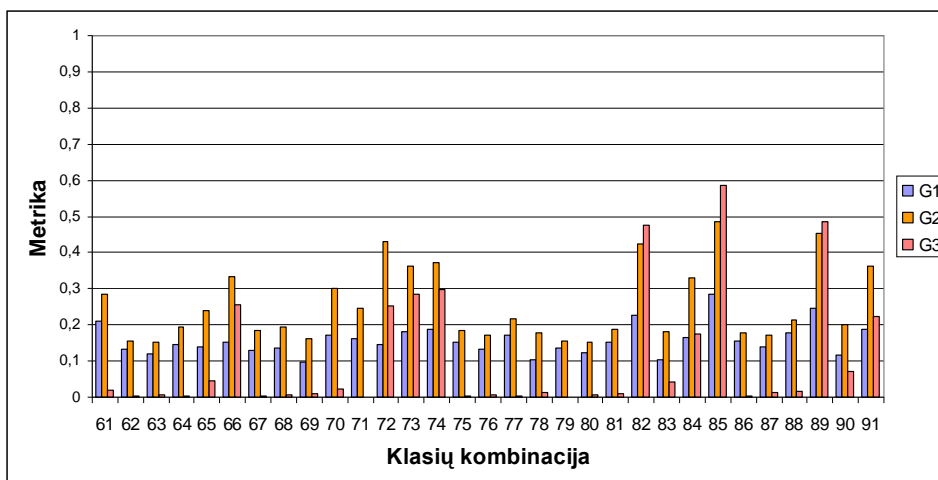
- 12 eilės TSPKK požymių sistemos atveju metrikų G1, G2, G3 vidurkių įverčiai pateikiami atitinkamai klasių kombinacijoms: 3.8 paveiksle pateikiama 1–30 klasių kombinacijoms, 3.9 paveiksle 31–60, tuo tarpu 3.10 paveiksle 61–91 klasių kombinacijoms.
- 12 eilės KK požymių sistemos atveju metrikų G1, G2, G3 vidurkių įverčiai pateikiami: 3.11 paveiksle 1–30 klasių kombinacijoms, 3.12 paveiksle 31–60, 3.13 paveiksle 61–91 klasių kombinacijoms.



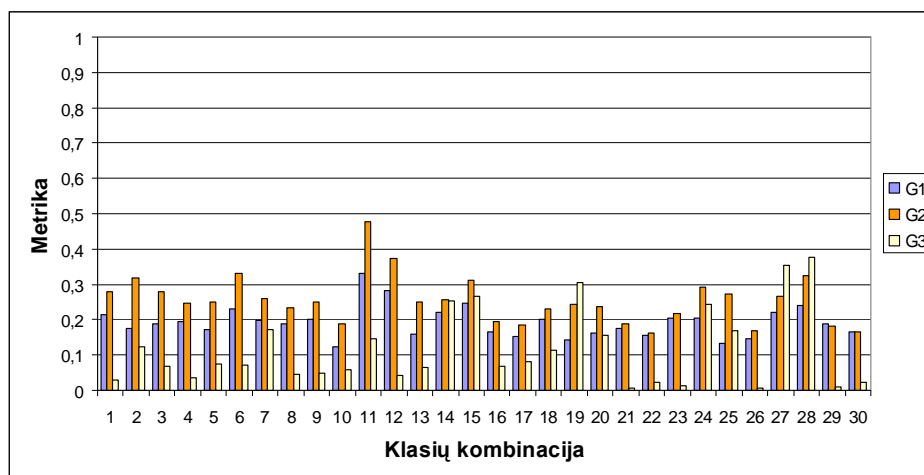
3.8 pav. Metrikų G1, G2, G3 vidurkių įverčiai 1–30 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM2 duomenys)



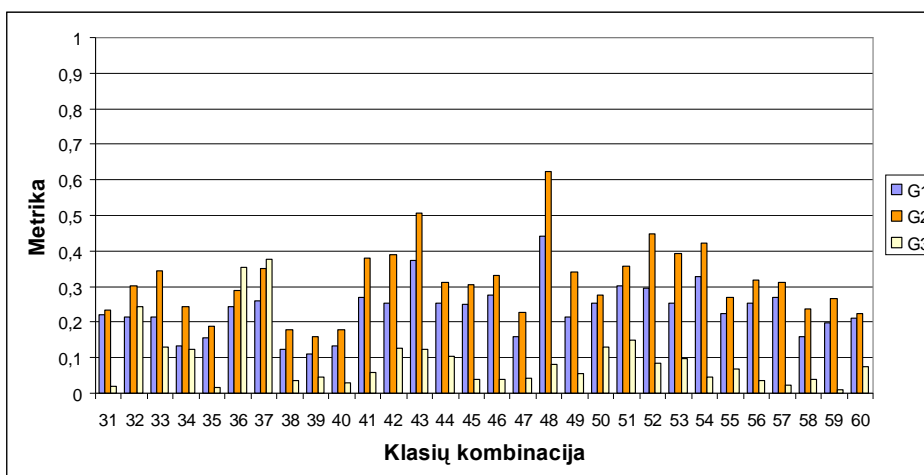
3.9 pav. Metrikų G1, G2, G3 vidurkių įverčiai 31–60 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM2 duomenys)



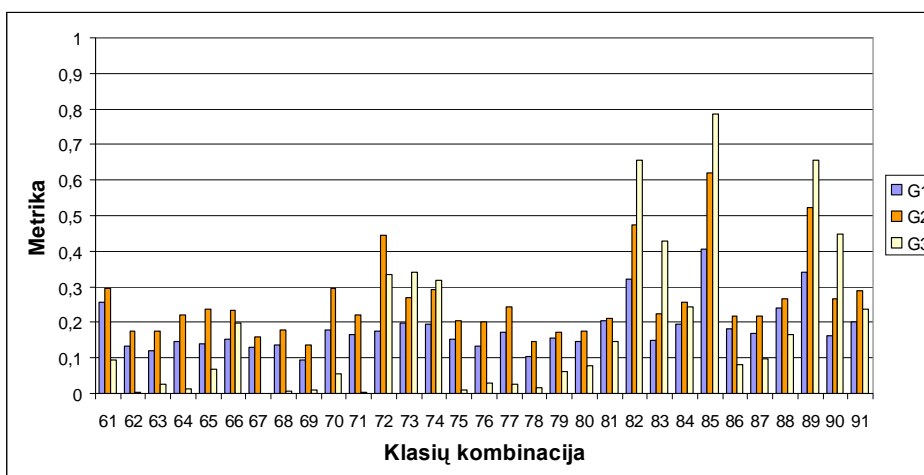
3.10 pav. Metrikų G1, G2, G3 vidurkių įverčiai 61–91 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM2 duomenys)



3.11 pav. Metrikų G1, G2, G3 vidurkių įverčiai 1–30 klasių kombinacijoms, KK požymių sistemos atveju (DM2 duomenys)



3.12 pav. Metrikų G1, G2, G3 vidurkių įverčiai 31–60 klasių kombinacijoms, KK požymių sistemos atveju (DM2 duomenys)



3.13 pav. Metrikų G1, G2, G3 vidurkių įverčiai 61–91 klasių kombinacijoms, KK požymių sistemos atveju (DM2 duomenys)

Naudojant antrąjį duomenų rinkinį DM2, iš eksperimentinių tyrimų rezultatų, gautų apskaičiuavus metrikų G1, G2, G3 vidurkių įverčius, yra matyti:

- 12 eilės TSPKK (3.8–3.10 pav.) ir 12 eilės KK (3.11–3.13 pav.) požymių sistemoms, kaip ir DM1 duomenų rinkinio atveju, pastebima metrikų vidurkių įverčių tendencija: dauguma atvejų metrikos G3 įgyjamos reikšmės ženkliai mažesnės nei metrikų G1 ir G2 reikšmės. Metrikos G1 reikšmės ne taip ženkliai mažesnės nei G2. Klases vertinant skirtingomis metrikomis, gaunama skirtinga metrikų įverčių reikšmių tendencija.
- Atsižvelgiant į kiekvieno tipo metrikos vidurkių įverčius atskirų požymių sistemų aspektu, kaip ir DM1 duomenų rinkinio atveju, pastebima tendencija: dauguma atvejų 12 eilės TSPKK požymių sistemos metrikų vidurkių įgyjami įverčiai mažesni nei 12 eilės KK sistemos, kur maža metrikos reikšmė nurodo gerą klasių atskiriamumą, didelė – blogą klasių atskiriamumą.

3.5 lentelėje pateikti 12 eilės TSPKK ir 12 eilės KK požymių sistemoms gauti metrikų vidurkių įverčių intervalai.

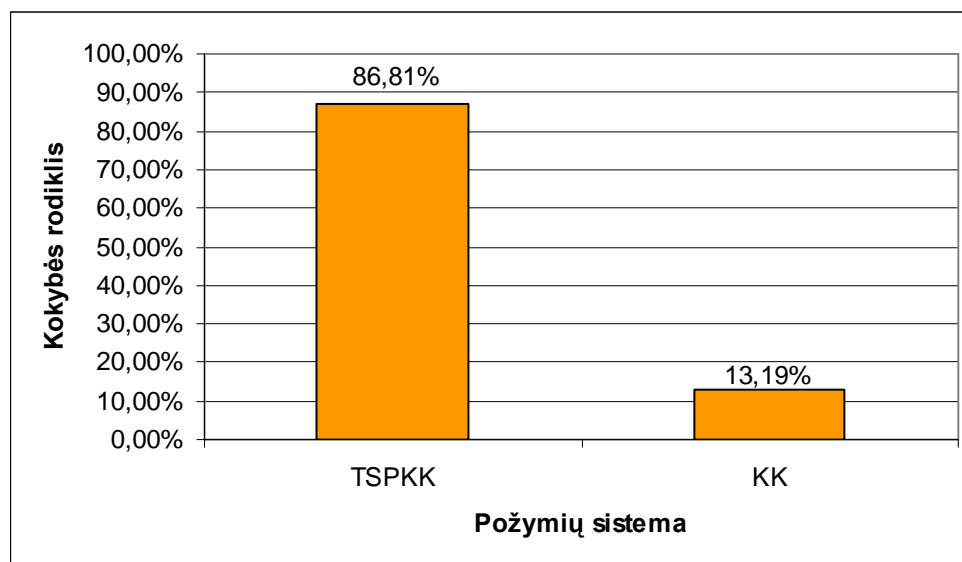
3.5 lentelė. Metrikų vidurkių įverčių intervalai (DM2 duomenų rinkinys)

Požymių sistema	Metrika	Metrikos vidurkio įverčių reikšmių intervalas	Klasės įgyjančios minimalią metrikos vidurkio reikšmę	Klasės įgyjančios maksimalią metrikos vidurkio reikšmę
TSPKK	G1	[0,10; 0,42]	s'-o:	r'-r
	G2	[0,15; 0,55]	t'-e	r'-r
	G3	[0,01; 0,58]	s'-r	t'-t
KK	G1	[0,10; 0,44]	s'-o:	r'-r
	G2	[0,14; 0,62]	s'-o:	r'-r
	G3	[0,01; 0,79]	s'-m	t'-t

Iš 3.5 lentelėje pateiktų metrikų G1, G2, G3 vidurkių įverčių intervalų yra matyti, kad:

- Minimalios metrikų reikšmės gaunamos skirtingoms klasių kombinacijoms, tarp jų trimis atvejais minimali reikšmė gaunama s'-o: klasių kombinacijai. Rezultatai rodo, kad šios klasės yra gerai atskiriamos.
- Tuo tarpu maksimalios metrikų reikšmės gaunamos r'-r ir t'-t klasių kombinacijoms. Taigi rezultatai rodo, kad šios klasės yra blogai atskiriamos.

Apskaičiavus metrikų G1, G2, G3 vidurkių įverčius atskiroms klasių kombinacijoms, randami požymių sistemų kokybės rodikliai. Požymių sistemų kokybės rodiklių rezultatai pateikti 3.14 paveiksle.



