



Kauno technologijos universitetas

Informatikos fakultetas

Balso atpažinimo technologijų taikymas kalbos sutrikimų lavinimui

Baigiamasis magistro studijų projektas

Dainius Arnastauskas

Projekto autorius

Prof. Tomas Blažauskas

Vadovas

Kaunas, 2022



Kauno technologijos universitetas

Informatikos fakultetas

Balso atpažinimo technologijų taikymas kalbos sutrikimų lavinimui

Baigiamasis magistro studijų projektas
Programų sistemų inžinerija (6211BX011)

Dainius Arnastauskas

Projekto autorius

Prof. Tomas Blažauskas

Vadovas

Prof. Eduardas Bareiša

Recenzentas

Kaunas, 2022



Kauno technologijos universitetas

Informatikos fakultetas

Dainius Arnastauskas

Balso atpažinimo technologijų taikymas kalbos sutrikimų lavinimui

Akademinio sąžiningumo deklaracija

Patvirtinu, kad:

1. baigiamąjį projektą parengiau savarankiškai ir sąžiningai, nepažeisdama(s) kitų asmenų autoriaus ar kitų teisių, laikydamasi(s) Lietuvos Respublikos autorių teisių ir gretutinių teisių įstatymo nuostatų, Kauno technologijos universiteto (toliau – Universitetas) intelektinės nuosavybės valdymo ir perdavimo nuostatų bei Universiteto akademinės etikos kodekse nustatytų etikos reikalavimų;
2. baigiamajame projekte visi pateikti duomenys ir tyrimų rezultatai yra teisingi ir gauti teisėtai, nei viena šio projekto dalis nėra plagijuota nuo jokių spausdintinių ar elektroninių šaltinių, visos baigiamojo projekto tekste pateiktos citatos ir nuorodos yra nurodytos literatūros sąrašė;
3. įstatymų nenumatytų piniginių sumų už baigiamąjį projektą ar jo dalis niekam nesu mokėjęs (-usi);
4. suprantu, kad išaiškėjus nesąžiningumo ar kitų asmenų teisių pažeidimo faktui, man bus taikomos akademinės nuobaudos pagal Universitete galiojančią tvarką ir būsiu pašalinta(s) iš Universiteto, o baigiamasis projektas gali būti pateiktas Akademinės etikos ir procedūrų kontrolieriaus tarnybai nagrinėjant galimą akademinės etikos pažeidimą.

Dainius Arnastauskas

Patvirtinta elektroniniu būdu

Autoriaus Arnastauskas, Dainius. Balso atpažinimo technologijų taikymas kalbos sutrikimų lavinimui. Magistro studijų baigiamasis projektas vadovas Prof. Tomas Blažauskas; Kauno technologijos universitetas, informatikos fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Informatikos mokslai, programų sistemų inžinerija.

Reikšminiai žodžiai: Android, Kotlin, Python, Django, API, programėlė, balso atpažinimas, balsas į tekstą.

Kaunas, 2022. 65 p.

Santrauka

Šiame darbe aprašyta balso atpažinimo ir vertimo į tekstą technologijos. Sukurta programinė įranga naudojanti „Google Cloud“ „speech-to-text“ servisu tyrinėti žmonių kalbą. Programinės įrangos surinkti duomenys panaudoti tyrimams atlikti.

Analitinėje dalyje pristatomi ir analizuojami balso atpažinimo metodai ir įgyvendinti sprendimai, pagrindžiamas „Google Cloud“ „speech-to-text“ pasirinkimas. Projektinėje dalyje aprašoma sistemos specifikacija ir testavimo atvejai. Tyrimo dalyje aprašyti atlikti tyrimai, metodika ir techninė įranga. Eksperimentinėje dalyje pristatomi naudotojų vertinimų tyrimai ir sistemos panaudojamumo tyrimas.

Author's Arnastauskas, Dainius. Voice Recognition Technologies Application in Speech Exercise. Master's Final Degree Project supervisor Prof. Tomas Blažauskas; Informatics faculty, Kaunas University of Technology.

Study field and area (study field group): computer sciences, software systems engineering.

Keywords: Android, Kotlin, Python, Django, API, app, speech recognition, speech to text.

Kaunas, 2022. 65 p.

Summary

This paper describes the technologies of voice recognition and translation into text. Developed software using Google Cloud Text-to-Speech to explore people's language. The data collected by the software was used for research.

Analytical chapter introduces and analyzes voice recognition methods and implementation solutions based on the choice of Google Cloud speech-to-text. The design part describes the system specification and test cases. The research part describes the research, methodology and technical equipment. The experimental chapter presents research on user's evaluations and research on the system usability.

Turinys

Lentelių sąrašas	3
Paveikslų sąrašas	4
Santrumpų ir terminų sąrašas	5
1. Įžanga	6
1.1. Dokumento paskirtis	6
1.2. Darbo tikslas	6
1.3. Mokslinis naujumas ir aktualumas	6
1.4. Uždaviniai	6
2. Analitinė dalis	7
2.1. Įvadas	7
2.2. Kalbos atpažinimo sritis.....	7
2.3. Kalbos atpažinimo taikymo sritys.....	7
2.4. Balso atpažinimo metodai.....	8
2.5. Atrinkti technologiniai sprendimai tikslui pasiekti.....	9
2.6. Levenšteino atstumas	11
2.7. Rinkos analizė.....	12
2.8. Įgyvendinimo problemos	14
2.9. Apibendrinančios išvados	14
3. Projektinė dalis	15
3.1. Reikalavimų specifikacija.....	15
3.2. Architektūros specifikacija	25
4. Tyrimo dalis	40
4.1. Tyrimo aprašymas.....	40
4.2. Techninė ir programinė įranga.....	41
4.3. Tyrimų atlikimo metodika	41
5. Eksperimentinė dalis	42
5.1. Eksperimentinių tyrimų tikslas	42
5.2. Daromų klaidų tyrimas	42
5.3. Aptiktų žodžių patikimumo tyrimas	44
5.4. Užsieniečių respondentų tyrimas	46
5.5. Sistemos panaudojamumo tyrimas	47
5.6. Eksperimentinio tyrimo išvados	50
6. Išvados	52
7. Literatūra	53
8. Priedai	55

Lentelių sąrašas

1 lentelė. Nagrinėjamų technologijų palyginimas	11
2 lentelė. Konkurentų palyginimas	13
3 lentelė. Valdyti visas užduotis panaudojimo atvejo specifikacija.	19
4 lentelė. Valdyti savo užduotis panaudojimo atvejo specifikacija.	19
5 lentelė. Valdyti naudotojus panaudojimo atvejo specifikacija.	20
6 lentelė. Peržiūrėti savo naudotojus panaudojimo atvejo specifikacija.....	20
7 lentelė. Peržiūrėti sesiją panaudojimo atvejo specifikacija.....	21
8 lentelė. Prisijungti panaudojimo atvejo specifikacija.	21
9 lentelė. Spręsti užduotis panaudojimo atvejo specifikacija.	21
10 lentelė. Matyti statistiką panaudojimo atvejo specifikacija.	22
11 lentelė Tyrimo įranga.....	41

Paveikslų sąrašas

1 pav. Kalbos sutrikimų lavinimo sistemos konteksto diagrama.....	17
2 pav. Panaudojimo atvejų diagrama	18
3 pav. Klasių diagrama	23
4 pav. Klasių diagrama	24
5 pav. Panaudojimo atvejų diagrama	26
6 pav. Programėlės paketų diagrama	28
7 pav. Saityno sąsajos paketų diagrama.....	29
8 pav. Programėlės klasių diagrama	30
9 pav. Saityno sąsajos klasių diagrama.....	31
10 pav. Matyti statistiką sekų diagrama.....	32
11 pav. Valdyti savo naudotojus sekų diagrama.....	33
12 pav. Prisijungti sekų diagrama	34
13 pav. Spręsti užduotis sekų diagrama	35
14 pav. Matyti statistiką komunikavimo diagrama.....	35
15 pav. Spręsti užduotis veiklos diagrama.....	36
16 pav. Mikrofono būsenų diagrama	37
17 pav. Diegimo diagrama.....	38
18 pav. Esybių ryšių pavyzdys.....	38
19 pav. Originalus SUS klausimynas (20).....	40
20 pav. Naudotojų grupių daromų klaidų kiekių vidurkiai.....	42
21 pav. Naudotojų grupių konkrečių žodžių klaidų kiekiai	43
22 pav. Užduočių patikimumo įvertinimo pasiskirstymas pagal raides	45
23 pav. Naudotojų grupių daromų klaidų kiekių vidurkiai.....	46
24 pav. SUS apklausos atsakymų lentelė.....	47
25 pav. Išmokstamumo ir panaudojamumo palyginimas	48
26 pav. SUS apklausos sveikų respondentų rezultatai.....	49
27 pav. SUS apklausos sutrikimų turinčių respondentų rezultatai	50

Santrumpų ir terminų sąrašas

Android – atviro kodo operacinė sistema, daugiausia naudojama išmaniuosiuose telefonuose.

Google Cloud speech-to-text - Google sukurtas balso vertimo į tekstą paslauga..

AWS S3 (Amazon Web Services Simple Storage Service) – Amazon kompanijos paslauga, kuri teikia objektų saugyklos paslaugas per interneto sąsają.

API (angl. Application Programming Interface) – tai sąsaja, kurią suteikia kompiuterinė sistema, biblioteka ar programa tam, kad programuotojas per kitą programą galėtų pasiekti jos funkcionalumą ar apsikeistų su ja duomenimis.

UI (angl. user interface) – visuma aparatinių ir programinių priemonių, sudarančių kompiuterio naudotojui patogias sąlygas valdyti operacinę sistemą ir taikomąsias programas.

1. Įžanga

1.1. Dokumento paskirtis

Šiame dokumente pristatomas magistro studijų metu sukurtas projektas ir su juo atlikti sistemos tyrimai.

1.2. Darbo tikslas

Naudojantis balso atpažinimo technologijomis, nustatyti kalbos raiškos skirtumus tarp sveikų ir sutrikimų turinčių žmonių, siekiant aptikti sutrikimus ir suteikti naudotojui reabilitaciją.

1.3. Mokslinis naujumas ir aktualumas

Šio darbo mokslinis naujumas ir aktualumas:

- Sukurta programinė įranga naudojantis „Google Cloud“ „speech-to-text“ paslauga vertinti naudotojo pasakytus tekstus ir fiksuoti įvairius parametrus.
- Atliktas „Google Cloud“ „speech-to-text“ paslaugos patikimumo vertinimo tyrimas su sveikais ir sutrikimų turinčiais žmonėmis,

1.4. Uždaviniai

Darbo tikslui pasiekti buvo suformuluoti tokie uždaviniai:

- Atlikti kalbos atpažinimo technologinių sprendimų analizę.
- Sukurti programinę įrangą galinčią vertinti naudotojų pasakytus tekstus.
- Atlikti sukurtos sistemos panaudojamumo tyrimus.

2. Analitinė dalis

2.1. Įvadas

Skyriaus paskirtis pasirengti projekto reikalavimų specifikavimui, projektavimui, susipažinti su taikymo sritimi, pasauliniais pasiekimais taikomojoje srityje.