3.14 pav. Požymių sistemos kokybės rodiklio įvertis gautas naudojant pateiktą metodą (DM2 duomenų rinkinys)

Iš eksperimentinių tyrimų rezultatų, gautų apskaičiavus požymių sistemų kokybės rodiklio įverčius, yra matyti, kad 12 eilės TSPKK požymių sistemos atveju kokybės rodiklis yra aukštesnis nei 12 eilės KK požymių sistemos atveju. Rezultatai rodo, kad naudojant pateiktą šnekos signalų atpažinimo požymių kokybės vertinimo metodą, DM2 duomenų rinkinio atveju, 12 eilės TSPKK yra nustatyta kokybiškąją požymių sistema.

Pateikto metodo teisingumui patikrinti, apskaičiuoti AK ir DLSK klasifikavimo klaidų įverčiai (3.6 lentelė).

3.6 lentelė. AK, DLSK klasifikavimo klaidų įverčiai (DM2 duomenų rinkinys)

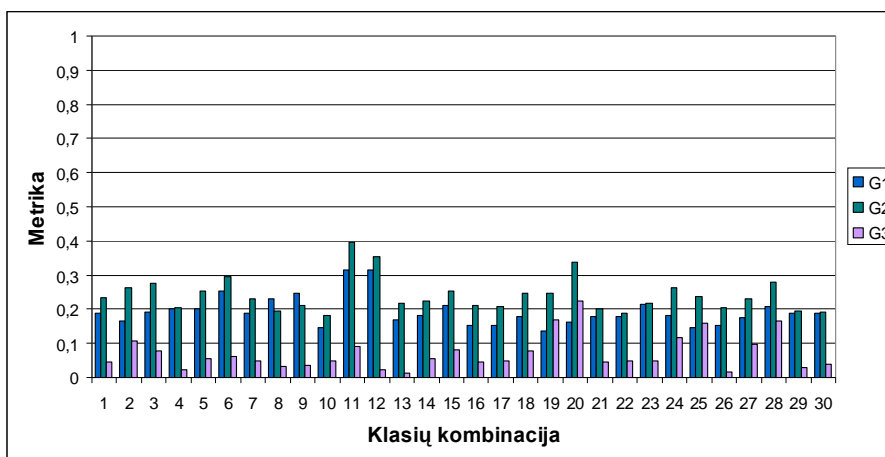
Klasifikatorius	Požymių sistema	Klaidos įvertis
AK	TSPKK	8,12 ± 0,93 %
	KK	13,12 ± 1,50 %
DLSK	TSPKK	2,48 ± 0,62 %
	KK	4,67 ± 0,73 %

Iš gautų rezultatų, apskaičiavus AK ir DLSK klasifikavimo klaidų įverčius, yra matyti:

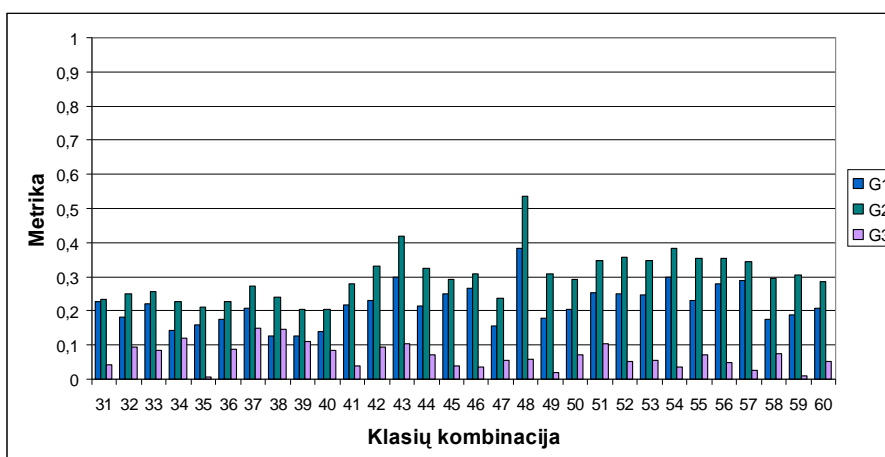
- Atpažinimo sistemai, naudojančiai TSPKK požymius ir AK klasifikatorių, gautas mažesnis klaidų įvertis nei atpažinimo sistemai, naudojančiai KK požymius ir AK klasifikatorių. Rezultatai rodo, kad naudojant AK klasifikatorių, TSPKK yra nustatyta kokybiškąją požymių sistema.
- Atpažinimo sistemai, naudojančiai TSPKK požymius ir DLSK klasifikatorių, gautas mažesnis klaidų įvertis nei atpažinimo sistemai, naudojančiai KK požymius ir DLSK klasifikatorių. Rezultatai rodo, kad naudojant DLSK klasifikatorių, TSPKK yra nustatyta kokybiškąją požymių sistema.

Pateikiami *eksperimentinių tyrimų rezultatai naudojant duomenų rinkinį DM3*. Pristatomi gauti metrikų G1, G2, G3 vidurkių įverčiai nagrinėjamų klasių kombinacijų atžvilgiu:

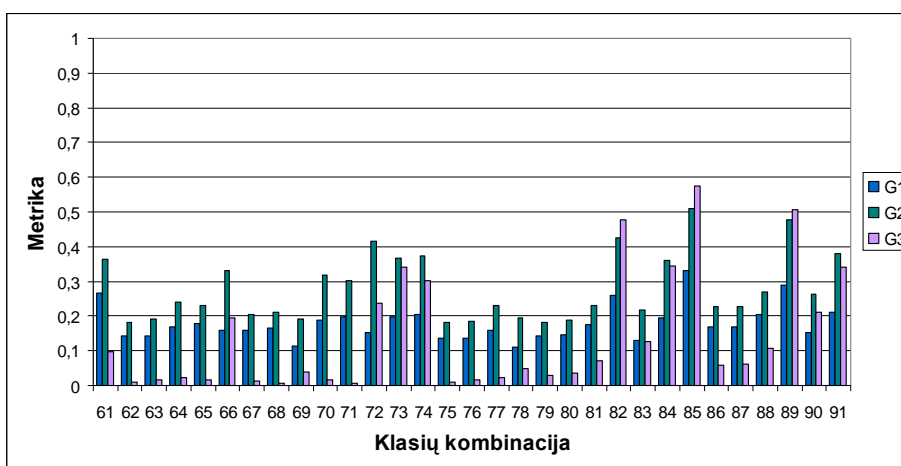
- 12 eilės TSPKK požymių sistemos atveju metrikų G1, G2, G3 vidurkių įverčiai pateikiami klasių kombinacijoms: 3.15 paveiksle 1–30 klasių kombinacijoms, 3.16 paveiksle 31–60, 3.17 paveiksle 61–91 klasių kombinacijoms.
- 12 eilės KK požymių sistemos atveju metrikų G1, G2, G3 vidurkių įverčiai pateikiami: 3.18 paveiksle 1–30 klasių kombinacijoms, 3.19 paveiksle 31–60 ir 3.20 paveiksle vaizduojami 61–91 klasių kombinacijoms.



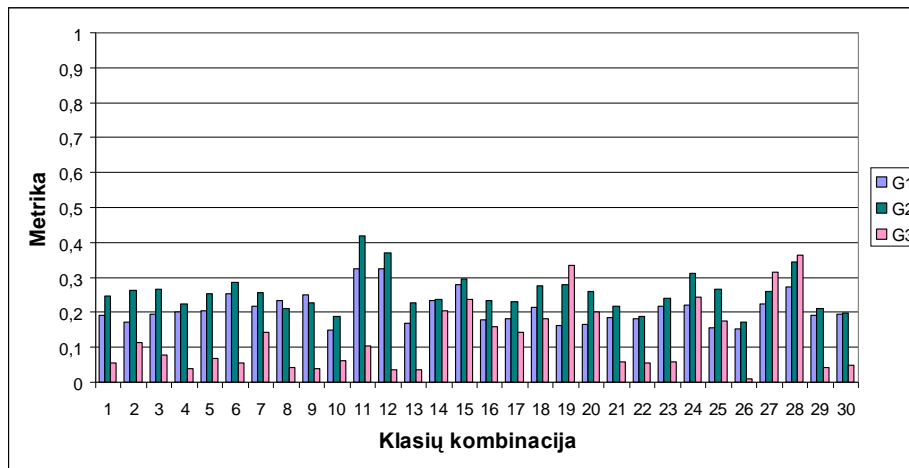
3.15 pav. Metrikų G1, G2, G3 vidurkių įverčiai 1–30 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM3 duomenys)



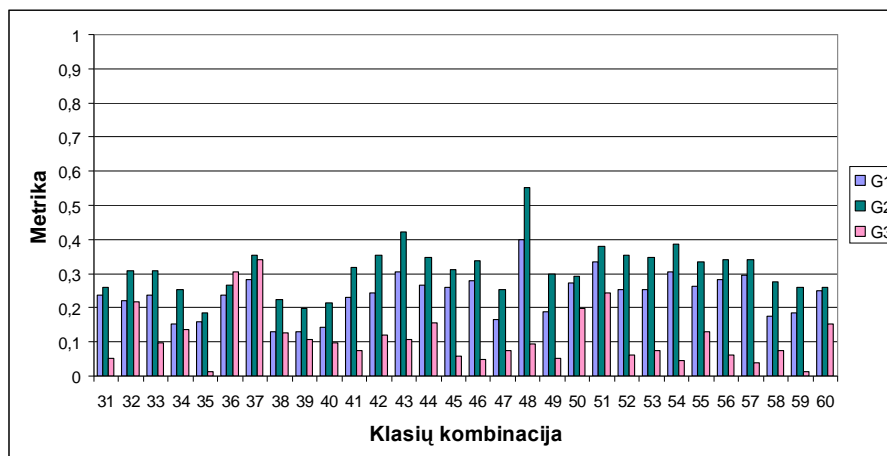
3.16 pav. Metrikų G1, G2, G3 vidurkių įverčiai 31–60 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM3 duomenys)



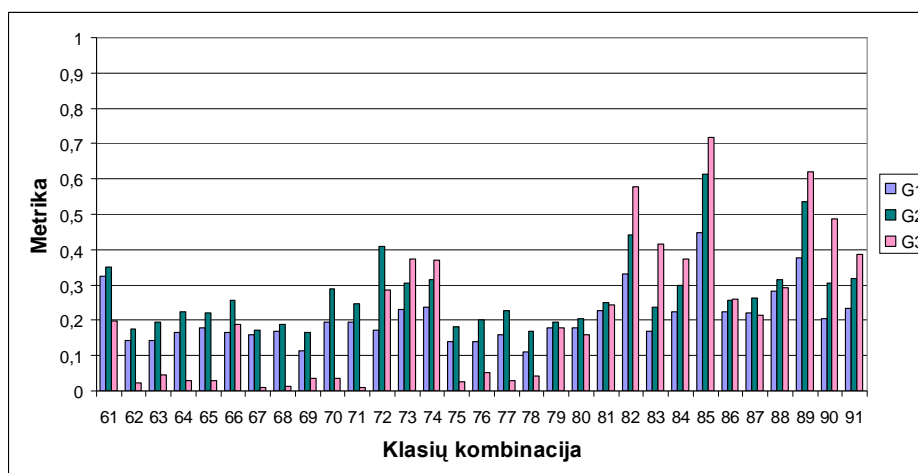
3.17 pav. Metrikų G1, G2, G3 įverčiai reikšmės 61–91 klasių kombinacijoms, TSPKK požymių sistemos atveju (DM3 duomenys)



3.18 pav. Metrikų G1, G2, G3 vidurkių įverčiai 1–30 klasių kombinacijoms, KK požymių sistemos atveju (DM3 duomenys)



3.19 pav. Metrikų G1, G2, G3 vidurkių įverčiai 31–60 klasių kombinacijoms, KK požymių sistemos atveju (DM3 duomenys)



3.20 pav. Metrikų G1, G2, G3 vidurkių įverčiai 61–91 klasių kombinacijoms, KK požymių sistemos atveju (DM3 duomenys)

Naudojant trečiąją duomenų rinkinį DM3, iš eksperimentinių tyrimų rezultatu, gautų apskaičiuotus metrikų G1, G2, G3 vidurkių įverčius matyti:

- 12 eilės TSPKK (3.15–3.17 pav.) ir 12 eilės KK (3.18–3.20 pav.) požymių sistemų atvejais, pastebima metrikų vidurkių įverčių tendencija (kaip ir DM1, DM2 duomenų rinkinių atvejais): dauguma atvejų metrikos G3 įgyjamos reikšmės daug mažesnės nei metrikų G1 ir G2. Metrikos G1 reikšmės ne taip ženkliai mažesnės nei G2. Klases vertinant skirtingomis metrikomis, gaunama skirtinga metrikų įverčių reikšmių tendencija.
- Atsižvelgiant į kiekvieno tipo metrikos vidurkių įverčius atskirų požymių sistemų aspektu, pastebima tendencija (kaip ir DM1, DM2 duomenų rinkiniams): dauguma atvejų 12 eilės TSPKK požymių sistemos metrikų vidurkių įgyjami įverčiai mažesni nei 12 eilės KK sistemos, kur maža metrikos reikšmė nurodo gerą klasių atskiriamumą, didelė – blogą klasių atskiriamumą.

3.7 lentelėje pateikti 12 eilės TSPKK ir 12 eilės KK požymių sistemoms apskaičiuoti metrikų vidurkių įverčių intervalai.

3.7 lentelė. Metrikų vidurkių įverčių intervalai (DM3 duomenų rinkinys)

Požymių sistema	Metrika	Metrikos vidurkio įverčių reikšmių intervalas	Klasės įgyjančios minimalią metrikos vidurkio reikšmę	Klasės įgyjančios maksimalią metrikos vidurkio reikšmę
TSPKK	G1	[0,11; 0,38]	s–o:	r'–r
	G2	[0,18; 0,54]	s–a	r'–r
	G3	[0,01; 0,57]	s'–r	t'–t
KK	G1	[0,11; 0,45]	s–o:	t'–t
	G2	[0,17; 0,62]	s'–o:	t'–t
	G3	[0,01; 0,72]	m–s	t'–t

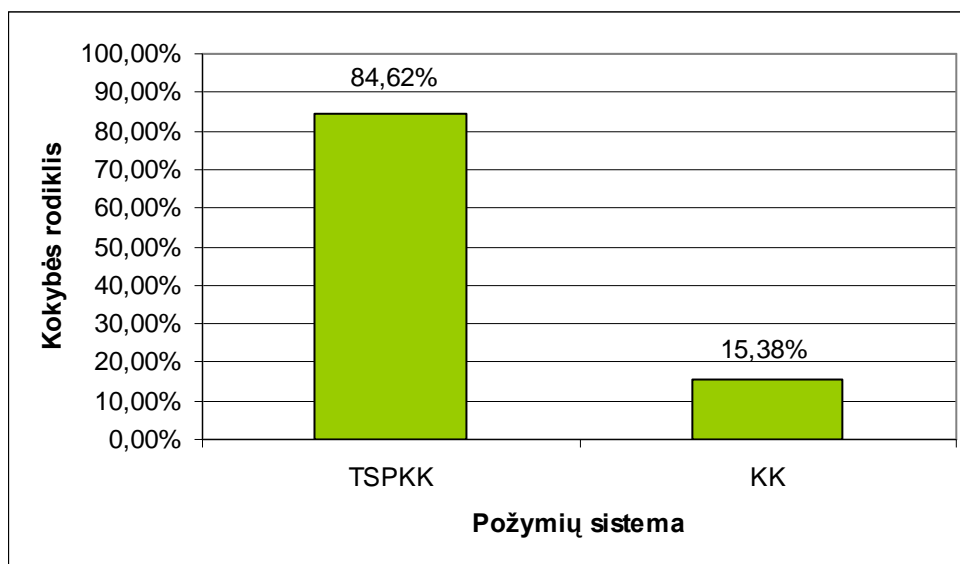
Iš 3.7 lentelėje pateiktų metrikų vidurkių įverčių intervalų yra matyti, kad:

- Minimalios metrikų reikšmės gaunamos skirtingoms klasių kombinacijoms, tarp jų dvejais atvejais minimali reikšmė gaunama s–o:

klasių kombinacijai ir vienas atvejis gaunamas s'–o: klasių kombinacijai. Tai rodo gerą šių klasių atskiriamumą.

- Maksimalios metrikų reikšmės gaunamos r'–r ir t'–t klasių kombinacijoms. Rezultatai rodo blogą šių klasių atskiriamumą.

Apskaičiavus metrikų G1, G2, G3 vidurkius atskiroms klasių kombinacijoms, randami požymių sistemų kokybės rodikliai (3.21 pav.).



3.21 pav. Požymių sistemos kokybės rodiklio įvertis gautas naudojant pateiktą metodą (DM3 duomenų rinkinys)

Iš rezultatų (3.21 pav.) matyti, kad 12 eilės TSPKK požymių sistemos atveju kokybės rodiklis yra aukštesnis nei 12 eilės KK požymių sistemos atveju. Taigi naudojant pateiktą metodą, DM3 duomenų rinkinio atveju, 12 eilės TSPKK – nustatyta kokybiškąją požymių sistema. 3.8 lentelėje pateikti apskaičiuoti AK ir DLSK klasifikavimo klaidų įverčiai.

3.8 lentelė. AK, DLSK klasifikavimo klaidų įverčiai (DM3 duomenų rinkinys)

Klasifikatorius	Požymių sistema	Klaidos įvertis
AK	TSPKK	10,09 ± 1,18 %
	KK	13,66 ± 1,55 %
DLSK	TSPKK	3,53 ± 0,63 %
	KK	5,70 ± 0,78 %

Iš eksperimentinių tyrimų rezultatų, gautų apskaičiavus AK ir DLSK klasifikavimo klaidų įverčius, yra matyti:

- Atpažinimo sistemai, naudojančiai TSPKK požymius ir AK klasifikatorių, gautas mažesnis klaidų įvertis nei atpažinimo sistemai, naudojančiai KK požymius ir AK klasifikatorių. Rezultatai rodo, kad naudojant AK klasifikatorių, TSPKK yra nustatyta kokybiškąją požymių sistema.
- Atpažinimo sistemai, naudojančiai TSPKK požymius ir DLSK klasifikatorių, gautas mažesnis klaidų įvertis nei atpažinimo sistemai, naudojančiai KK požymius ir DLSK klasifikatorių. Rezultatai rodo, kad naudojant DLSK klasifikatorių, TSPKK yra nustatyta kokybiškąją požymių sistema.

Eksperimentinių tyrimų rezultatai patvirtino pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą AK ir DLSK atpažinimo sistemų atvejais.

3.3. Skyriaus rezultatai ir išvados

- Atlikti eksperimentiniai tyrimai, kurių tikslas – eksperimentiškai patvirtinti pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumą. Eksperimentiniams tyrimams atlikti panaudota sukurta eksperimentinė bazė.
- Eksperimentiškai patvirtintas pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumas artimiausio kaimyno ir dinaminio laiko skalės kraipymo atpažinimo sistemų pavyzdžiais.
- Parodyta, kad pirmojo duomenų rinkinio DM1 atveju (kalbėtojas vyras), kai apskaičiuotas tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 93,41 %, o kepstro koeficientų požymių sistemos kokybės rodiklis – 6,59 %, gauti tokie signalų atpažinimo rezultatai:

1. Atpažinimo sistemai, naudojančiai tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, gauta $9,14 \pm 1,11$ % klaidų, o atpažinimo sistemai, naudojančiai kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, gauta $13,48 \pm 1,51$ % klaidų.
 2. Atpažinimo sistemai, naudojančiai tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, gauta $3,63 \pm 0,76$ % klaidų, o atpažinimo sistemai, naudojančiai kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, gauta $6,17 \pm 1,06$ % klaidų.
- Parodyta, kad antrojo duomenų rinkinio DM2 atveju (kalbėtoja moteris), kai apskaičiuotas tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 86,81 %, o kepstro koeficientų požymių sistemos kokybės rodiklis – 13,19 %, gauti tokie signalų atpažinimo rezultatai:
 1. Atpažinimo sistemai, naudojančiai tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, gauta $8,12 \pm 0,93$ % klaidų, o atpažinimo sistemai, naudojančiai kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, gauta $13,12 \pm 1,50$ % klaidų.
 2. Atpažinimo sistemai, naudojančiai tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, gauta $2,48 \pm 0,62$ % klaidų, o atpažinimo sistemai, naudojančiai kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, gauta $4,67 \pm 0,73$ % klaidų.
 - Parodyta, kad trečiojo duomenų rinkinio DM3 atveju (du kalbėtojai vyrai, dvi kalbėtojos moterys), kai apskaičiuotas tiesinės suvokimo

prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 84,62 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 15,38 %, gauti tokie signalų atpažinimo rezultatai:

1. Atpažinimo sistemai, naudojančiai tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, gauta $10,09 \pm 1,18$ % klaidų, o atpažinimo sistemai, naudojančiai kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, gauta $13,66 \pm 1,55$ % klaidų.
2. Atpažinimo sistemai, naudojančiai tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, gauta $3,53 \pm 0,63$ % klaidų, o atpažinimo sistemai, naudojančiai kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, gauta $5,70 \pm 0,78$ % klaidų.

Bendrosios išvados

1. Pateiktas šnekos signalų atpažinimo požymių kokybės vertinimo metodas, grindžiamas metrikų naudojimu.

2. Pateiktas metrikų rinkinys šnekos signalų atpažinimo požymių kokybei vertinti, susidedantis iš trijų metrikų: požymių kiekio klasės ribose, klasių artimiausių kaimynų atstumų santykio, klasės ribos peržengimo kiekio.

3. Parodyta, kad šnekos signalų atpažinimo požymių kokybė Euklido erdvėje aprašo atpažinimo sistemų kokybę ir nereikalauja klasifikavimo eksperimentų vykdymo.

4. Įvesta požymių kokybės rodiklio skalė 0 %–100 %, kur 0 % rodo žemiausią požymių sistemos kokybę, o 100 % – aukščiausią požymių sistemos kokybę.

5. Parodyta, kad šnekos signalų požymių kokybės vertinimo metodo algoritmo sudėtingumas yra $O(2R \log 2R)$, kai dinaminio laiko skalės kraipymo atpažinimo sistemos kokybės vertinimo algoritmo sudėtingumas yra $O(R^2)$, kur R – šnekos signalo objekto vektorių skaičius.

6. Sukurta šnekos signalų atpažinimo požymių kokybės vertinimo tyrimams pritaikyta eksperimentinė bazė.

7. Patvirtintas pateikto šnekos signalų atpažinimo požymių kokybės vertinimo metodo teisingumas artimiausio kaimyno ir dinaminio laiko skalės kraipymo atpažinimo sistemų pavyzdžiais.

8. Parodyta, kad pirmojo duomenų rinkinio atveju (kalbėtojas vyras), kai tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 93,41 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 6,59 %, signalų atpažinimo rezultatai buvo tokie:

- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $9,14 \pm 1,11$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $13,48 \pm 1,51$ % klaidų.
- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $3,63 \pm 0,76$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $6,17 \pm 1,06$ % klaidų.

9. Parodyta, kad antrojo duomenų rinkinio atveju (kalbėtoja moteris), kai tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 86,81 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 13,19 %, signalų atpažinimo rezultatai buvo tokie:

- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $8,12 \pm 0,93$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $13,12 \pm 1,50$ % klaidų.
- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $2,48 \pm 0,62$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $4,67 \pm 0,73$ % klaidų.