Kalbos reabilitacijos programinė įranga dažniausiai yra skirta žmonėms su klausos sutrikimais, tobulinti jų kalbėjimo gebėjimus. Pradedama nuo paprastų raidžių garsų toliau skiemenų ir pereinant prie sudėtingų spontaniškų žodžių ir frazių. Tokia reabilitacija yra atliekama su terapeuto pagalba, tačiau programinė įranga leistų savarankišką mokymąsi ir padėtų žmonėms, kurie negali susimokėti už tokias sesijas su terapeutu. Mokantis savarankiškai nereikėtų kontaktuoti su žmonėmis to pasėkoje sumažėtų rizika užsikrėsti virusu.

Kuriama programėlė pateiks užduotis su atitinkamais pratimais ir galimais pratimo atsakymais. Naudotojo pratimo teisingumas bus vertinamas pagal balso atpažinimo patikimumo procentą. Patikimumo procentas yra lygus gaunamų taškų skaičiui už pratimą. Naudotojas galės pažiūrėti savo statistiką, ten matyti kiek kartų bandė užduotį ir jos taškų vidurkį. Administratoriai ir logopedai galės sukurti užduotis ir valdyti naudotojus.

2.2. Kalbos atpažinimo sritis

Kalbos atpažinimas informatikos pasaulyje yra žinomas kaip automatinis kalbos atpažinimas (ASR), kompiuterio kalbos atpažinimas arba kalba į tekstą (STT). Jis yra informatikos ir kompiuterinės lingvistikos pogrupis, kuriame kuriama naujausi algoritmai ir technologijos išversti sakinę kalbą į tekstą. Ši sritis apima žinias ir mokslinius tyrimus kompiuterių mokslo, kalbotyros ir kompiuterių inžinerijos srityse.

Kalbos atpažinimo technologijos sparčiai patobulėjo su keliomis naujovių bangomis. Gilaus mašininio mokymosi, didžiųjų duomenų pasiekimai buvo pritaikyti tobulinant kalbos atpažinimą. Apie tai byloja šioje srityje paskelbtų akademinų darbų gausa ir kompanijos, kurios pritaikė įvairius gilau mokymosi metodus diegiant kalbos atpažinimo sistemas.

2.3. Kalbos atpažinimo taikymo sritys

2.3.1. Kasdieniniame gyvenime ir ugdyme

Kalbos atpažinimas galėtų būti panaudotas mokantis kalbą. Pritaikant šias technologijas būtų galima išmokti taisyklingo tarimo ir lavinti savo kalbėjimo įgūdžius. (1)

Kalbos atpažinimas gali leisti mokiniams, turintiems sutrikimų tapti geresniais rašytojais. Tiesiog kalbėdami garsiai, jie rašytų tekstą, nereikėtų rūpintis rašyba, skyryba ir kita rašymo mechanika. (2)

2.3.2. Žmonėms su negalia

Balso atpažinimo technologija leidžia „laisvomis rankomis“ kontroliuoti įvairius įtaisus. Tai ypatingai svarbu žmonėms turintiems sutrikimų. Žmonės su negalia galėtų pasinaudoti programomis sukurtomis pasinaudojant balso atpažinimo technologijas.

Kurtiesiems ar neprigirdintiems asmenims balso atpažinimo programinė įranga galėtų tiesiogiai generuoti pokalbių tekstą. Pavyzdžiui: konferencijos, paskaitos, religinės apeigos. (3)

Akliesiems balso atpažinimo technologijos galėtų padėti kontroliuoti elektrinius įrenginius. Ši technologija palengvintų ir padarytų gyvenimą malonesnį. (4)

2.3.3. Kitos taikymo sritys

Balso atpažinimo technologijos galėtų būti pritaikytos:

- Kuriant automatinius subtitrus.
- Kuriant automatinius vertėjus.
- Namų atomizavime.
- Robotikoje.
- Ir kt.

2.4. Balso atpažinimo metodai

Akustikos ir kalbos modeliavimai yra svarbūs šiuolaikinių statistiškai pagrįstų kalbos atpažinimo algoritmų dalys. Paslėpti Markovo modeliai (HMM) yra plačiai naudojami daugelyje sistemų. Kalbos modeliavimas taip pat naudojamas daugelyje kitų natūralios kalbos apdorojimo programų, tokių kaip dokumentų klasifikavimas ar statistinis mašininis vertimas. Toliau apžvelgsiu algoritmus ir modelius naudojamus kalbos atpažinime.

2.4.1. Paslėpti Markovo modeliai (HMM)

Paslėptieji Markovo modeliai yra vienas iš populiariausių mašininio mokymosi ir statistikos, kurioje modeliuojamos sekos, metodų. Statistinis Markovo modelis naudojamas kalbos atpažinime, kuriame laikoma, kad modeliuojama sistema yra Markovo procesas - vadinkime jį X - su nepastebimomis („paslėptomis“) būsenomis. HMM daro prielaidą, kad yra dar vienas procesas Y , kurio elgesys „priklauso“ nuo X . Tikslas yra sužinoti apie X stebint Y . HMM numato, kad kiekvienam laiko egzemplioriui n_0 sąlyginis Y_{n_0} tikimybių pasiskirstymas atsižvelgiant į istoriją $\{X_n = x_n\}_{n <= n_0}$ neturi priklausyti nuo $\{x_n\}_{n <= n_0}$. (5) HMM yra toks populiarus, kalbos atpažinime, nes jis gali būti automatiškai apmokomas, bei paprastas.

2.4.2. Dinaminis laiko skalės kraipymas (DTW)

Šis metodas yra skirtas laiko sekų, kurios gali skirtis greičiu panašumų matavimui. Viena šio algoritmo panaudojimo sritis yra automatinis kalbos atpažinimas. Kitos taikymo sritys galėtų būti kalbėtojo atpažinimas, parašo atpažinimas, įvairių judesių panašumų aptikimas. Bet kokius duomenis, kuriuos galima paversti linijiniais, galima analizuoti naudojant DTW tačiau dabar jį iš esmės pakeitė sėkmingesnis HMM metodas.

2.4.3. Neuroniniai tinklai

Nuo devintojo dešimtmečio neuroniniai tinklai buvo naudojami daugeliu kalbos atpažinimo aspektų, tokių kaip fonemų klasifikacija, fonemų klasifikacija naudojant daugiafunkcinius evoliucinius algoritmus, izoliuoto žodžio atpažinimas, garso ir vaizdo kalbos atpažinimas, garso ir vaizdo garsiakalbių atpažinimas ir garsiakalbių pritaikymas. (6) Neuroniniai tinklai daro mažiau prielaidų apie statistines ypatybes nei HMM bei turi keletą savybių, kurios paverčia neuroninius tinklus patraukliu kalbos atpažinimo modeliu.

2.4.4. Gilūs grįžtamieji ir pasikartojantys neuroniniai tinklai

Gilus grįžtamasis neuroninis tinklas (DNN) yra dirbtinis neuroninis tinklas, turintis kelis paslėptus vienetų sluoksnius tarp įvesties ir išvesties sluoksnių. (7) Panašiai kaip sekliuose neuroniniuose tinkluose, DNN gali modeliuoti sudėtingus netiesinius ryšius. DNN architektūros sukuria kompozicinius modelius, kur papildomi sluoksniai suteikia galimybę komponuoti žemesnių sluoksnių ypatybes, suteikdami didžiulį mokymosi pajėgumą ir tokiu būdu potencialą modeliuoti sudėtingus kalbos duomenų modelius. (8) Galimybė modeliuoti sudėtingus duomenų modelius šį metodą padaro daug kartų pranašesni negu paprastus neuroninius tinklus.

2.4.5. Automatinis kalbos atpažinimas nuo galo iki galo (End-to-end)

Tradiciniai fonetiniai (t. y. visi HMM pagrįsti modeliai) metodai reikalauja atskirų komponentų ir mokymo tarimui, akustikai ir kalbos modeliui. Modeliai nuo galo iki galo kartu mokosi visų kalbos atpažinimo komponentų. Tai yra supaprastina mokymo ir diegimo procesus. Pavyzdžiui, visoms HMM pagrįstoms sistemoms reikalinga n kalbos modelių, o tipiškas n kalbos modelis atmintyje dažnai užima keletą gigabaitų, todėl jų diegimas mobiliuosiuose įrenginiuose yra nepraktiškas. Išspręsti šiai problemai šiuolaikinės komercinės kompanijos kaip „Google“ ir „Apple“ ASR sistemas diegia „debesyje“ ir reikalauja tinklo ryšio. Didelio masto CNN-RNN-CTC architektūrą 2018 m. pristatė „Google DeepMind“, kuri pasiekė 6 kartus geresnius rezultatus nei žmonių ekspertai. (9)

2.5. Atrinkti technologiniai sprendimai tikslui pasiekti

2.5.1. Google „Speech-to-text“

Google „Speech-toText“ API pasižymi įspūdingu tikslumu. Viena iš API įspūdingo tikslumo priežasčių yra galimybė pasirinkti tarp skirtingų mašininio mokymosi modelių, atsižvelgiant į tai, kam naudojama programa. Tai daro Google „Speech-to-Text“ paslaugą tinkamu sprendimu įvairioms programoms, ne vien trumpai žiniatinklio paieškai.

„Speech-to-Text“ API taip pat turi įspūdingą išplėstinių skyrybos parinkčių nustatymą. Tai naudojama kurti naudingesnėms transkripcijoms, gauti mažiau praleistų sakinių ar skyrybos klaidų.

Google „Speech-to-Text“ API nėra nemokamas. Kalbos atpažinimo funkcija yra nemokama tik 60 minučių kas mėnesį. Viršijus minučių limitą įsigalioja 0,006 USD kaina už 15 sekundžių. (10)

Pliusai:

- Atpažįsta daugiau nei 120 kalbų.
- Keli mašininio mokymosi modeliai, skirti didesniai tikslumui.
- Automatinis kalbos atpažinimas.
- Teksto transkripcija.
- Tikrinio daiktavardžio atpažinimas.
- Duomenų privatumas.
- Foninio garso pašalinimas iš telefono skambučių ir vaizdo įrašų.

Minusai:

- Kainuoja pinigus.
- Ribotas susikurtas žodynas.

2.5.2. Microsoft pažintinės paslaugos

Pagrindinis dalykas, skiriantis „Microsoft Cognitive Services“ „Speech-to-Text API“, yra „Speaker Recognition“ funkcija. Tai yra garso programinės įrangos, tokios kaip veido atpažinimas, klausomoji versija. Tai kaip tinklainės skenavimas naudotojo balso garsui. Tai daroma nepaprastai lengvai skirtingo lygio naudotojams.

Ta pati balso atpažinimo galimybė leidžia programinei įrangai prisitaikyti prie konkretaus naudotojo kalbos stilių ir modelių. Ši paslauga taip pat siūlo daugiau pasirinktinio žodyno variantų nei „Google“, ir pateikia tai kaip papildomą naudą. (11)

Pliusai:

- Didesnis duomenų saugumas naudojant balso atpažinimo algoritmus.
- Transkripcija realiuoju laiku.
- Vertimas realiuoju laiku.
- Pritaikomas žodynas.
- Teksto į kalbą galimybės natūraliems kalbos modeliams.

Minusai:

- Įtaisyti apribojimai dėl API, sukurto bendro pobūdžio tikslams.
- Naudoja mikroservisus, kurie gali būti naudingi sprendžiant atskiras problemas, tačiau netinka didesnių problemų sprendimui.