10. Parodyta, kad trečiojo duomenų rinkinio atveju (du kalbėtojai vyrai, dvi kalbėtojos moterys), kai tiesinės suvokimo prognozės kepstro koeficientų požymių sistemos kokybės rodiklis buvo 84,62 %, o kepstro koeficientų požymių sistemos kokybės rodiklis buvo 15,38 %, signalų atpažinimo rezultatai buvo tokie:

- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $10,09 \pm 1,18$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir artimiausio kaimyno klasifikatorių, darė $13,66 \pm 1,55$ % klaidų.
- Atpažinimo sistema, naudojanti tiesinio suvokimo prognozės kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $3,53 \pm 0,63$ % klaidų, o atpažinimo sistema, naudojanti kepstro koeficientų požymius ir dinaminio laiko skalės kraipymo klasifikatorių, darė $5,70 \pm 0,78$ % klaidų.

11. Tikslinga tęsti atpažinimo požymių kokybės vertinimo metodų tyrimus sprendžiant mechaninių ir biologinių dinaminių sistemų būsenų atpažinimo problemas.

Literatūra ir šaltiniai

- Agarwal, P. K.; Edelsbrunner, H.; Schwarzkopf, O.; Welzl, E. 1991. Euclidean minimum spanning trees and bichromatic closest pairs, *Discrete and Computational Geometry (Springer)*, 6 (1): 407–422.
- Ai-bing, N.; Liang, M.; Xiao-hua, X. 2007. Solving degree – constrained minimum spanning tree with a new algorithm, *Proceedings of 2007 International Conference on Management Science & Engineering (14th)*, 381–386.
- Aissa, M.; Ben Mnaouer, A.; Belghith, A.; Shaqra. 2009. A novel extension of Kruskal's algorithm in multicast routing, *IEEE 34th Conference on Local Computer Networks (LCN 2009)*, 289–292.
- Alpert, C. J.; Hut, T. C.; Huang, J. H. and Kahng, A. B. 1993. A Direct Combination of the Prim and Dijkstra Constructions for Improved Performance-Driven Global Routing, *in Proceedings of ISCAS'1993*, 1869–1872.
- Alsmadi, M. K.; Omar, K. B.; Noah, S. A.; Almarashdah, I. 2009. Performance Comparison of Multi-layer Perceptron (Back Propagation, Delta Rule and Perceptron) algorithms in Neural Networks, *IEEE International Advance Computing Conference (IACC 2009)*, 296–299.

- Anusuya, M. A.; Katti, S. K. 2011. Classification Techniques used in Speech Recognition Applications: A review, (*IJCSIS*) *International Journal of Computer Science and Information Security* 6(3): 181–205.
- Anusuya, M. A.; Katti, S. K. 2009. Speech Recognition by Machine: A Review, (*IJCSIS*) *International Journal of Computer Science and Information Security* 2(4): 910–954.
- Arauzo–Azofra, A.; Aznarte, J. L. M.; Benitez, J. M. 2009. Empirical Study of Individual Feature Evaluators and Cutting Criteria for Feature Selection in Classification, *Ninth International Conference on Intelligent Systems Design and Applications*, 541–546.
- Arndt, C. 2001. *Information Measures: Information and its description in Science and Engineering*. Berlin: Springer. 547 p.
- Balagani, K. S.; Phoha, V. V. 2010. On the Feature Selection Criterion Based on an Approximation of Multidimensional Mutual Information, *IEEE Transactions on pattern analysis and machine intelligence* 32(7): 1342–1343.
- Bao, M.; Guan, L.; Li, X.; Tian, J.; Yang, J. 2006. Feature Extraction Using Histogram Entropies of Euclidean Distances for Vehicle Classification, *International Conference on Computational Intelligence and Security* V.1: 668–673.
- Bareisa, E.; Jusas, V.; Motiejunas, K.; Seinauskas, R. 2007. The Criteria of Functional Delay Test Quality Assessment, *IEEE 10th Euromicro Conference on Digital System Design Architectures, Methods and Tools (DSD 2007)*, 207–214.
- Basu, M.; Ho, T. K. 2006. *Data complexity in pattern recognition*. Springer London Ltd. 300 p.
- Battiti, R. 1994. Using mutual information for selecting features in supervised neural net learning, *IEEE Transactions on Neural Networks* 5(4): 537–550.

- Bernado–Mansilla, E.; Ho, T. K. 2005. Domain of competence of XCS classifier system in complexity measurement space, *IEEE Transactions on Evolutionary Computation* 9(1): 82–104.
- Bernado–Mansilla, E.; Ho, T. K. and Orriols A. 2006. Data Complexity and Evolutionary Learning. *Data Complexity in Pattern Recognition*, Springer, 115-134.
- Bernado–Mansilla, E.; Ho, T. K. 2004. On Classifier Domains of Competence, *Proceedings of the 17th International Conference on Pattern Recognition*, Cambridge, UK, 136–139.
- Berndt, D. J.; Clifford, J. 1994. Using dynamic time warping to find patterns in time series, *AAAI-94 Workshop on Knowledge Discovery in Databases (KDD-94)*, 359–370.
- Bock, F. 1971. An algorithm to construct a minimum direct spanning tree in a direct network, *Developments in Operations Research*, Gordon and Breach, New York, 29–44.
- Bohme, J. F. 1974. The cepstrum as a generalized function, *IEEE Transactions*, 20(5): 650–653.
- Buchin, K.; Mulzer, W. 2009. Delaunay triangulations in $O(\text{sort}(n))$ time and more, *Proc. 50th IEEE Symposium on Foundations of Computer Science*, 139–148.
- Boersma, P.; Weenink, D. 2000. PRAAT. Available at: <http://www.fon.hum.uva.nl/praat/>, prisijungta 2010.05.01.
- Callahan, P. and Kosaraju, S. 1993. Faster algorithms for some geometric graph problems in higher dimensions, *In Fourth annual ACM-SIAM Symposium on Discrete Algorithms*, 291–300.
- Chee, L. S.; Ai, O. C.; Hariharan, M.; Yaacob, S. 2009. MFCC based Recognition of Repetitions and Prolongations in Stuttered Speech using k-NN and LDA, *Proceedings of 2009 IEEE student Conference on Research and Development (SCOReD 2009)*, 146–149.

- Chen, C. H. 1971. Theoretical Comparison of a Class of Feature Selection Criteria in Pattern Recognition, *IEEE transaction on computers* C-20(9): 1054–1056.
- Chen, X.; Tian, J.; Cheng, J.; Yang, X. 2004. Segmentation of Fingerprint Images Using Linear Classifier, *EURASIP Journal on Applied Signal Processing*, *EURASIP JASP* 4: 480–494.
- Cheriton, D. and Tarjan, R. E. 1976. Finding minimum spanning trees, *In SIAM Journal on Computing* 5: 724–741.
- Childers, D. G.; Skinner, D. P and Kemerait, R. C. 1977. The cepstrum: A guide to processing, *Proc. IEEE* 65: 1428–1443.
- Cooley, J. W.; Lewis, P. A. W.; Welch, P. D. 1967. Historical notes on the fast Fourier transform, *IEEE Transactions on Audio and Electroacoustics* 15(2): 76–79.
- Costa, P.; Grouffaud, J.; Larzabal, P.; Clergeot, H.; 1999. Estimation of the number of signals from features of the covariance matrix: a supervised approach, *IEEE Transactions on Signal Processing* 47(11): 3108–3115.
- Cover, T. M.; Thomas, J. A. 2006. *Elements of Information Theory, Second Edition*. John Wiley and Sons. 748 p.
- Cover, T. M. and Hart, P. E. 1967. Nearest neighbour pattern classification, *IEEE Transactions on Information Theory* IT-13(1): 21–27.
- Chung, J, R.; Kwon, K. and Choe, Y. 2009. Evolution of Recollection and Prediction in Neural Networks, *Proceedings of International Joint Conference on Neural Networks, Atlanta, Georgia, USA*, 571–577.
- Davis, S. and Mermelstein, P. 1980. Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentence, *IEEE Transactions on Acoustics, Speech and Signal Processing* 28(4): 357–366.

- Doan, S.; Horiguchi, S. 2004. An efficient feature selection using multi-criteria in text categorization. *Proceedings of the IEEE Fourth International Conference on Hybrid Intelligent Systems*, 86–91.
- Duda, R. O.; Hart, P. E and Stork, D. G. 2001. *Pattern classification, Second edition*. John Wiley. 654 p.
- Dowdy, S.; Wearden, S.; Chilko, D. 2004: *Statistics for research, third edition*. New York: John Wiley. 627 p.
- Dzemyda, G.; Kurasova, O.; Žilinskas, J. 2008. *Daugiamačių duomenų vizualizavimo metodai*. Matematikos ir informatikos institutas. Mokslo Aidai, Vilnius 2008. 204 p.
- Eidukas, D., Kalnius, R. 2010. Models Quality of Electronics Products, *Electronics and Electrical Engineering*, 2(98): 29–34.
- Elizondo, D. A.; Birkenhead, R.; Gamez, M.; Garcia, N.; Alfaro, E. 2009. Estimation of Classification Complexity. *Proceedings of International Joint Conference on Neural Networks*, 764–770.
- Fontaine, V.; Ris, C.; Boite, J. M. 1997. Nonlinear Discriminant Analysis for Improved Speech Recognition, *Proceedings ICASSP 96*, 2071–2074.
- Furui, S. 2001. *Digital Speech Processing, Synthesis, and Recognition*. Second edition, Revised and Expanded. Signal Processing and Communications. 452 p.
- Gavrilis, D.; Tsoulos, I. G.; Dermatas, E. 2008. Selecting and constructing features using grammatical evolution, *Pattern Recognition Letters* 29(9): 1358–1365.
- Goyani, M.; Dave, N.; Patel, N. M. 2010. Performance Analysis of Lip Synchronization Using LPC, MFCC and PLP Speech Parameters, *International Conference on Computational Intelligence and Communication Networks*, 582–587.

- Golipour, L.; O'Shaughnessy, D. 2009. Context-independent phoneme recognition using a K-Nearest Neighbour classification approach, *IEEE International Conference on Acoustics Speech and Signal Processing*, 1341–1344.
- Graham, R. L.; Hell, P. 1985. History of the Minimum Spanning Tree Problem, *Annals of the History of Computing* 7(1): 43–57.
- Guyon, I., Elisseeff, 2003. A. An introduction to variable and feature selection, *The Journal of Machine Learning Research* V. 3: 1157–1182.
- Gunawan, W.; Hasegawa-Johnson, M. 2001. PLP coefficients can be quantized at 400 bps, *IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (ICASSP '01)* V.1: 77–80.
- Guo, W.; Gao, H. L.; Chen, G.; Yu, L. 2009. Particle Swarm Optimization for the Degree-constrained MST problem in WSN Topology Control, *The International Conference on Machine Learning and Cybernetics (ICMLC 2009)*, *IEEE Press, Baoding China* V.3: 1793–1798.
- Guo, X.; Yin, Y.; Dong, C.; Yang, G.; Zhou, G. 2008. On the Class Imbalance Problem, *Fourth International Conference on Natural Computation*, 192–201.
- Han, W.; Chan, C. F.; Choy, C. S.; Pun, K. P. 2006. An Efficient MFCC Extraction Method in Speech Recognition, *IEEE International Symposium on Circuits and Systems. ISCAS 2006. Proceedings*, 145–148.
- Heideman, M. T.; Johnson, D. H.; and Burrus, C. S. 1984. Gauss and the history of the FFT, *IEEE Acoust., Speech, Signal Processing Mag.*, 1: 14–21.
- Hermansky, H. 1990. Perceptual linear predictive (PLP) analysis of speech, *Journal of Acoustical Society of America* 87(4): 1738–1752.
- Hermansky, H.; Hanson, B. A.; Wakita, H. 1985. Perceptually based linear predictive analysis of speech, *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '85* V.10: 509–512.
- Ho, T. K. 2001. Data complexity analysis for classifier combination, *Springer Berlin, Heidelberg, Multiple classifier systems* V. 2096: 53–67.

- Ho, T. K. 2000. Complexity of classification problems and comparative advantages of combined classifiers, *Proceedings of the 1st International Workshop Multiple Classifier Systems*, 97–106.
- Ho, T. K. 2002. A data complexity analysis of comparative advantages of decision forest constructors, *Pattern Analysis and Applications V. 5*: 102–112.
- Ho, T. K.; Basu, M. 2002. Complexity measures of supervised classification problems, *IEEE transactions on pattern analysis and machine intelligence* 24 (3): 289–300.
- Hoekstra, A. and Duin, R. P. W. 1996. On the Nonlinearity of Pattern Classifiers, *Pattern Recognition, Proceedings of the 13th International Conference* 4: 271–275.
- Hossan, M. A.; Memon, S.; Gregory, M. A. 2010. A Novel Approach for MFCC Feature Extraction, *4th International Conference on Signal Processing and Communication Systems (ICSPCS)*, 1.
- Hua, O.; Ji, A.; He, O. 2010. Multiple Real-valued K Nearest Neighbor Classifiers System by Feature Grouping, *IEEE International Conference on Systems Man and Cybernetic (SMC)*, 3922–3925.
- Huang, F.; Gao, G.; Wang, Y. 2009. Comparison of Prim and Kruskal on Shanghai and Shenzhen 300 Index hierarchical structure tree, *International Conference on Web Information Systems and Mining*, 237–241.
- Hui, S.; Žark. H. 1994. The Widrow-Hoff Algorithm For McCulloch-Pitts Type Neurons, *IEEE Transactions On Neural Networks* 5(6): 924–929.
- Yao, A. C.-C. 1989. Lower bounds for algebraic computation trees with integer inputs, *Proc. 30th Annual Symposium on Foundations of Computer Science (FOCS 1989)*, 308–313.
- Yuan, P.; Lin, M.; Xiangli, K.; Zhengqing, L.; Lei, W. 2010. A Study On Echo Feature Extraction Based On The Modified Relative Spectra (RASTA) and Perception Linear Prediction (PLP) Auditory Model, *IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS) V.2*: 657–661.

- Jaromczyk, J. E. and Toussaint, G. T. 1992. Relative neighborhood graphs and their relatives, *Proceedings of the IEEE*, 80(9): 1502–1517.
- Joanes, D and Gill, C. A. 1998. Comparing measures of sample skewness and kurtosis, *Journal of the Royal Statistical Society (Series D): The Statistician* 47 (1): 183–189.
- Junqua, J. C.; Haton, J. P. 1996. *Robustness in automatic speech recognition: Fundamentals and applications*. Kluwer Academic Publishers Norwell, MA, USA. 440 p.
- Juang, B. H. and Rabiner, L. R. 2005. Automatic Speech Recognition - A Brief History of the Technology, *Elsevier Encyclopedia of Language and Linguistics, Second Edition*.
- Kerroum, M. A.; Hammouch, A.; Aboutajdine, D.; Bellaachia, A. 2008. Using the maximum Mutual Information criterion to textural Feature Selection for satellite image classification. *IEEE Symposium on Computers and Communications*. ISCC 2008, 1005–1009.
- Krishna, B.; Kaliaperumal, B. 2011. Efficient Genetic-Wrapper Algorithm Based Data Mining for Feature Subset Selection in a Power Quality Pattern Recognition Application, *The International Arab Journal of Information Technology* Vol. 8(3): 397–405.
- Kruskal, J. B. 1956. On the Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem, *In Proceedings of the American Mathematical Society* 7(1): 48–50.
- Larranaga, P.; Lozano, J. A. 2002. *Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation (Genetic Algorithms and Evolutionary Computation)*. Kluwer Academic Publishers. 382 p.
- Lebourgeois, F.; Emptoz, H. 1996. Pretopological approach for supervised learning, *Proceeding of the Thirteenth International Conference on Pattern Recognition*, 256–260.

- Leung, Y. W. and Yum, T.S. 1991. Efficient algorithms for multiple destinations routing, *IEEE International Conference on Communications V.3*: 1311–1317.
- Li, Z.; Ditang, F. 1998. The research on speech feature representation method and distance measure method, *Pattern recognition, 9th International Conference V. 1*: 631–633.
- Lileikytė, R.; Telksnys, L. 2011. Quality Estimation Methodology of Speech Recognition Features, *Electronics and Electrical Engineering*, 4 (110), 113–116.
- Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering*. 2(118). ISSN 1392–1215.
- Liu, L.; Han, B.; Ren, X.; Gao, Z. 2010. Learning algorithm for the state feedback artificial neural network, *Sixth International Conference on Natural Computation (ICNC 2010) V.1*: 357–361.
- Macia, N.; Bernado–Mansilla, E.; Orriols–Puig, A. 2008. Preliminary approach on synthetic data sets generalization based on class separability measure, *IEEE International Conference on Pattern Recognition*, 1–4.
- Makhoul, J. 1975. Linear Prediction: A Tutorial Review, *Proceedings of the IEEE* 63(4): 561–580.
- Malina, W. 1981. On an Extended Fisher Criterion for Feature Selection, *IEEE Transactions on pattern analysis and machine intelligence* 3(5): 611–614.
- Miche, Y.; Bas, P.; Lendasse, A.; Jutten, C. and Simula, O. 2007. Extracting relevant features of steganographic schemes by feature selection techniques, *In Proceedings of the 3rd Wavilla Challenge (Wacha '07)*, 1–15.
- Miyamoto, T.; Uchimura, S.; Hamamoto, Y.; Iizuka, N.; Oka, M.; Yamada-Okabe, H. 2003. Comparative study of feature selection methods on microarray data, *IEEE EMBS Asian-Pacific Conference on Biomedical Engineering*, 82–83.

- Myers, C., Rabiner, L., Rosenberg, A. 1980. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition, *IEEE Transactions on Acoustics, Speech, and Signal Processing* 28(6): 623–625.
- Myers, C.; Rabiner, L.; Rosenberg, A. 1980. An investigation of the use of dynamic time warping for word spotting and connected speech recognition, *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '80*, 173–177.
- Mollineda, R. A. and Sánchez, J.S and Sotoca J.M. 2005. Data Characterization for Effective Prototype Selection, *Pattern Recognition and Image Analysis* 3523(2005): 27–34.
- Narasimhan, G.; Zachariasen, M. and Zhu, J. 2000. Experiments with computing geometric minimum spanning trees, *In Proceedings of ALENEX'00*, 183–196.
- Oppenheim, A. V. 1970. Speech spectrograms using the fast Fourier transform, *IEEE Spectrum* 7: 57–62.
- Peng, H. C.; Long, F. and 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* 27(8): 1226–1238.
- Pinto, J. P. 2010. Multilayer Perceptron Based Hierarchical Acoustic Modeling for Automatic Speech Recognition, PhD thesis, *EPFL Switzerland*.
- Puzicha, J.; Buhmann, J.M.; Rubner, Y.; Tomasi, C.; 1999. Empirical evaluation of dissimilarity measures for color and texture. *Computer Vision, The Proceedings of the Seventh IEEE International Conference on V.2*: 1–8.
- Prim, R. C. 1957. Shortest connection networks and some generalizations, *In Bell System Technical Journal* 36: 1389–1401.
- Rabiner, L. R.; Juang, B. H. 1993. *Fundamentals of speech recognition*. Prentice-Hall, New Jersey. 507 p.

- Rabiner, L. R.; Rosenberg, A. E.; Levinson, S.E. 1978. Considerations in dynamic time warping algorithms for discrete word recognition, *IEEE transactions on Acoustics, Speech and Signal Processing* 26(6): 575–582.
- Raškinis, A.; Raškinis G.; Kazlauskienė A. 2004. Lietuvių bendrinės šneamosios kalbos garsų fondo kūrimo principai, *Lituanistica* 4 (60): 53–62.
- Raškinis, G.; Kazlauskienė, A. 2009. Phone frequency in standard Lithuanian. *Respectus philologicus* 16 (21): 169–182.
- Rosenblatt, F. 1958. The perceptron: A probabilistic model for information storage and organization in the brain, *Psychological Review* 65(6): 386–408.
- Sakoe, H.; Chiba, S. 1978. Dynamic programming algorithm optimization for spoken word recognition, *IEEE Transactions on Acoustics, Speech and Signal Processing* 26(1): 43–49.
- Sakoe, H. 1979. Two-level DP-matching - A dynamic programming-based pattern matching algorithm for connected word recognition, *IEEE Transactions on Acoustics, Speech and Signal Processing* 27(6): 588–595.
- Sandon, P. A. 1988. A local interaction heuristic for adaptive networks, *Proceedings of the IEEE International Conference on Neural Networks* V.1: 24–27.
- Schafer, R. W.; Rabiner, L. R. 1975. Digital representations of speech signals, *Proceedings of the IEEE* 63(4): 662–677.
- Sohn, S. Y. 1999. Meta analysis of classification algorithms for pattern recognition, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11): 1137–1144.
- Soryani, M.; Rafat, N. 2005. Application of Genetic Algorithms to Feature Subset Selection in a Farsi OCR, *Proceedings of the 2005 IEEE Congress on Evolutionary Computation CEC* 3: 2514–2521.
- Sotoca, J. M.; Sanchez, J. S.; Mollineda, R. A. 2005. A review of data complexity measures and their applicability to pattern classification problems,

- Actas del III Taller Nacional de Minería de Datos y Aprendizaje. TAMIDA*, 77–83.
- Sudhakar, T. D.; Srinivas, K. N. 2011. Power System Restoration Based On Kruskal's Algorithm, *1st International Conference on Electrical Energy Systems*, 281–287.
- Targamadžė, A.; Petrauskienė, R.; Rubliauskas, D. 2010. Influence of technologies on quality of distance learning, *Electronics and Electrical Engineering, Kaunas: Technologija*, 6(102): 131–134.
- Theodoridis, S.; and Koutroumbas, K. 1999. *Pattern Recognition*, by Academic Press. 625 p.
- Tierney, J. 1980. A study of LPC analysis of speech in additive noise, *IEEE Transactions on Acoustics, speech, and signal processing* 28(4): 389–397.
- Van der Merwe, C. J.; du Preez, J. A. 1991. Calculation of LPC-based cepstrum coefficients using mel-scale frequency warping, *Communications and Signal Processing, COMSIG 1991*, 17–21.
- Vandervalk, B. P.; McCarthy, E. L.; Wilkinson, M. D. 2009. Optimization of Distributed SPARQL Queries Using Edmonds' Algorithm and Prim's Algorithm, *International Conference on Computational Science and Engineering V.1*: 330–337.
- Voitovetsky, I.; Guterman, H.; Cohen, A. 1997. Unsupervised speaker classification using self-organizing maps (SOM), *Neural networks for signal processing VII, Processing of the 1997 IEEE Workshop*, 578–587.
- Wachter, M. D.; Demuynck, K.; Compernelle, D. V. 2007. Outlier Correction for Local Distance Measures in Example Based Speech Recognition, *IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP '07*, 433–436.
- Webb, A. R. 2002. *Statistical Pattern Recognition*. John Wiley, NJ, 2nd edition. 496 p.

- Widrow, G. and Hoff, M. E. 1960. Adaptive switching circuits, *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record, Part 2*, 96–104.
- Wittenburg, G.; Dziengel, N.; Wartenburger, C.; Schiller, J. 2010. A System for Distributed Event Detection in Wireless Sensor Networks, *In Proceedings of the 9th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN '10)*, 94–104.
- Xiao-hua, X. ; Ai-bing, N. ; Liang, M. 2009. A new Verification Algorithm for Minimum Spanning Tree based on Reduction and Merge Technology, *proceedings of 2009 4th International Conference on Computer Science & Education*, 469–473.
- Zuo, W.; Zhang, D.; Wang, K. 2008. On kernel difference-weighted k-nearest neighbor classification, *Pattern Analysis Applications* 11: 247–257.

Autoriaus publikacijos disertacijos tema

Straipsniai recenzuojamuose mokslo žurnaluose

Lileikytė, R.; Telksnys, L. 2011. Quality Estimation Methodology of Speech Recognition Features, *Electronics and Electrical Engineering*. 4(110), 113–116. ISSN 1392–1215.

Lileikytė, R.; Telksnys, L. 2012. Quality Measurement of Speech Recognition Features in Context of Nearest Neighbour Classifier, *Electronics and Electrical Engineering*. 2(118). ISSN 1392–1215.