2.5.3. IBM Watson

„IBM Watson Speech to Text“ „API“ remiasi hipotezių generavimu ir vertinimu rengdama atsakymą. Paslauga taip pat gali atskirti kelis kalbėtojus, todėl jis tinka daugumai transkripcijos užduočių. Yra galimybės netgi nustatyti daugybę filtrų, pašalinti nešvarumus, pridėti pasitikėjimo žodžių ir formatavimo parametrų, kad kalbos į tekstą veikimas būtų tikslesnis. (12)

Pliusai:

- Apdoroja nestruktūrizuotus duomenis.
- Pagerina produktyvumą teikiant svarbius duomenis.
- Gali apdoroti didelius duomenų kiekius.
- Lengva įsidiesti ir pradėti naudotis.

Minusai:

- Tiesiogiai nepalaiko struktūrinių duomenų.
- Reikalauja priežiūros.
- Palaiko tik ribotą kalbų skaičių.
- Užima laiko visiškam įgyvendinimui.

2.5.4. Technologijos pasirinkimo pagrindimas

Apibendrintas technologijų palyginimas pateiktas 1 lentelėje. Žalia spalva pažymėti teigiami dalykai, geltona ne esminiai ir raudona spalva neigiami.

1 lentelė. Nagrinėjamų technologijų palyginimas

	Google	Microsoft	IBM
Mokamas	Yra limitas	Yra limitas	Yra limitas
Palaiko lietuvių kalbą	Taip	Ne	Ne
Lengva pritaikyti programėlėse	Taip	Iš dalies	Ne
Prisitaikantys metodai	Taip	Ne	Ne
Reikalauja papildomo laiko įgyvendinimui	Ne	Ne	Taip

Iš jos matome, kad „Google Cloud“ „Speech-to-Text“ paslauga labiausiai atitinka išsikeltą tikslo poreikius. Paslaugos nemokamo limitą turėtų užtekti tyrimui įvykdyti.

2.6. Levenšteino atstumas

Levenšteino atstumas yra skirtas paskaičiuoti kiek mažiausiai pakeitimų reikės padaryti, kad viena simbolių seka atitiktų kitą. Šis dviejų aibių atstumas nustato kiek raidžių skiriasi du žodžiai. Pavyzdžiui atstumas tarp **ratas** ir **batas** yra 1, kadangi reikia žodžio **ratas** raidę pakeisti į **b** raidę. Levenšteino atstumas tarp **a** ir **b** aibių yra skaičiuojamas $lev_{a,b}(|a|, |b|)$, kur

$$lev_{a,b}(i, j) = \begin{cases} \max(i, j) & \text{if } \min(i, j) = 0, \\ \min \begin{cases} lev_{a,b}(i-1, j) + 1 \\ lev_{a,b}(i, j-1) + 1 \\ lev_{a,b}(i-1, j-1) + 1_{(a_i \neq b_j)} \end{cases} & \text{kitu atveju kur kai } a_i \neq b_j \end{cases}$$

indikacijos funkcija yra padidinama 1, o kitu atveju, kai $a_i = b_j$ indikacijos funkcija būna lygi 0 ir $lev_{a,b}(i, j)$ yra atstumas tarp pirmojo i elemento ir aibės a ir pirmojo j elemento iš b aibės.

Šis algoritmas bus naudingas lyginant žmonių ištartus žodžius su žodžiais, kuriuos reikėjo ištarti. Pasirinkus tinkamą lūžio tašką galima užskaityti panašius atsakymus.

2.7. Rinkos analizė

2.7.1. „English speak practice“

Programėlė pateikia kasdienius anglų pokalbius, kurie padeda pagerinti anglų kalbos klausymo ir kalbėjimo įgūdžius. Joje yra labai daug interaktyvių praktinių įrankių, o tai garantuoja, kad nebus nuobodu mokytis naudojantis šia programėle. Pokalbiai surinkti iš įprasto kasdienio gyvenimo, jie gali vykti mokykloje, ligoninėje, maisto parduotuvėje arba autobusų stotyje, ir t.t. Surinkti pokalbiai yra išversti į 17 kalbų. Pokalbiai yra automatiškai sinchronizuoti su tekstu todėl iškart galima matyti ir subtitrus. Programėlės tikslas yra sukurti pamokas, kurios tinka visiems, įskaitant pradedančiuosius. (13) Programėlėje trūksta paprastų tarimo užduočių ir atgalinio ryšio su naudotoju. Užduotys skirtos tik praktikuotis, bet pats negali žinoti ar gerai atlieki užduotis. Kuriamoje programėlėje planuojama įgyvendinti paprastas tarimo ir sudėtingas užduotis, o užduoties įvertinimui panaudoti balso atpažinimo technologijas.

2.7.2. „Ello english learning“

Tai unikali anglų kalbos mokymosi Android programėlė, teikianti klausymo ir skaitymo užduotis. Tikri žmonės, kalba įdomiomis temomis su autentiškais akcentais iš viso pasaulio. Programėlėje yra daugiau negu 1500 dialogų, kuriuos galima skaityti, klausyti ir mokytis. Kiekviename dialoge yra žodyno skyrius su garso komentarais. Taip pat galima žiūrėti vaizdo įrašus. Galima mokytis neprisijungus: programa saugo pamokų failus, galima atsisiuntti garso ir vaizdo failus. (14) Programėlė turi daug funkcionalumo: balso atpažinimas, automatinis vertimas, atkūrimas užrakintame ekrane ir kt. Pritaikius balso atpažinimo technologijas gaunamas užduoties įvertinimas iškart ją užbaigus. Šios funkcijos pagalba gaunamas atgalinis ryšys, to siekiama ir planuojamoje kurti programėlėje.

2.7.3. „Tell a story“

Mokomasis užduočių žaidimas kalbos ir loginio mąstymo lavinimui. Kiekvienos užduoties metu yra pristatoma paveikslukų seka, iš kurių reikia sukurti istoriją. Lavinamojo žaidimo pradžioje paveikslukai yra sumaišyti. Vaikas turi paveikslukus sudėlioti teisinga seka ir papasakoti, ką šie paveikslukai rodo. Žaidimas efektyviai lavina kalbėjimą užbaigtais ir prasmingais sakiniais, nes vaikas, remiantis pastebėtais faktais, prieina prie loginių išvadų. Kad paveikslukus sudėti teisinga seka, visų pirma vaikai turi sukurti priežasties, laiko, erdvės ir kitus loginius ryšius tarp paveikslukų bei pademonstruoti gana aukšto lygio abstrakciją. Paveikslukų nagrinėjimas lavina dėmesį ir koncentraciją. (15) Pasirinktas įdomus ir įtraukiantis užduočių pakeikimo būdas, tačiau nėra pritaikyta jokių technologijų papasakotos istorijos įvertinimui. Svarbu sudėlioti paveikslėlius teisinga tvarka ir gali pereiti prie kitos užduoties. Pritaikius balso atpažinimo technologijas būtų galima atpažinti ar panaudoti raktiniai žodžiai pasakojant istoriją.

2.7.4. Logopediniai pratimai

Internetinė svetainė orientuota į mokymąsi tarti atskirus garsus. Iš pradžių mokomasi taisyklingai ištarti atskirą garsą, vėliau garsas įtvirtinamas skiemenyse, žodžiuose ir sakiniuose. Paskutinis etapas – įtvirtinti garsą kasdienėje kalboje ir išmokyti jį atskirti nuo panašiai skambančių garsų. Kiekvieno garso mokymui yra tinkamą liežuvio ir lūpų mankštelę. (16) Svetainėje pateikiami pratimai ir užduotys, tačiau nėra jokio kalbos atpažinimo ir užduočių įvertinimo. Užduotys pateiktos lietuvių kalba. Svetainės tikslas padėti kalbėjimo ir kalbos sutrikimų turintiems vaikams.

2.7.5. Tyrinėtų programų apibendrinimas

Apibendrintas tyrinėtų programų palyginimas pateiktas 2 lentelėje. Žalia spalva pažymėti teigiami dalykai, geltona ne esminiai ir raudona spalva neigiami.

2 lentelė. Konkurentų palyginimas

	Kuriama programinė įranga	<i>English Speak Practice</i>	<i>Ello English Learning</i>	<i>Tell story a</i>	Logopediniai pratimai
Yra mokymosi užduotys	Taip	Taip	Taip	Taip	Taip
Galima matyti visus užduočių įrašus	Ne	Taip	Taip	Taip	Taip
Gaunami rezultatai kaip gerai įvykdyta užduotis	Taip	Ne	Taip	Ne	Ne
Galima pasiklausti teisingo tarimo	Taip	Taip	Taip	Ne	Taip
Parsisiunčiamas užduočių turinys	Ne	Taip	Taip	Ne	Ne
Yra paprastų tarimo užduočių	Taip	Ne	Taip	Ne	Taip

Tyrinėjant rinkoje esančius sprendimus, buvo apžvelgta kalbos mokymosi ir istorijų pasakojimo programos, bei logopedinių pratimų internetinis puslapis. „English speak practice“ ir „Elio english learning“ yra programėlės, kuriose galima išmokyti anglų kalbos. Programėlių

tikslas tas pats, tačiau tarpusavyje konkurentai turi ir skirtumų. „English Speak Practice“ orientuotas į šiek tiek jau suprantančius anglų kalbą žmones ir žinančius tarimą. „Ello English Learning“ pateikia galimybes pradėti mokytis nuo paprastų dalykų kaip pradedančiajam. Kuriamoje sistemoje bus mokomasi parastų dalykų: raidžių tarimo skiemenų ir sakinių. Kaip ir „Ellio“ sistemoje planuojama pasitelkti balso atpažinimo technologijas ir įvykdžius užduotį informuoti apie jos rezultatus. „Tell a story“ programėlė orientuota į saviugdą. Programėlės užduotys yra sudėti paveikslėlius tinkama tvarka ir papasakoti istoriją. Iš to negaunamas joks grįžtamasis ryšis. Logopedinių pratimų internetinėje svetainėje pateikiamos kalbėjimo užduotys. Jos skirtos lavinti kalbą, tarti atitinkamus garsus, tačiau naudotojas negauna jokio grįžtamojo ryšio apie vykdomas užduotis. Kuriamą programėlę neturės viso funkcionalumo, kurį turi konkurentai, tačiau jo nereikia dėl sistemos specifikos.

2.8. Įgyvendinimo problemos

Planuojama kurti programėlę naudosis Google debesų „Speech-to-text“ paslauga, už kurios veikimą atsako Google. Programėlėje gali kilti nenumatyta situacija, kai ši paslauga dėl kažkokių tai priežasčių neveiks. Tokiu atveju reikėtų naudotoją informuoti apie sutrikusios paslaugos veikimą ir nefiksuoti jo rezultatų. Neveikiant „Speech-to-text“ paslaugos pagrindinis programėlės funkcionalumas būtų nepasiekiamas. Paslaugos funkcionalumas tampa mokamas išnaudojus nemokamą limitą. Google debesų „Speech-to-text“ paslaugos šnekamosios kalbos atpažinimas gali būti nepakankamai tinkamas žmonių su kalbos sutrikimais reabilitacijai.

2.9. Apibendrinančios išvados

1. Atlikus srities apžvalgą buvo nustatyta kalbos atpažinimo taikymo sritys ir metodai. Nuspręsta toliau patyrinti „end-to-end“ metodą, nes jis leis laiku pasiketi projekto tikslus.
2. Renkantis technologinius sprendimus tikslui pasiekti buvo rasti trys vienas kito verti konkurentai, nustatyti jų privalumai ir trūkumai. Realizacijai buvo pasirinkta – „Google Cloud“ „Speech-to-Text“ paslauga, nes ji turi įgyvendintą kokybišką lietuvių kalbos atpažinimą, ir jos ekosistema tinkama mūsų projektui.
3. Atlikus rinkos analizę buvo nustatyta, kad ne visos tyrinėtos programos turi mūsų tyrimui reikalingą funkcionalumą. Esminis apžvelgtų programų trūkumas yra tai, kad jos neturi lietuvių kalbos. Todėl bus kuriama kalbos lavinimo programa, kurios esminė savybė bus lietuvių kalbos atpažinimas.