Lileikytė, R.; Telksnys, L. Quality estimation of speech recognition features for Dynamic Time Warping classifier, *Information Technology and Control*. ISSN 1392–124X. (Priimta spausdinimui).

Priedai

A priedas. Eksperimentinių tyrimų fonemų kombinacijos

1A lentelėje pateiktos klasių kombinacijos numeraciją atitinkančios fonemų kombinacijos, naudotos eksperimentinių tyrimų metu (čia „ ‚ “ žymi minkštą priebalsį, „ : “ – ilgą balsį).

1A lentelė. Eksperimentinių tyrimų fonemų kombinacijos

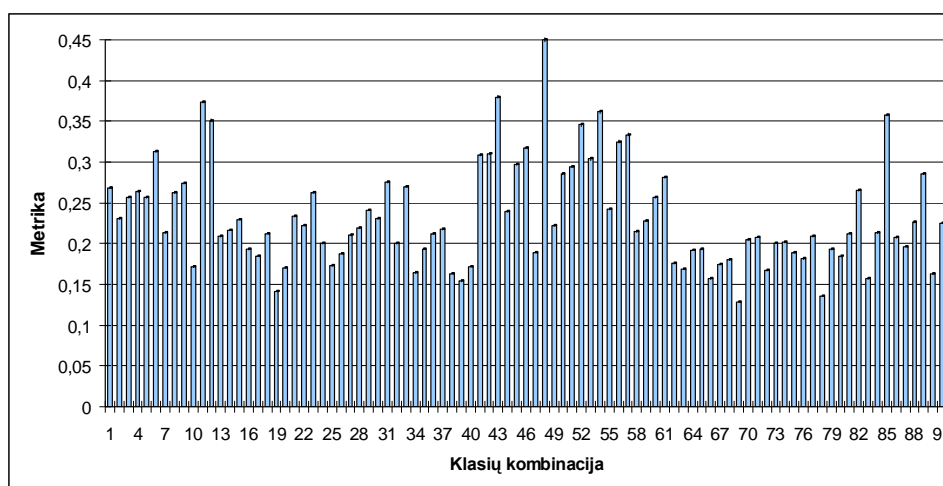
Nr.	Kombinac.	Nr.	Kombinac.	Nr.	Kombinac.	Nr.	Kombinac.
1	2	3	4	5	6	7	8
1.	<i>a-i</i>	24.	<i>m-k</i>	47.	<i>r'-o:</i>	70.	<i>s'-r'</i>
2.	<i>e-a</i>	25.	<i>m-o:</i>	48.	<i>r'-r</i>	71.	<i>s'-r</i>
3.	<i>e-i</i>	26.	<i>m-s</i>	49.	<i>r'-s</i>	72.	<i>s'-s</i>
4.	<i>j-a</i>	27.	<i>m-t'</i>	50.	<i>r'-t'</i>	73.	<i>s'-t'</i>
5.	<i>j-e</i>	28.	<i>m-t</i>	51.	<i>r'-t</i>	74.	<i>s'-t</i>
6.	<i>j-i</i>	29.	<i>n'-a</i>	52.	<i>r-a</i>	75.	<i>s-a</i>
7.	<i>j-k</i>	30.	<i>n'-e</i>	53.	<i>r-e</i>	76.	<i>s-e</i>
8.	<i>j-m</i>	31.	<i>n'-i</i>	54.	<i>r-i</i>	77.	<i>s-i</i>
9.	<i>j-n'</i>	32.	<i>n'-k</i>	55.	<i>r-k</i>	78.	<i>s-o:</i>
10.	<i>j-o:</i>	33.	<i>n'-m</i>	56.	<i>r-m</i>	79.	<i>t'-a</i>
11.	<i>j-r'</i>	34.	<i>n'-o:</i>	57.	<i>r-n'</i>	80.	<i>t'-e</i>

1A lentelės tęsinys

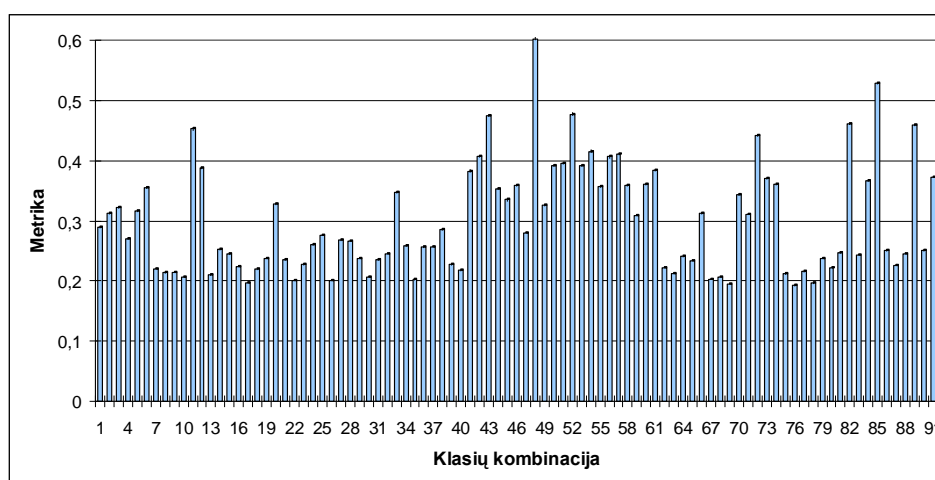
1	2	3	4	5	6	7	8
12.	<i>j-r</i>	35.	<i>n'-s</i>	58.	<i>r-o:</i>	81.	<i>t'-i</i>
13.	<i>j-s</i>	36.	<i>n'-t'</i>	59.	<i>r-s</i>	82.	<i>t'-k</i>
14.	<i>j-t'</i>	37.	<i>n'-t</i>	60.	<i>r-t'</i>	83.	<i>t'-o:</i>
15.	<i>j-t</i>	38.	<i>o:-a</i>	61.	<i>r-t</i>	84.	<i>t'-s</i>
16.	<i>k-a</i>	39.	<i>o:-e</i>	62.	<i>s'-a</i>	85.	<i>t'-t</i>
17.	<i>k-e</i>	40.	<i>o:-i</i>	63.	<i>s'-e</i>	86.	<i>t-a</i>
18.	<i>k-i</i>	41.	<i>r'-a</i>	64.	<i>s'-i</i>	87.	<i>t-e</i>
19.	<i>k-o:</i>	42.	<i>r'-e</i>	65.	<i>s'-j</i>	88.	<i>t-i</i>
20.	<i>k-s</i>	43.	<i>r'-i</i>	66.	<i>s'-k</i>	89.	<i>t-k</i>
21.	<i>m-a</i>	44.	<i>r'-k</i>	67.	<i>s'-m</i>	90.	<i>t-o:</i>
22.	<i>m-e</i>	45.	<i>r'-m</i>	68.	<i>s'-n'</i>	91.	<i>t-s</i>
23.	<i>m-i</i>	46.	<i>r'-n'</i>	69.	<i>s'-o:</i>		

B priedas. Metrikų vidurkių įverčiai

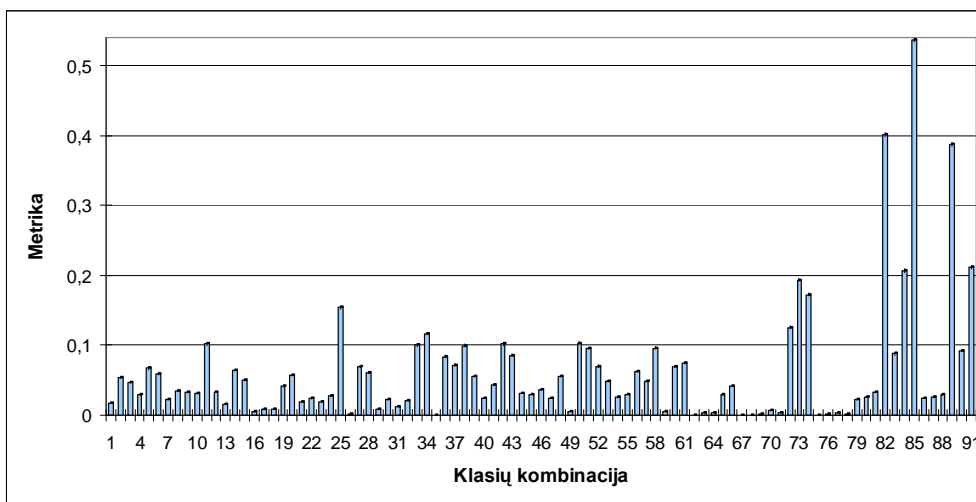
Priedo paveiksluose pateiktos eksperimentinių tyrimų metu apskaičiuotos metrikų G1, G2, G3 vidurkių įverčių reikšmės su 80 % pasikliautinumo intervalais. Apskaičiavimai atlikti 12 eilės TSPKK ir 12 eilės KK požymių sistemoms, trims duomenų rinkiniams – DM1, DM2, DM3. Priedo (1B)–(6B) paveiksluose pateiktos metrikų G1, G2, G3 vidurkių ir pasikliautinųjų intervalų reikšmės DM1 duomenų rinkiniui, (7B)–(12B) paveiksluose – DM2 rinkiniui ir (13B)–(18B) paveiksluose – DM3 rinkiniui. Pasikliautinumo intervalai itin maži, siekia vieną tūkstantą, nes vienai klasių porai atlikta metrikų apskaičiavimų tarp 10 000 pavyzdžių kombinacijų.



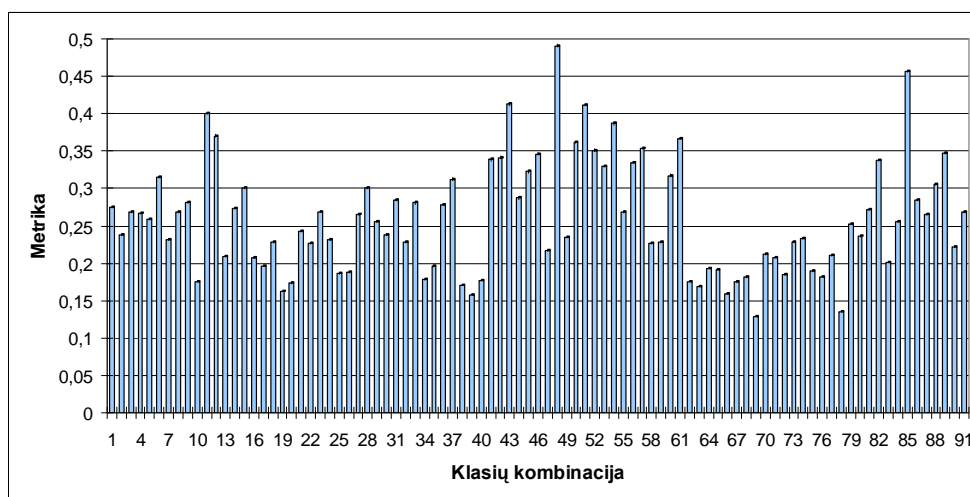
1B pav. Metrikos G1 vidurkių įverčiai (TSPKK, DM1)



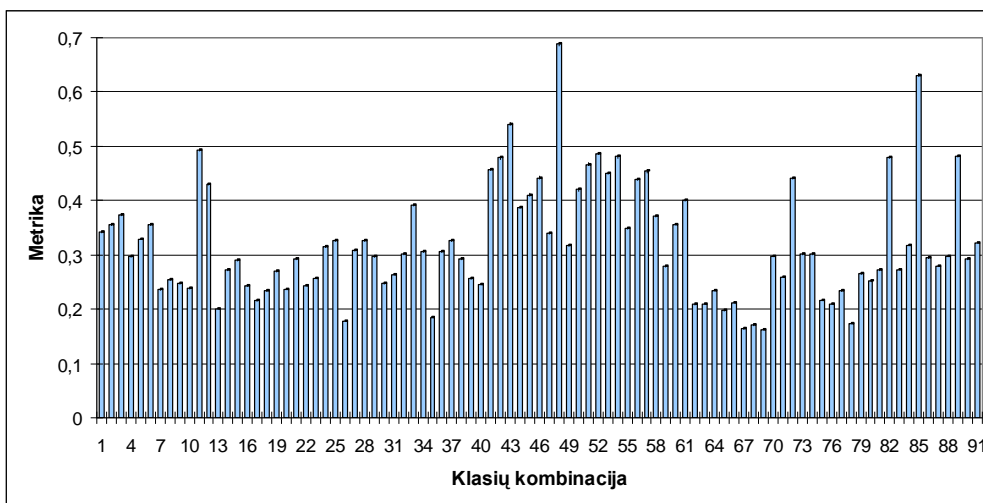
2B pav. Metrikos G2 vidurkių įverčiai (TSPKK, DM1)



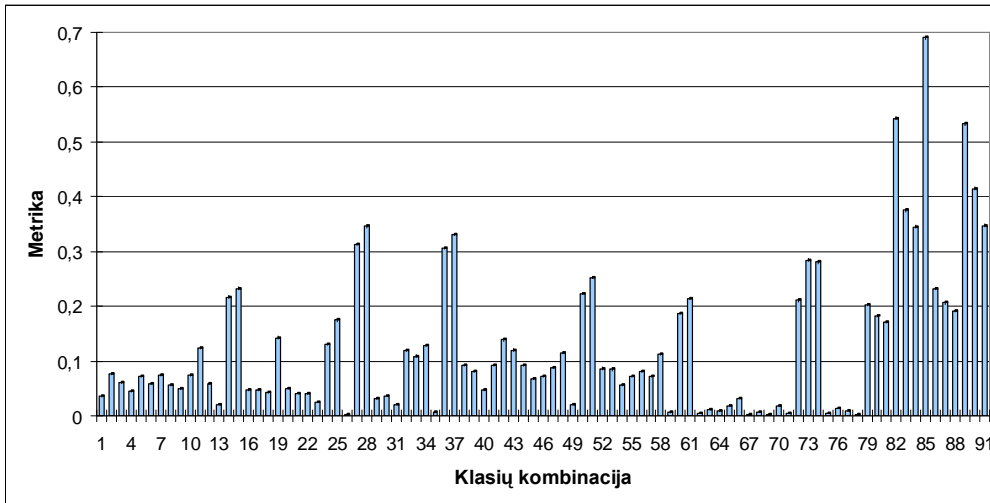
3B pav. Metrikos G3 vidurkių įverčiai (TSPKK, DM1)



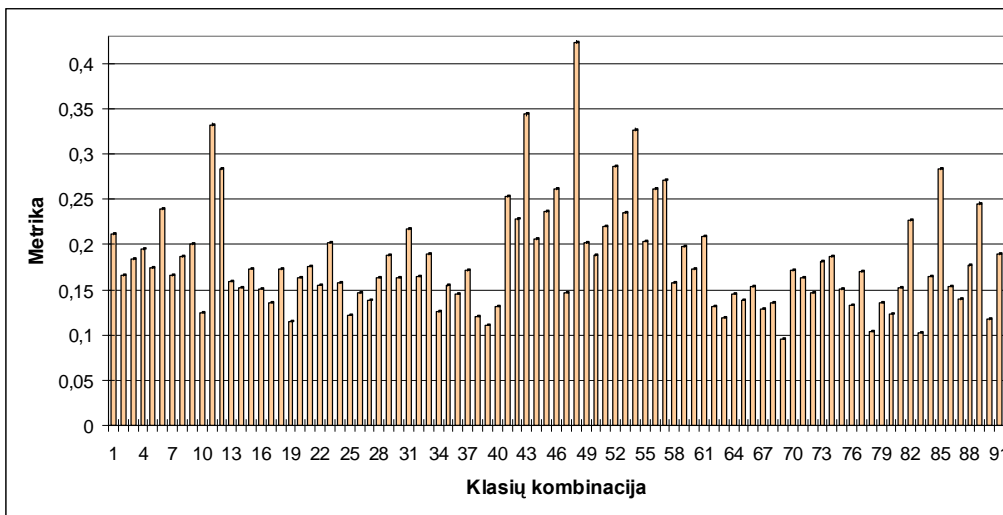
4B pav. Metrikos G1 vidurkių įverčiai (KK, DM1)



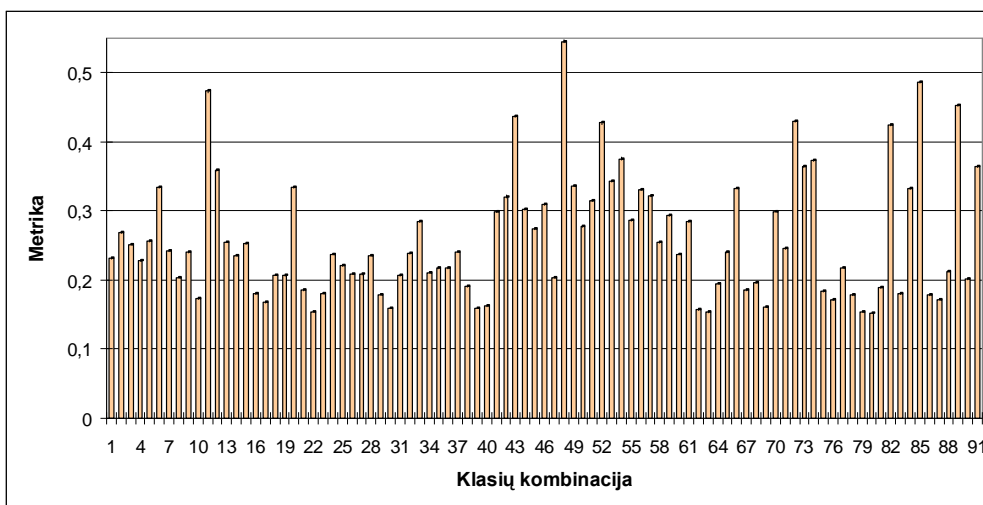
5B pav. Metrikos G2 vidurkių įverčiai (KK, DM1)



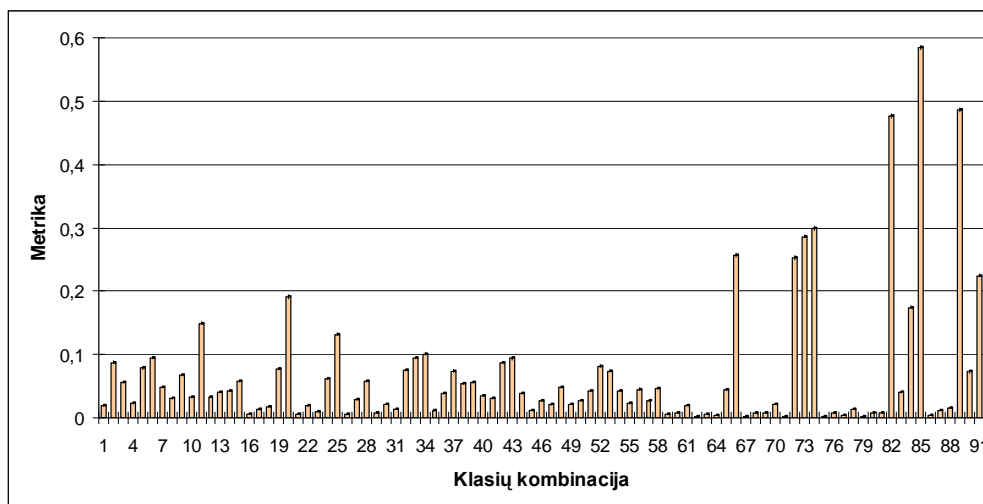
6B pav. Metrikos G3 vidurkių įverčiai (KK, DM1)



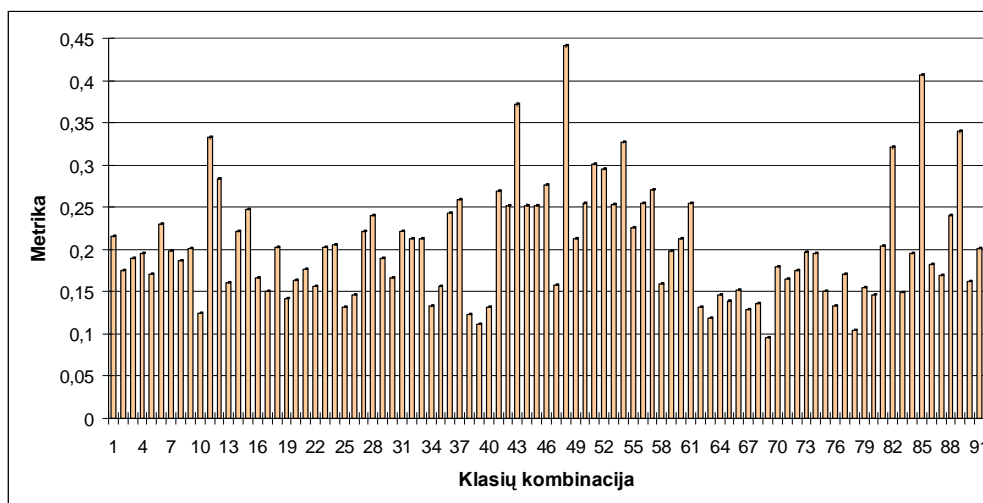
7B pav. Metrikos G1 vidurkių įverčiai (TSPKK, DM2)



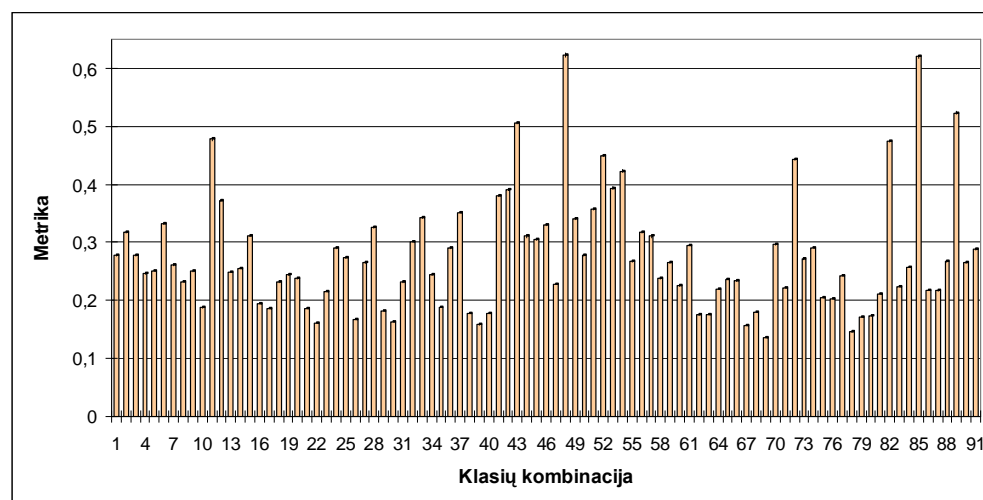
8B pav. Metrikos G2 vidurkių įverčiai (TSPKK, DM2)



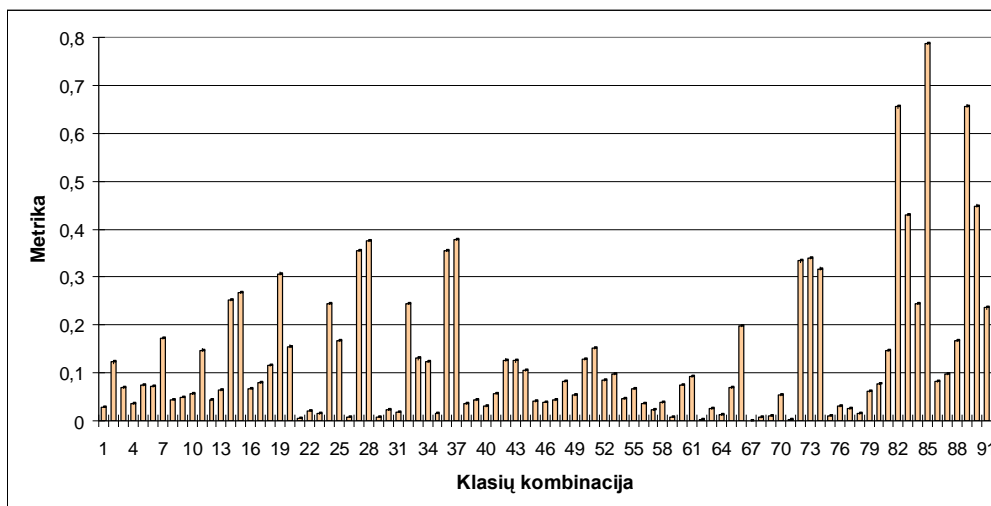
9B pav. Metrikos G3 vidurkių įverčiai (TSPKK, DM2)