3. Projektinė dalis

3.1. Reikalavimų specifikacija

3.1.1. Sistemos paskirtis

3.1.1.1. Projekto kūrimo pagrindas (pagrindimas)

Kalbos reabilitacijos programinė įranga dažniausiai yra skirta žmonėms su klausos sutrikimais, tobulinti jų kalbėjimo gebėjimus. Pradedama nuo paprastų raidžių garsų toliau skiemenų ir pereinant prie sudėtingų spontaniškų žodžių ir frazių. Tokia reabilitacija yra atliekama su terapeuto pagalba, tačiau programinė įranga leistų savarankišką mokymąsi ir padėtų žmonėms, kurie negali susimokėti už tokias sesijas su terapeutu. Mokantis savarankiškai nereikėtų kontaktuoti su žmonėmis to pasėkoje sumažėtų rizika užsikrėsti virusu.

3.1.1.2. Sistemos tikslai (paskirtis)

Darbo tikslas yra automatinis kalbos raiškumo įvertinimas pasitelkiant Google debesų „Speech-to-Text“ paslauga. Galutinė sistema pateiks užduotis su atitinkamais pratimais ir galimais pratimo atsakymais. Naudotojo pratimo teisingumas bus vertinamas pagal balso atpažinimo patikimumo procentą ir pritaikius Levenšteino atstumą. Levenšteino atstumo algoritmo pagalba bus nustatoma ar aptiktas tekstas atitinka reikalingą atsakymą, pritaikius lūžio tašką. Patikimumo procentas yra lygus gaunamų taškų skaičiui už pratimą. Naudotojas galės pažiūrėti savo statistiką, ten matyti kiek kartų bandė užduotį ir jos taškų vidurkį. Administratoriai ir logopedai galės sukurti užduotis ir valdyti naudotojus.

3.1.2. Užsakovai, pirkėjai ir kiti sistema suinteresuoti asmenys

Projekto pasiūlymas suformuluotas pagal Tomo Blažausko užsakymą.

Sistemą galėtų naudoti reabilitacijos įstaigos, kurios teikia kalbos sutrikimų lavinimo paslaugas. Taip savo paslaugas galėtų teikti nuotoliniu būdu, kas padėtų sumažinti lankytojų skaičių, juos būtų galima konsultuoti nuotoliniu būdu.

3.1.3. Naudotojai

Pagrindiniai programėlės naudotojai yra žmonės turintys kalbėjimo sutrikimų, norintys lavinti kalbą. Tai gali būti vaikai besimokantys tarti konkrečius garsus ar vyresni žmonės. Klientai galėtų būti logopedai, kurie norėtų pateikti savo užduotis programėleje pacientams.

3.1.4. Įpareigojantys apribojimai. (Mandated constraints)

3.1.4.1. Apribojimai sprendimui

Realizacijai keliami reikalavimai:

- Minimali Android versija 5.0.
- Programėlė lietuvių kalba.
- Programėlė veiktų tinkamai net jei vienu metu ją naudotų 100 žmonių.

- Programėlė turi būti kuriama „Kotlin“ programavimo kalba.
- Programa turi atpažinti sveiko žmogaus kalbą bent 80% tikslumu.
- Balsui verti į tekstą turi būti naudojama „Google Cloud“ „speech-to-text“ paslauga.
- Naudotojo atsakymo tekstą palyginti su teisingo atsakymo tekstu naudojantis Levenšteino atstumo algoritmu.
- Užskaityti naudotojo atsakymus, kurių Levenšteino atstumas mažesnis negu ketvirtadalis teisingo atsakymo teksto ilgio.

3.1.4.2. Diegimo aplinka

Programėlė diegiama telefonuose kurių Android versija ≥ 5.0 . Saityno sąsajai naudojamas „Heroku“ serveris. Duomenų bazei naudojama „Postgresql“. Paveikslėlių talpinimui naudojama AWS S3.

3.1.4.3. Bendradarbiaujančios sistemos

Balso įrašui versti į tekstą bus naudojama „Google Cloud“ „speech-to-text“ paslauga. Paveikslėlių talpinimui naudojama AWS S3.

3.1.4.4. Komerciniai specializuoti programų paketai

Programavimui bus naudojama „Android Studio“ ir „PyCharm“ programinė įranga.

3.1.4.5. Numatoma darbo vietos aplinka

Darbo vietoje privalo būti interneto ryšys. O pati vieta gali būti bet kur.

3.1.4.6. Sistemos kūrimo terminai

Pirmasis prototipas – 2021-05-17.

Galutinė sistemos versija – 2021-12-23.

3.1.4.7. Sistemos kūrimo biudžetas

Sistemos kūrimui nėra skirta pinigų. Biudžetas nulinis.

3.1.5. Terminų žodynas

„Android“ – atviro kodo operacinė sistema, daugiausia naudojama išmaniuosiuose telefonuose.

„Speech-to-Text“ - „Google“ sukurta balso vertimo į tekstą programa „Android“ operacinei sistemai.

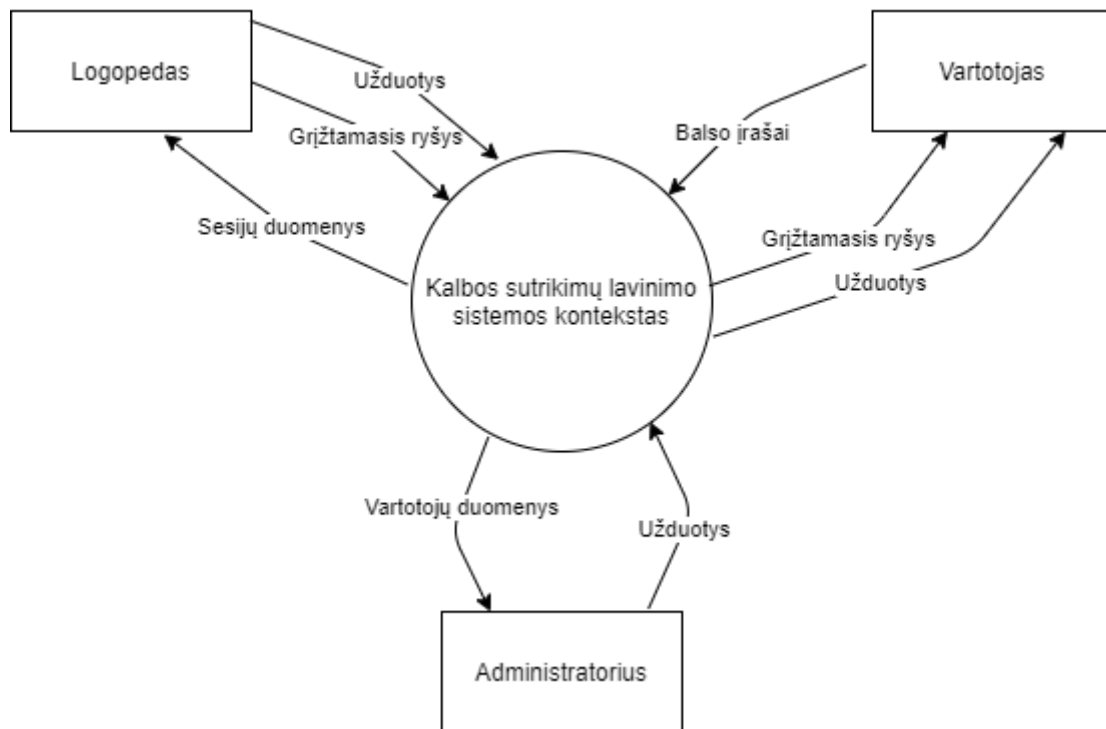
AWS S3 (Amazon Web Services Simple Storage Service) – Amazon kompanijos paslauga, kuri teikia objektų saugyklos paslaugas per interneto sąsają

3.1.6. Svarbūs faktai ir prielaidos

60 min per mėnesį galima nemokamai naudoti „Google Cloud“ „speech-to-text“ paslauga.

3.1.7. Veiklos sudėtis (The scope of the work)

3.1.7.1. Veiklos kontekstas (pateikiama konteksto diagrama)



1 pav. Kalbos sutrikimų lavinimo sistemos konteksto diagrama

3.1.7.2. Veiklos padalinimas

Naudotojas turėtų matyti galimas užduotis, peržiūrėti jų detalią informaciją ir pasirinkti spręsti užduotį. Sistemai naudotojas tiesks balso įrašus kuriuos sistema analizuos ir pateiks grįžtamąjį ryšį.

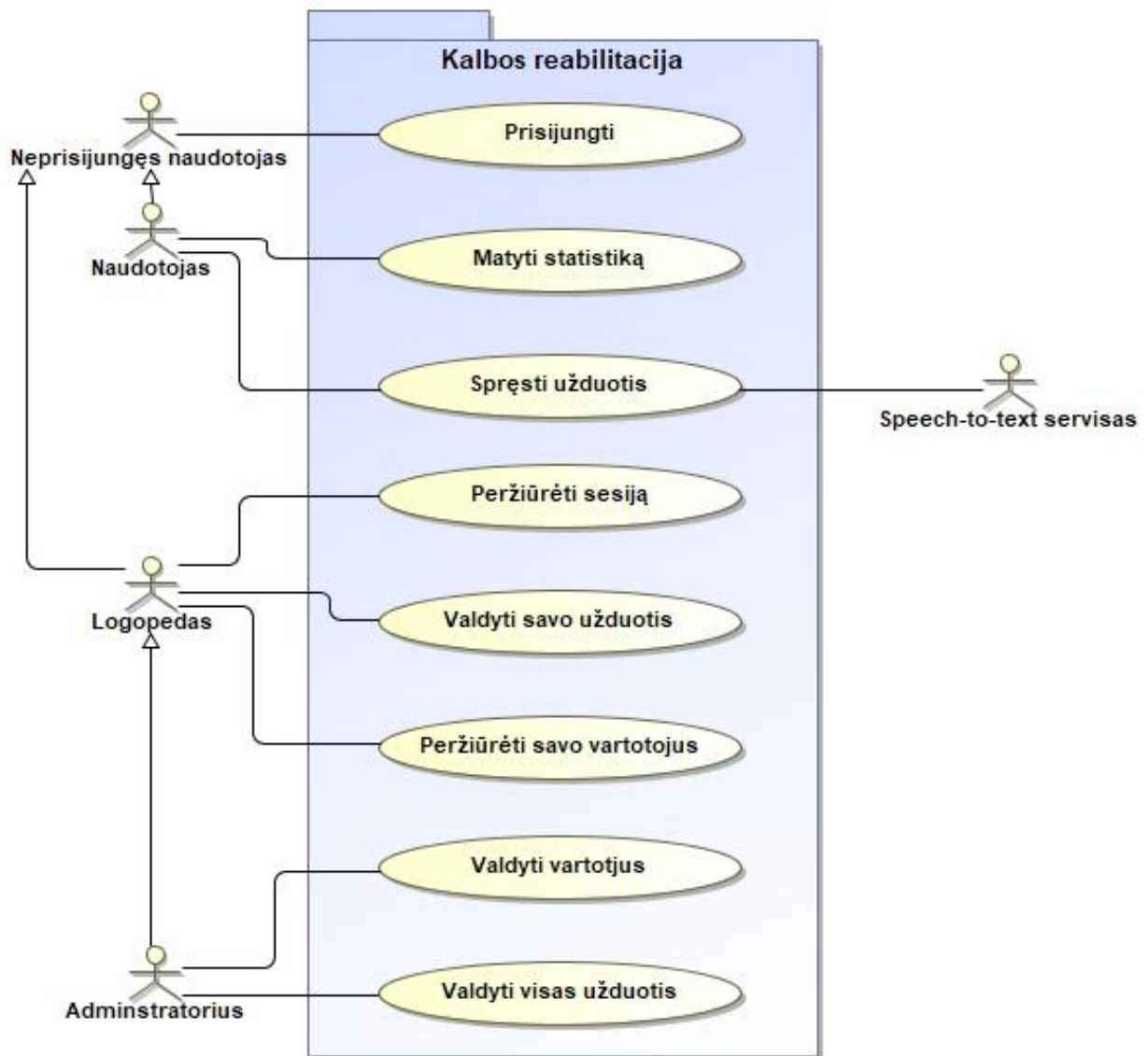
Logopedas turėtų galimybę teikti savo užduotis, peržiūrėti jam priskirtų naudotojų sesijas ir pateikti grįžtamąjį ryšį.