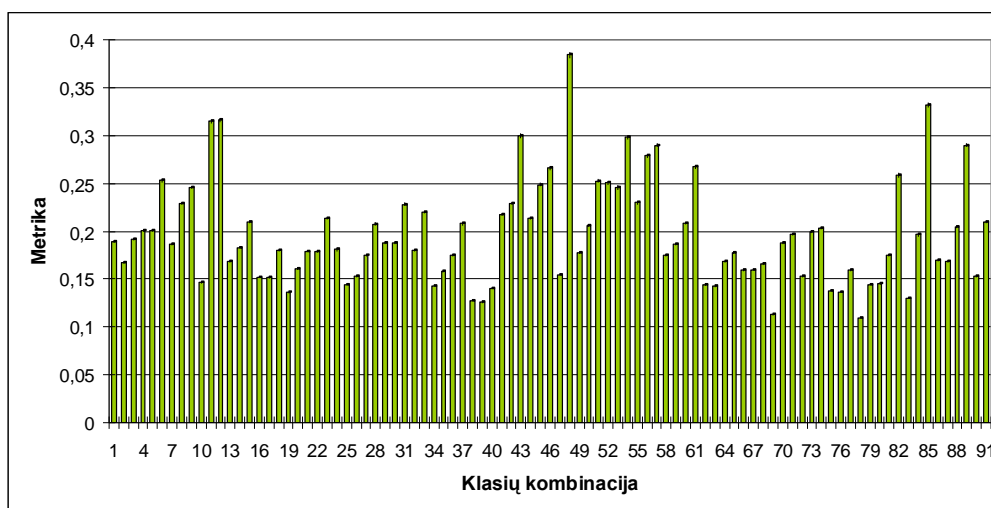
10B pav. Metrikos G1 vidurkių įverčiai (KK, DM2)



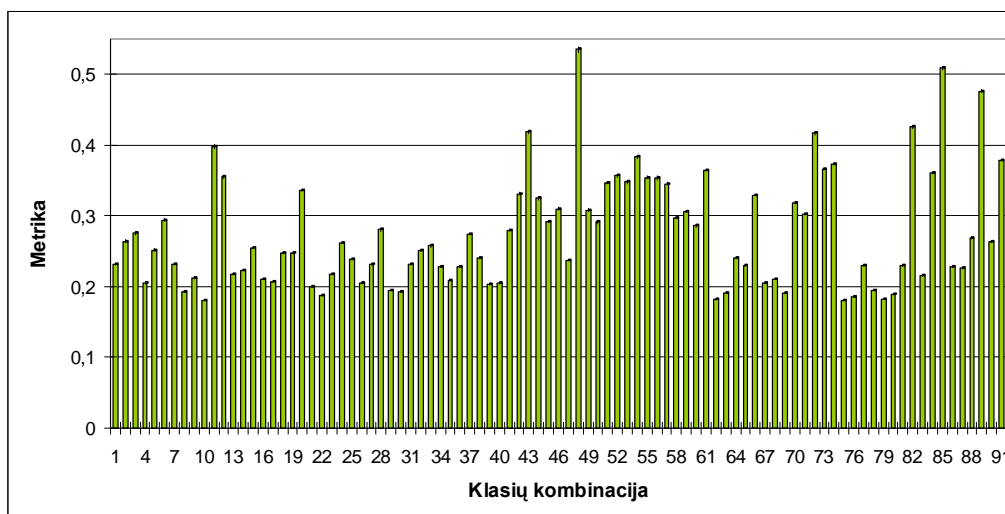
11B pav. Metrikos G2 vidurkių įverčiai (KK, DM2)



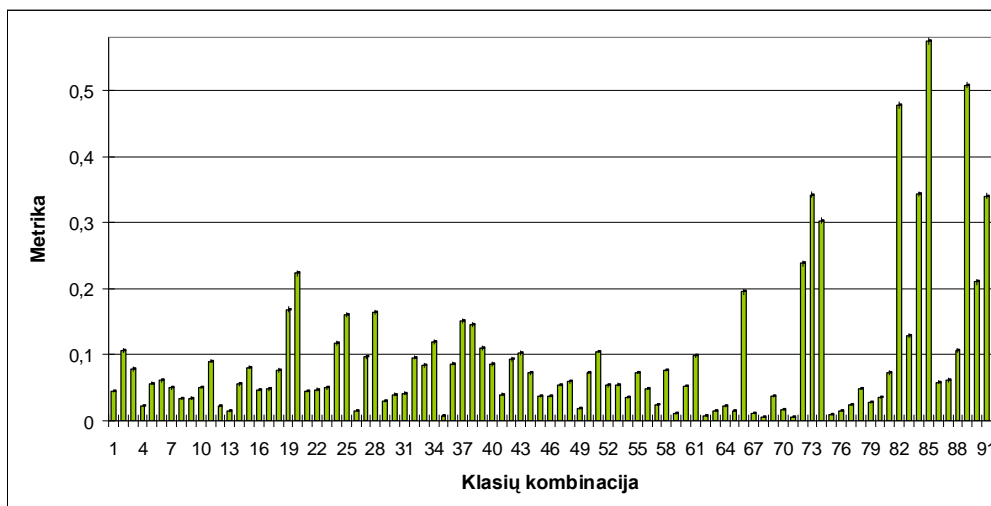
12B pav. Metrikos G3 vidurkių įverčiai (KK, DM2)



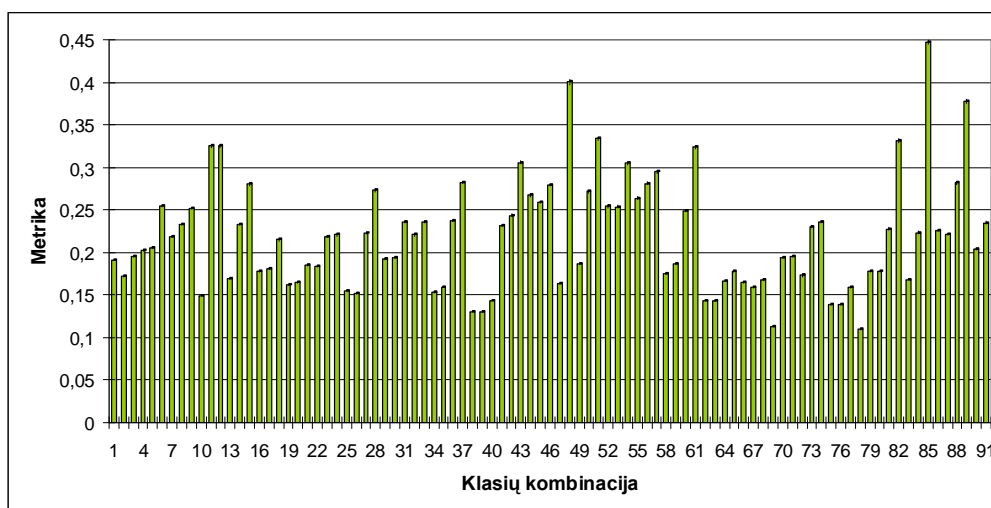
13B pav. Metrikos G1 vidurkių įverčiai (TSPKK, DM3)



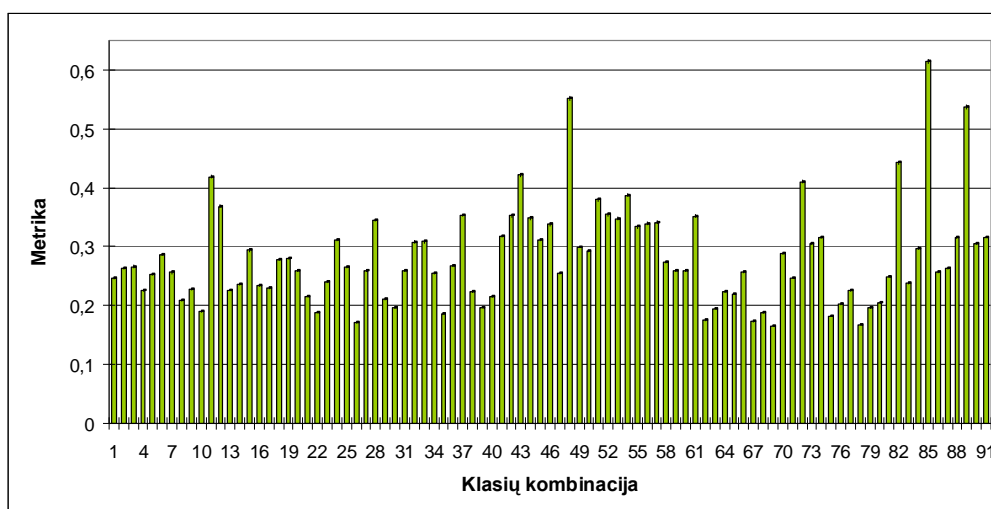
14B pav. Metrikos G2 vidurkių įverčiai (TSPKK, DM3)



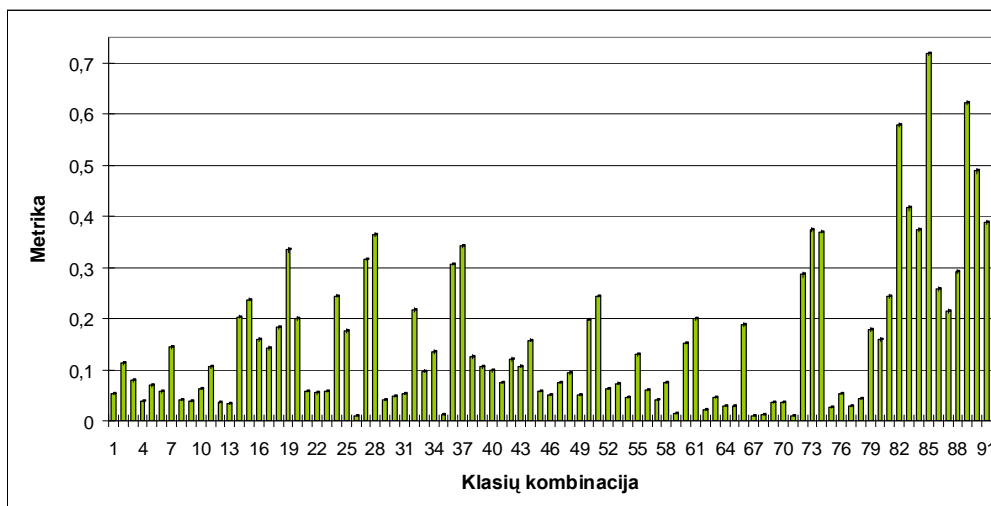
15B pav. Metrikos G3 vidurkių įverčiai (TSPKK, DM3)



16B pav. Metrikos G1 vidurkių įverčiai (KK, DM3)



17B pav. Metrikos G2 vidurkių įverčiai (KK, DM3)



18B pav. Metrikos G3 vidurkių įverčiai (KK, DM3)

Rasa LILEIKYTĖ

ŠNEKOS ATPAŽINIMO POŽYMIŲ KOKYBĖS VERTINIMAS

Daktaro disertacija

Technologijos mokslai,
Informatikos inžinerija (07T)

Rasa LILEIKYTĖ

QUALITY ESTIMATION OF SPEECH RECOGNITION FEATURES

Doctoral Dissertation

Technological Sciences,
Informatics Engineering (07T)