Administratorius turėtų galimybę valdyti naudotojus ir pateikti užduotis. Taip pat administratorius turėtų turėti visas logopedo galimybes.

Svarbiausia sistemos dalis yra balso vertimas į tekstą. Pagal tai sistema pateiks grįžtamąjį ryšį naudotojui.

3.1.8. Sistemos sudėtis (The scope of the product)

3.1.8.1. Sistemos ribos



2 pav. Panaudojimo atvejų diagrama

Naudotojas – naudosis programėle užduotims spręsti stebėti statistikai.

Logopedas – naudosis programėle valdyti savo užduotimis ir naudotojams, taip pat priskirti užduotims ir peržiūrėti naudotojų sesijoms.

Administratorius – naudosis programėle valdyti visiems naudotojoms bei užduotims.

Speech-to-text servisas – verčia žmogaus pasakytus žodžius į tekstą, be jo neįmanoma spręsti užduočių.

3.1.8.2. Panaudojimo atvejų sąrašas

Kuriamos programų sistemos panaudojimo atvejai:

- Valdyti visas užduotis (lentelė 3).
- Valdyti savo užduotis (lentelė 4).
- Valdyti naudotojus (lentelė 5).
- Peržiūrėti savo naudotojus (lentelė 6).
- Peržiūrėti sesiją (lentelė 7).
- Prisijungti (lentelė 8).
- Spręsti užduotis (lentelė 9).
- Matyti statistiką (lentelė 10).

3 lentelė. Valdyti visas užduotis panaudojimo atvejo specifikacija.

Panaudojimo atvejis		Valdyti visas užduotis.
Tikslas. Peržiūrėti, kurti naujas, redaguoti arba trinti jau esamas užduotis.		
Prieš sąlyga		Naudotojas turi būti prisijungęs prie sistemos ir turėti administratoriaus rolę.
Aktoriai		Administratorius.
Susiję panaudojimo atvejai	Apimami PA	
	Išplečiantys PA	
	Specializuojami PA	
Sužadinimas		Naudotojo pasirinkimas meniu.
Po sąlyga		Sukūrus naują įrašą rodoma detali to įrašo informacija. Paredagavus įrašą jis yra išsaugomas ir grįžtama į peržiūrą. Ištrynus įrašą grįžtama į sąrašą.
Pagrindinis scenarijus		Naudotojas teisingai suveda duomenis ir išsaugo formą.
Alternatyvus scenarijus		Naudotojas neteisingai suveda formos duomenis.

4 lentelė. Valdyti savo užduotis panaudojimo atvejo specifikacija.

Panaudojimo atvejis		Valdyti savo užduotis.
Tikslas. Peržiūrėti, kurti naujas, redaguoti arba trinti sukurtas užduotis.		
Prieš sąlyga		Naudotojas turi būti prisijungęs prie sistemos ir turėti logopedo arba administratoriaus rolę.
Aktoriai		Logopedas, Administratorius.
Susiję panaudojimo atvejai	Apimami PA	
	Išplečiantys PA	
	Specializuojami PA	
Sužadinimas		Naudotojo pasirinkimas meniu.

Po sąlyga	Sukūrus naują įrašą rodoma detali to įrašo informacija. Paredagavus įrašą jis yra išsaugomas ir grįžtama į peržiūrą. Ištrynus įrašą grįžtama į sąrašą.
Pagrindinis scenarijus	Naudotojas teisingai suveda duomenis ir išsaugo formą.
Alternatyvus scenarijus	Naudotojas neteisingai suveda formos duomenis.

5 lentelė. Valdyti naudotojus panaudojimo atvejo specifikacija.

Panaudojimo atvejis	Valdyti naudotojus.
Tikslas. Peržiūrėti, redaguoti arba trinti naudotojus.	
Prieš sąlyga	Naudotojas turi būti prisijungęs prie sistemos ir būti administratoriaus rolės.
Aktoriai	Administratorius.
Susiję panaudojimo atvejai	Apimami PA
	Išplečiantys PA
	Specializuojami PA
Sužadinimas	Naudotojo pasirinkimas meniu.
Po sąlyga	Paredagavus informaciją forma yra išsaugoma ir grįžtama į peržiūrą. Ištrynus naudotoją grįžtama į sąrašą.
Pagrindinis scenarijus	Naudotojas teisingai suveda duomenis ir išsaugo formą.
Alternatyvus scenarijus	Naudotojas neteisingai suveda formos duomenis.

6 lentelė. Peržiūrėti savo naudotojus panaudojimo atvejo specifikacija.

Panaudojimo atvejis	Peržiūrėti savo naudotojus.
Tikslas. Peržiūrėti priskirtus naudotojus.	
Prieš sąlyga	Naudotojas turi būti prisijungęs prie sistemos ir būti administratoriaus arba logopedo rolės.
Aktoriai	Administratorius, logopedas
Susiję panaudojimo atvejai	Apimami PA
	Išplečiantys PA
	Specializuojami PA
Sužadinimas	Naudotojo pasirinkimas meniu.
Po sąlyga	Logopedas mato priskirtų naudotojų informaciją.
Pagrindinis scenarijus	Logopedas pasirinkęs meniu punktą mato priskirtų naudotojų informaciją.
Alternatyvus scenarijus	Nėra.

7 lentelė. Peržiūrėti sesiją panaudojimo atvejo specifikacija.

Panaudojimo atvejis		Peržiūrėti sesiją
Tikslas. Peržiūrėti kaip naudotojai vykdo užduotis. Kokius rezultatus gauna.		
Prieš sąlyga		Naudotojas turi būti prisijungęs prie sistemos ir turi būti administratoriaus arba logopedo rolės.
Aktoriai		Administratorius, logopedas.
Susiję panaudojimo atvejai	Apimami PA	
	Išplečiantys PA	
	Specializuojami PA	
Sužadinimas		Pasirinkus naudotoją pasirenkama peržiūrėti sesiją.
Po sąlyga		Sesija pažymima kaip peržiūrėta.
Pagrindinis scenarijus		Paliekamas atsiliepimas apie naudotojo sesiją.
Alternatyvus scenarijus		Nepaliekamas atsiliepimas apie naudotojo sesiją.

8 lentelė. Prisijungti panaudojimo atvejo specifikacija.

Panaudojimo atvejis		Prisijungti
Tikslas. Prisijungti prie sistemos pasinaudoti jos funkcionalumu.		
Prieš sąlyga		Naudotojas turi instaliuotą programėlę
Aktoriai		Administratorius, Logopedas, Naudotojas
Susiję panaudojimo atvejai	Apimami PA	
	Išplečiantys PA	
	Specializuojami PA	
Sužadinimas		Atidaroma programėlė.
Po sąlyga		Naudotojas prijungiamas prie sistemos.
Pagrindinis scenarijus		Naudotojas atidaro programėlę. Suveda teisingai duomenis. Prijungiamas prie sistemos.
Alternatyvus scenarijus		Suvesti blogi duomenys arba serveris nepasiekiamas. Tada išvedamas atitinkamas klaidos pranešimas.

9 lentelė. Spręsti užduotis panaudojimo atvejo specifikacija.

Panaudojimo atvejis		Spręsti užduotis
Tikslas. Spręsti priskirtas užduotis ir gauti atgalinį ryšį apie užduoties teisingumą.		
Prieš sąlyga		Naudotojas turi būti prisijungęs prie sistemos.
Aktoriai		Naudotojas
	Apimami PA	
	Išplečiantys PA	

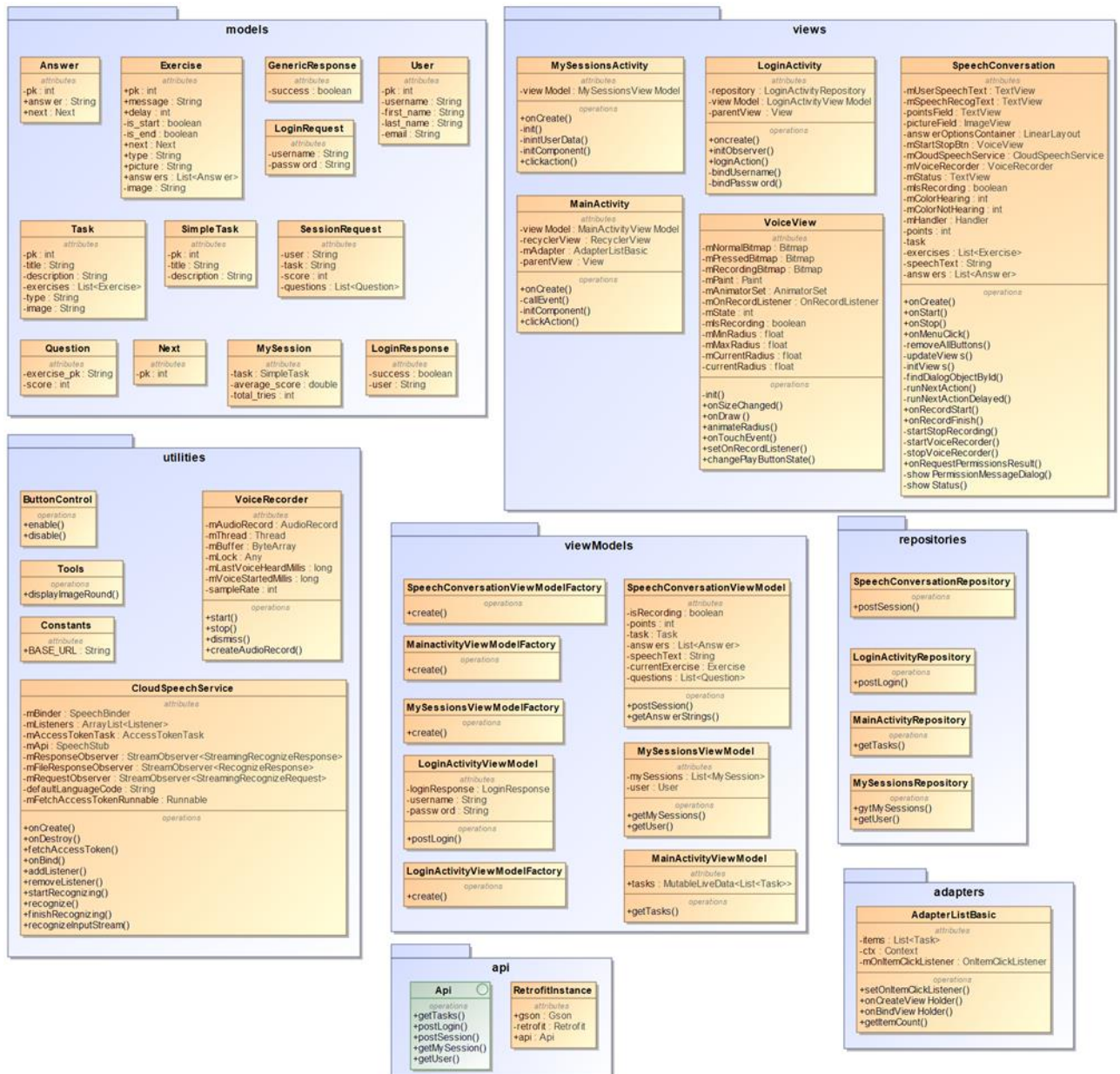
Susiję panaudojimo atvejai	Specializuojami PA	
Sužadinimas		Pasirinkama užduotis spręsti.
Po sąlyga		Išvedamas užduoties įvykdymo rezultatas.
Pagrindinis scenarijus		Naudotojas pasirenka užduotį. Ją įvykdo. Gauna atgalinį ryšį, kaip teisingai įvykdė užduotį.
Alternatyvus scenarijus		„Speech-to-text“ servisas yra neapsiekiamas. Tuo atveju turėtų neleisti vykdyti užduoties.

10 lentelė. Matyti statistiką panaudojimo atvejo specifikacija.

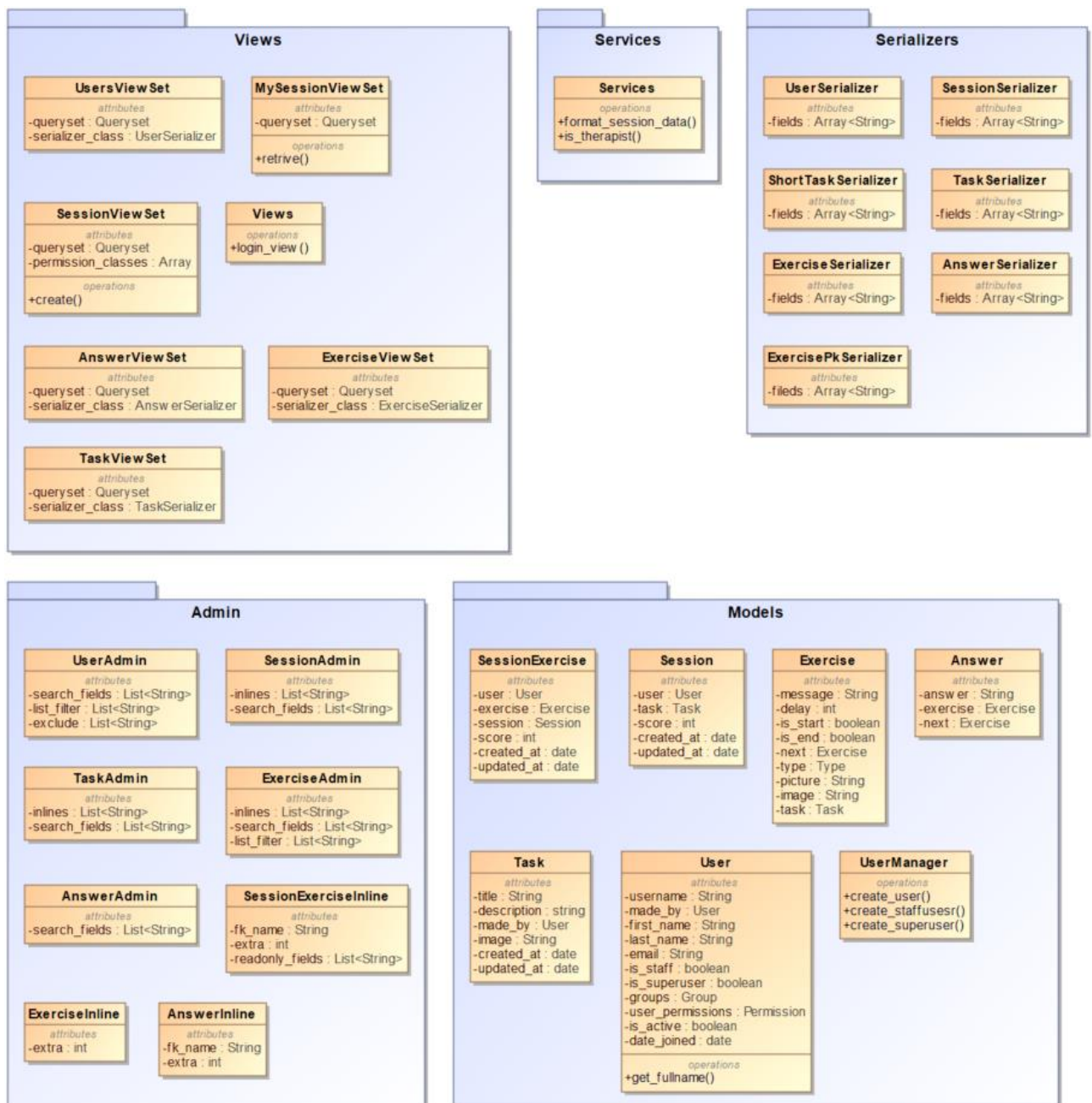
Panaudojimo atvejis		Matyti statistiką
Tikslas.		Matyti statistiką apie įvykdytas ir likusias užduotis. Kiek laiko praleista prie kiekvienos užduoties, kokie rezultatai.
Prieš sąlyga		Naudotojas turi būti prisijungęs prie sistemos.
Aktoriai		Naudotojas
Susiję panaudojimo atvejai	Apimami PA	
	Išplečiantys PA	
	Specializuojami PA	
Sužadinimas		Naudotas pasirenka meniu punktą.
Po sąlyga		Naudotojas mato savo užduočių informaciją.
Pagrindinis scenarijus		Naudotojui pasirinkus meniu punktą peržiūrėti savo informaciją išvedama statistika apie įvykdytas užduotis.
Alternatyvus scenarijus		Nėra.

3.1.9. Funkciniai reikalavimai duomenims

3-4 pav. pateikiama kuriamos programėlės ir saityno sąsajos klasių diagramos.



3 pav. Klasių diagrama



4 pav. Klasių diagrama

3.1.10. Reikalavimai panaudojamumui (Usability)

Sistemos panaudojamumo nefunkciniai reikalavimai:

- Sistemos naudotojo sąsaja turi būti pasiekama naudojant Android išmanųjį įrenginį su ≥ 5 Android versija.
- Sistema turi turėti Lietuvių kalbos naudotojo sąsają.
- Sistema turi būti parasta naudotis asmenims turintiems žemą informacinio raštingumo lygį.

3.1.11. Reikalavimai vykdymo charakteristikoms (Performance)

Sistemos vykdymo charakteristikų nefunkciniai reikalavimai:

- Sistemoje turi būti numatyta galimybė vykdyti duomenų gavimą asinchroniškai, t. y. duomenų apdorojimas neturi daryti įtakos sistemos greitimeikiai.
- Sistema turi išvesti visų matematinių skaičiavimų rezultatus sveiku skaičiumi.

3.1.12. Reikalavimai sistemos priežiūrai (Maintainability and portability)

Sistemos priežiūros nefunkciniai reikalavimai:

- Sistemoje turi būti diegiami periodiniai atnaujinimai, taisomos klaidos visą sutarties laikotarpį.

3.1.13. Reikalavimai saugumui (Security)

Sistemos saugumo nefunkciniai reikalavimai:

- Sistema turi būti apsaugota nuo neautentifikuotos prieigos, nesankcionuoto naudotojo sąsajos perėmimo, nesankcionuoto duomenų perėmimo ar jų įterpimo.
- Sistema turi būti apsaugota nuo žalingo kodo įterpimo (angl. „Injection“, XSS) ir kitų saugumo pažeidimų, kurie įvardijami OWASP TOP 10 (<https://www.owasp.org>) sąrašė (arba lygiaverčiame).

3.2. Architektūros specifikacija

3.2.1. Architektūros pateikimas

Dokumente kuriamos sistemos architektūra pateikiama skirtingais aspektais: panaudojimo atvejų vaizdu, statiniu vaizdu, dinaminiu arba procesų vaizdu, ir išdėstymo vaizdu. Jų pateikimui naudojama unifikuota modeliavimo kalba (UML). Sistemos atvaizdavimui šiais vaizdais yra naudojamosi standartinėmis UML diagramomis:

- Panaudos atvejų vaizdas:
 - Panaudos atvejų diagrama.
- Statinis vaizdas:
 - Sistemos išskaidymo į paketus diagrama;
 - Klasių diagrama.
- Dinaminis (procesų) vaizdas:
 - Bendravimo diagrama;
 - Sekų diagrama;
 - Veiklos diagrama
- Išdėstymo vaizdas:
 - Išdėstymo diagrama.

3.2.2. Architektūros tikslai ir apribojimai

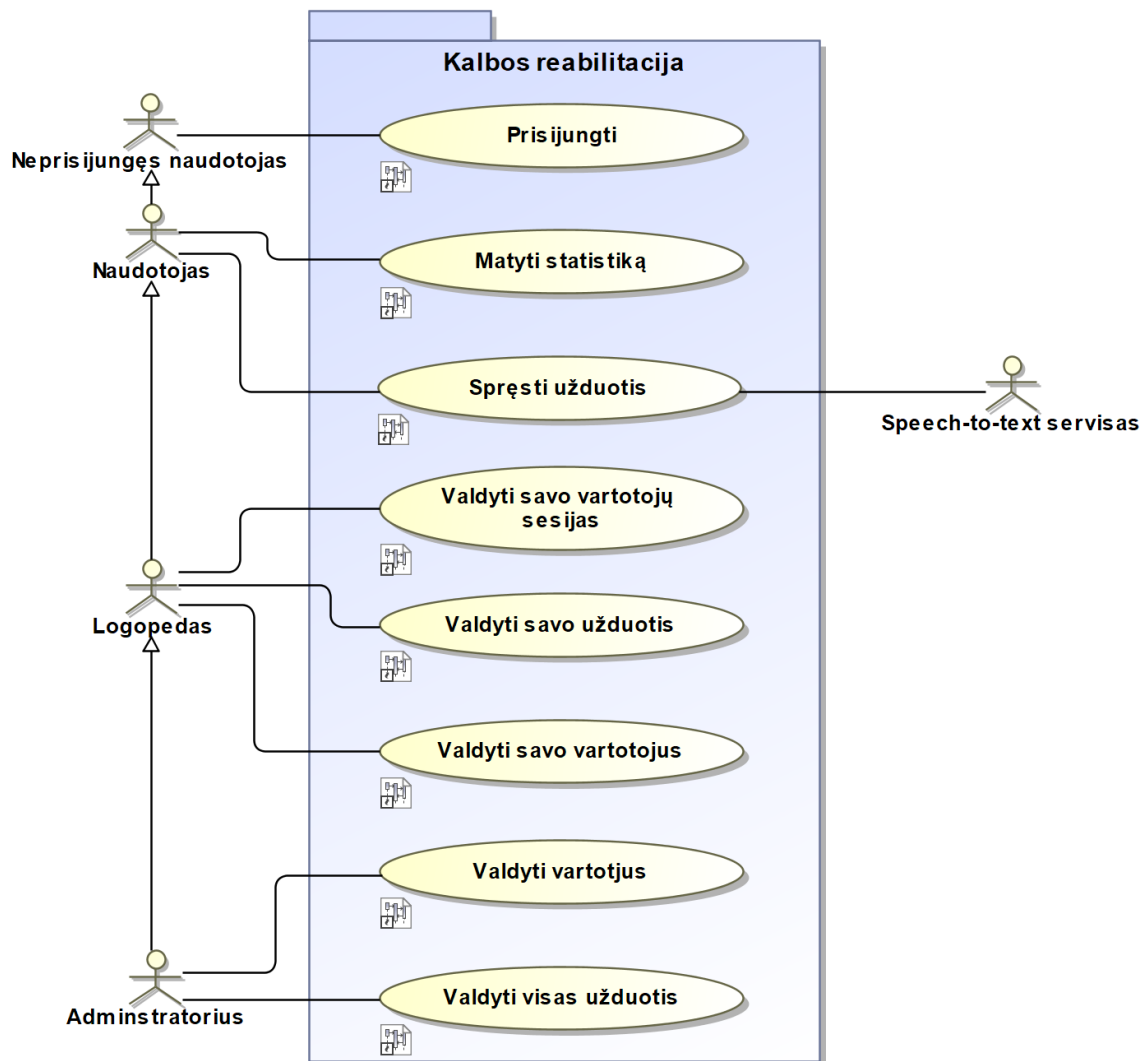
Reikalavimai ir apribojimai, kurie turi įtakos sistemos architektūrai:

- Minimali Android versija 5.0.
- Programėlė turi būti kuriama „Kotlin“ programavimo kalba.

- Programėlės sąsaja turi būti lietuvių kalba.
- Programėlėje turi turėti integraciją su „Google Cloud“ „speech-to-text“ paslauga.
- Saityno sąsaja turi turėti integraciją su AWS S3 paslauga.
- Saityno sąsaja ir duomenų bazė talpinama „Heroku“ serveryje

3.2.3. Panaudojimo atvejų vaizdas

Pateikiami esminiai panaudojimo atvejai ir naudotojų funkcijos.



5 pav. Panaudojimo atvejų diagrama

Naudotojas – naudosis programėle užduotims spręsti ir stebėti statistikai.

Logopedas – naudosis saityno sąsaja valdyti savo užduotimis ir naudotojams, taip pat valdyti naudotojų sesijoms.

Administratorius – naudosis saityno sąsaja valdyti visiems naudotojams bei užduotims.

Speech-to-text servisas – verčia žmogaus pasakytus žodžius į tekstą, be jo neįmanoma spręsti užduočių.

3.2.4. Statinis sistemos vaizdas

Sistemos statiniui vaizdui pavaizduoti pateikiama 6 paveikslėlyje pavaizduota programėlės paketų diagrama, kurią sudaro aštuoni paketai.

„Views“ pakete talpinama naudotojo sąsajos komponentai. MVVM architektūros „views“ yra atsakingi už duomenų, gautų iš valdiklio, rodymą. „View“ yra transformuotas modelis į naudotojo sąsają.

„Model“ paketo klasėse yra saugomi duomenys ir susijusi logika. Tai reiškia duomenis, kurie perduodami tarp valdiklio komponentų ar bet kurios kitos susijusios verslo logikos.

„ViewModels“ paketuose saugomos klasės, kurių pagrindinė paskirtis išvesti duomenis į ekraną ir reaguoti į naudotojo sąveika su programėle. „ViewModel“ klasės yra atsakingos už funkcijų, komandų, metodų pateikimą, kad būtų kontroliuojama „views“ klasių būseną. Šio paketo klasės taip pat atsakingos už „models“ valdymą ir įvykių suaktyvinimą „views“ paketų klasėse.

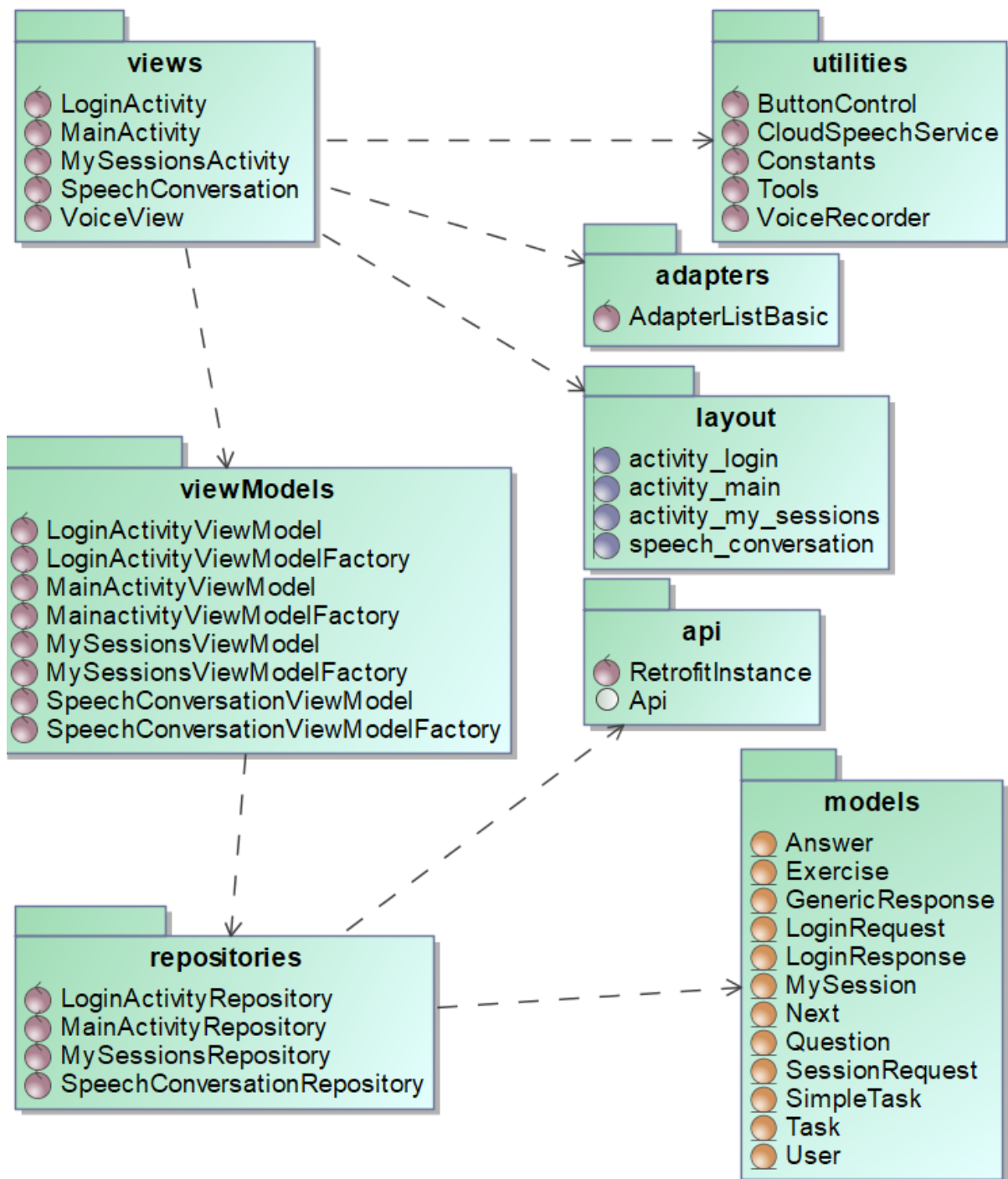
„Repository“ projektavimo schema palengvina verslo logikos ir prieigos prie duomenų sluoksnių atsiejimą. Duomenys yra tiesiog atiduodami nežinant kaip vėliau jie bus atvaizduojami.

„Api“ pakete esančios klasės atsakingos už komunikavimą su duomenų baze, duomenų išsiuntimą ir gavimą.

„Utilities“ pakete yra pagalbinės klasės. Tai konstantos, įvairūs įrankiai.

„Adapters“ pakete yra klasės, kurios padeda atvaizduoti elementus ekrane tam tikru būdu.

„Layout“ pakete egzistuoja elementai, kurie yra naudotojo sąsajos brėžiniai. Struktūruotai nurodo kaip duomenys bus atvaizduoti.



6 pav. Programėlės paketų diagrama

7 paveikslėlyje pavaizduota saityno sąsajos paketų diagrama, kurią sudaro šeši paketai.

„Views“ pakete klasės atsakingos už komunikavimą, duomenų išsiuntimą ir gavimą.

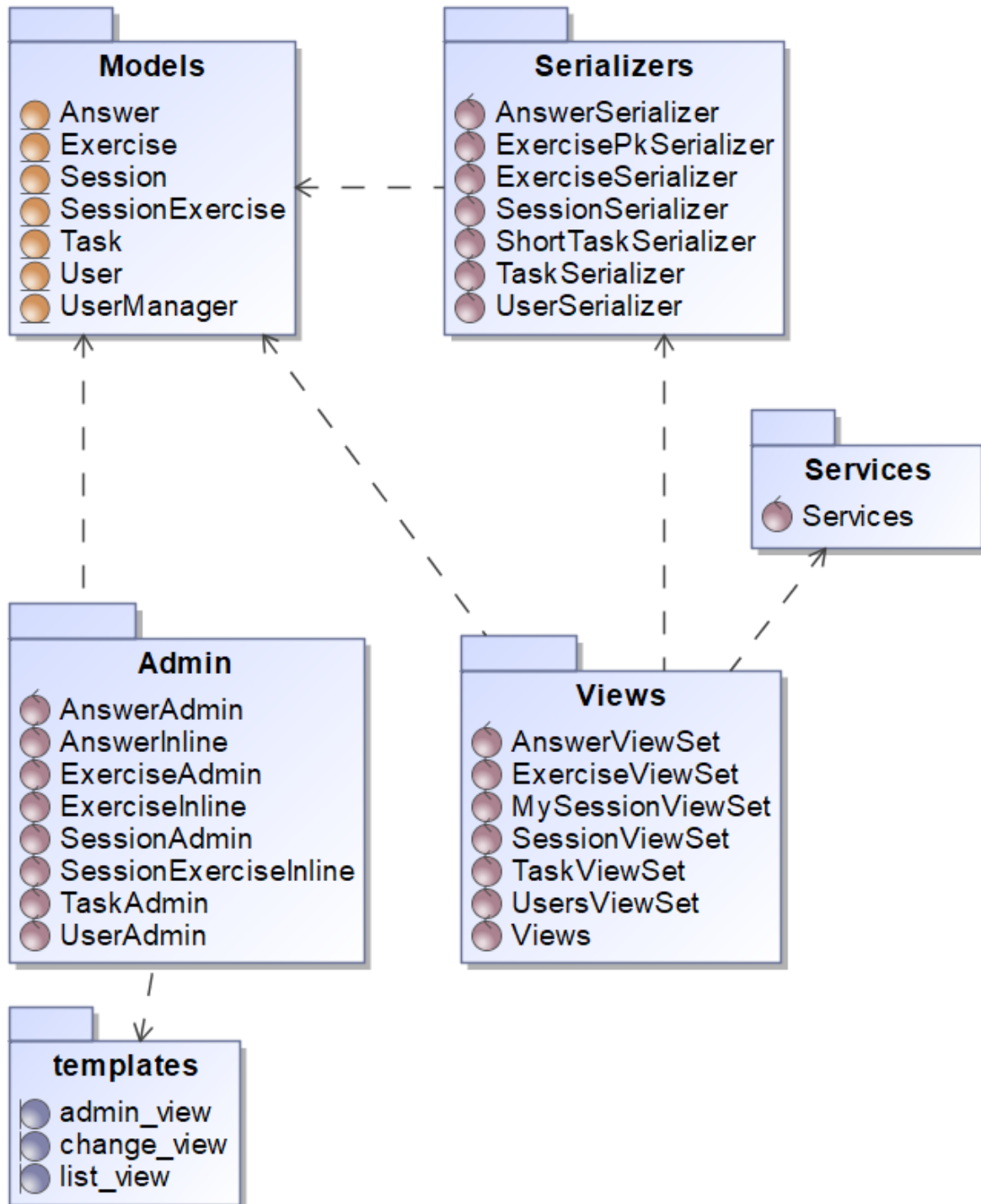
„Services“ pakete yra saugomos klasės ir funkcijos, kurios palengvina darbą. Tam tikri pasikartojantys kodo elementai yra išskelti į paslaugas, kad jais būtų galima pasinaudoti.

„Serializers“ pakete yra saugomos klasės kurios atsakingos už duomenų serializavimą t. y. duomenų validavimą, duomenų vertiną į JSON formatą ir iš JSON į modelį.

„Model“ paketo klasėse yra saugomi duomenys ir susijusi logika. Tai reiškia duomenis, kurie perduodami tarp valdiklio komponentų ar bet kurios kitos susijusios verslo logikos.

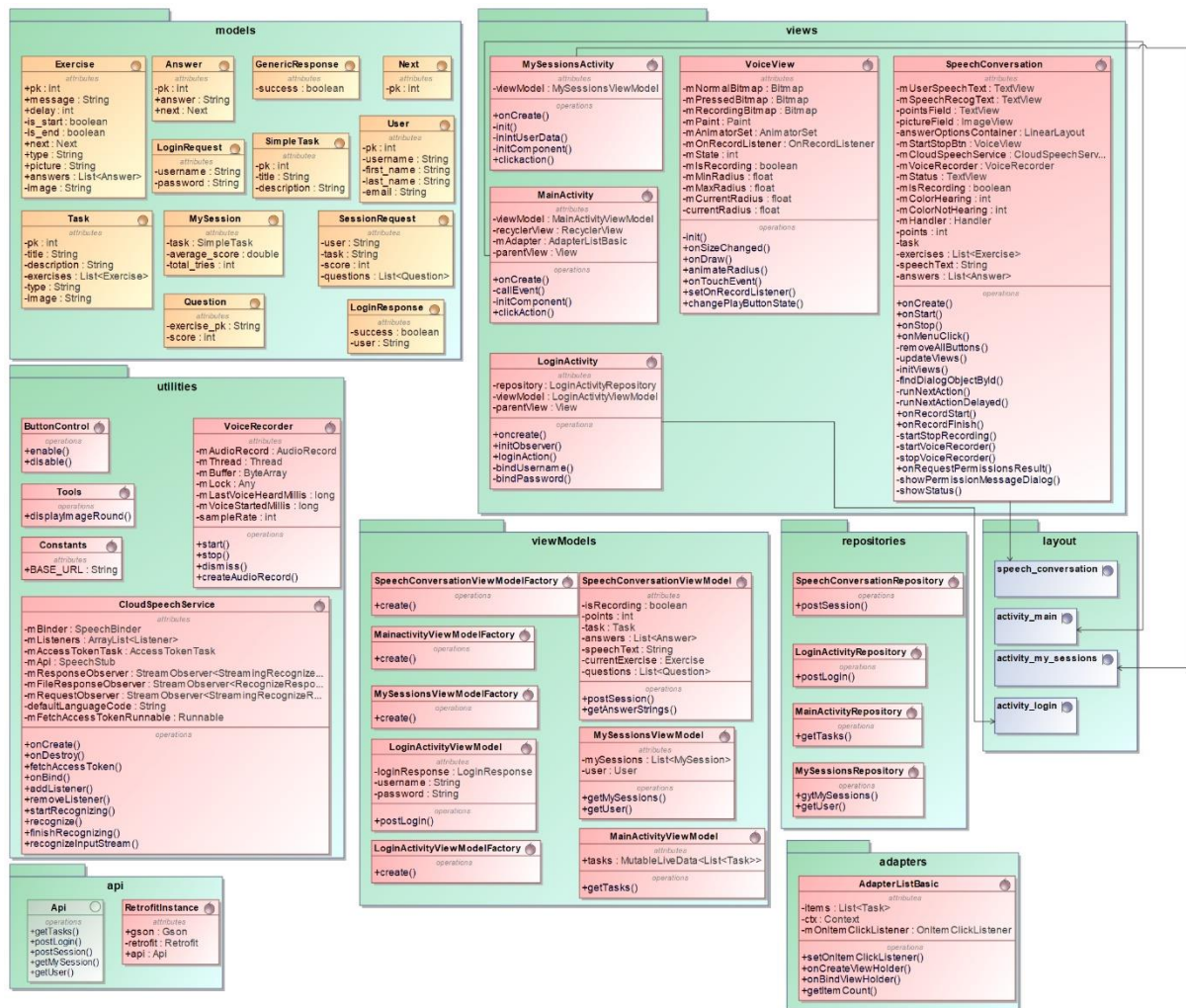
„Admin“ pakete esančios klasės atsakingos už sukūrimo, peržiūros redagavimo ir ištrynimo funkcijas.

„Template“ pakete egzistuoja elementai, kurie yra naudotojo sąsajos brėžiniai. Struktūruotai nurodo kaip duomenys bus atvaizduoti.



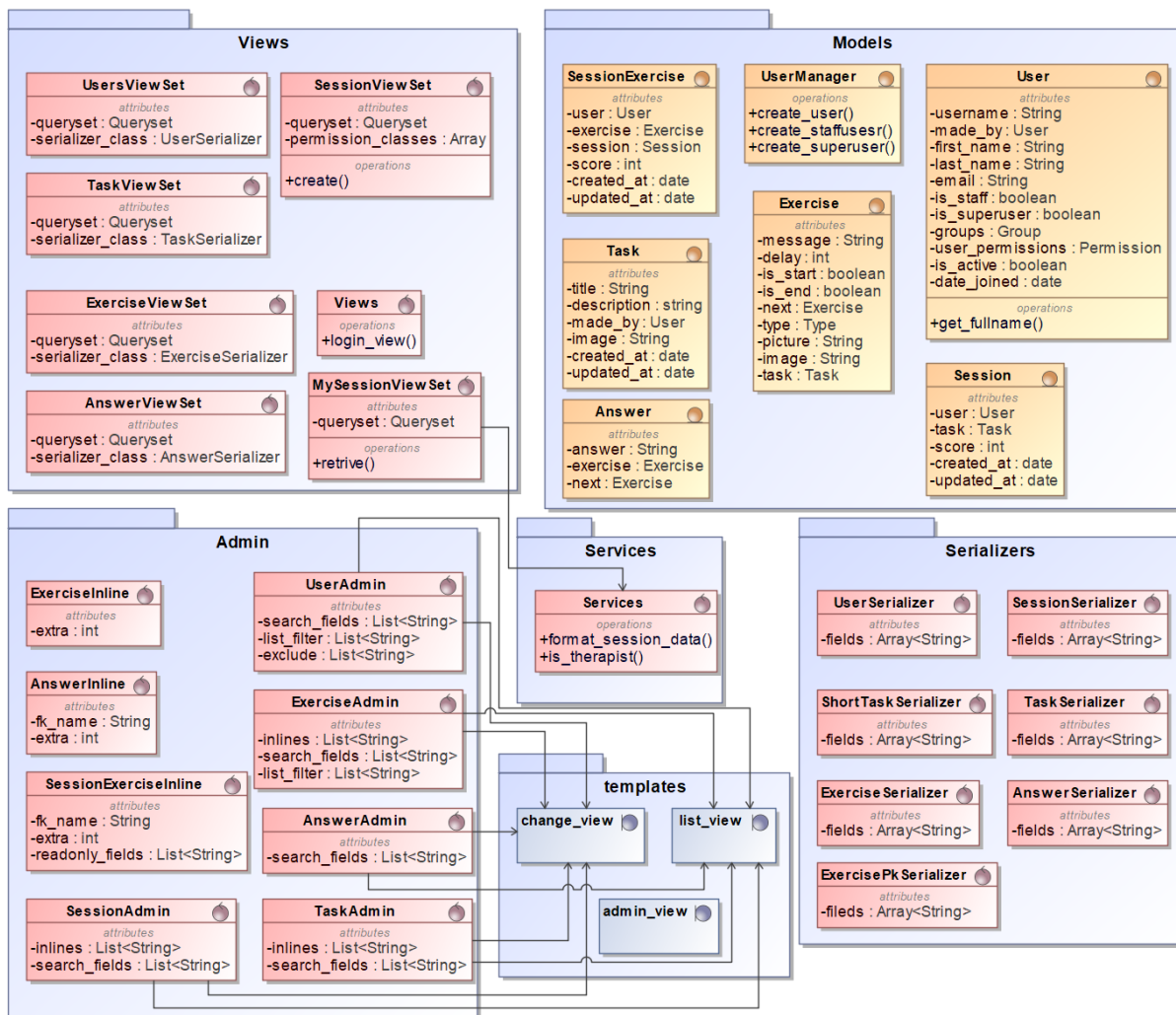
7 pav. Saityno sąsajos paketų diagrama

8 paveikslėlyje yra pateikiama programėlės klasių diagrama. Šioje diagramoje galima matyti detalizuotus paketus ir jų klasių turimus atributus ir metodus.



8 pav. Programėlės klasių diagrama

9 paveikslėlyje yra pateikiama saityno sąsajos klasių diagrama. Šioje diagramoje galima matyti detalizuotus paketus ir jų klasių turimus atributus ir metodus.



9 pav. Saityno sąsajos klasių diagrama

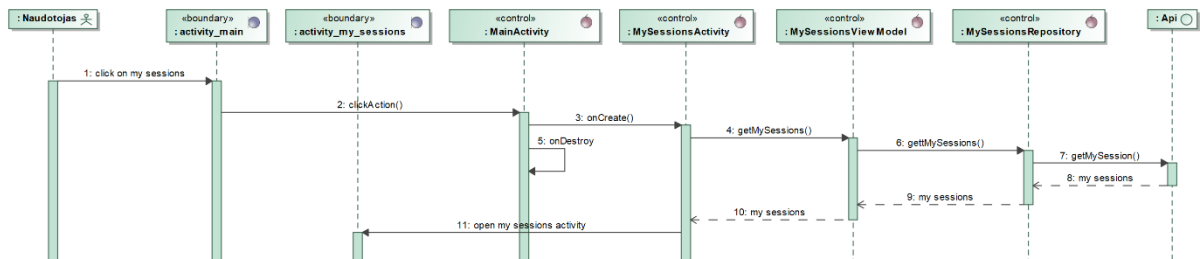
3.2.5. Sistemos dinaminis vaizdas

Sistemos dinaminiam vaizdai pavaizduoti pateikiamos sąveikos, būsenų ir veiklos diagramos. Sąveikai atvaizduoti pasirinkta sekų ir bendravimo diagramos.

3.2.5.1. Sekų diagramos

Sekų diagramos (11 - 18 pav.) pateikiamos svarbiausiems panaudojimo atvejams. Šiose diagramose matosi kokius medus kviečia klasės vykdant tam tikrus panaudojimo atvejus.

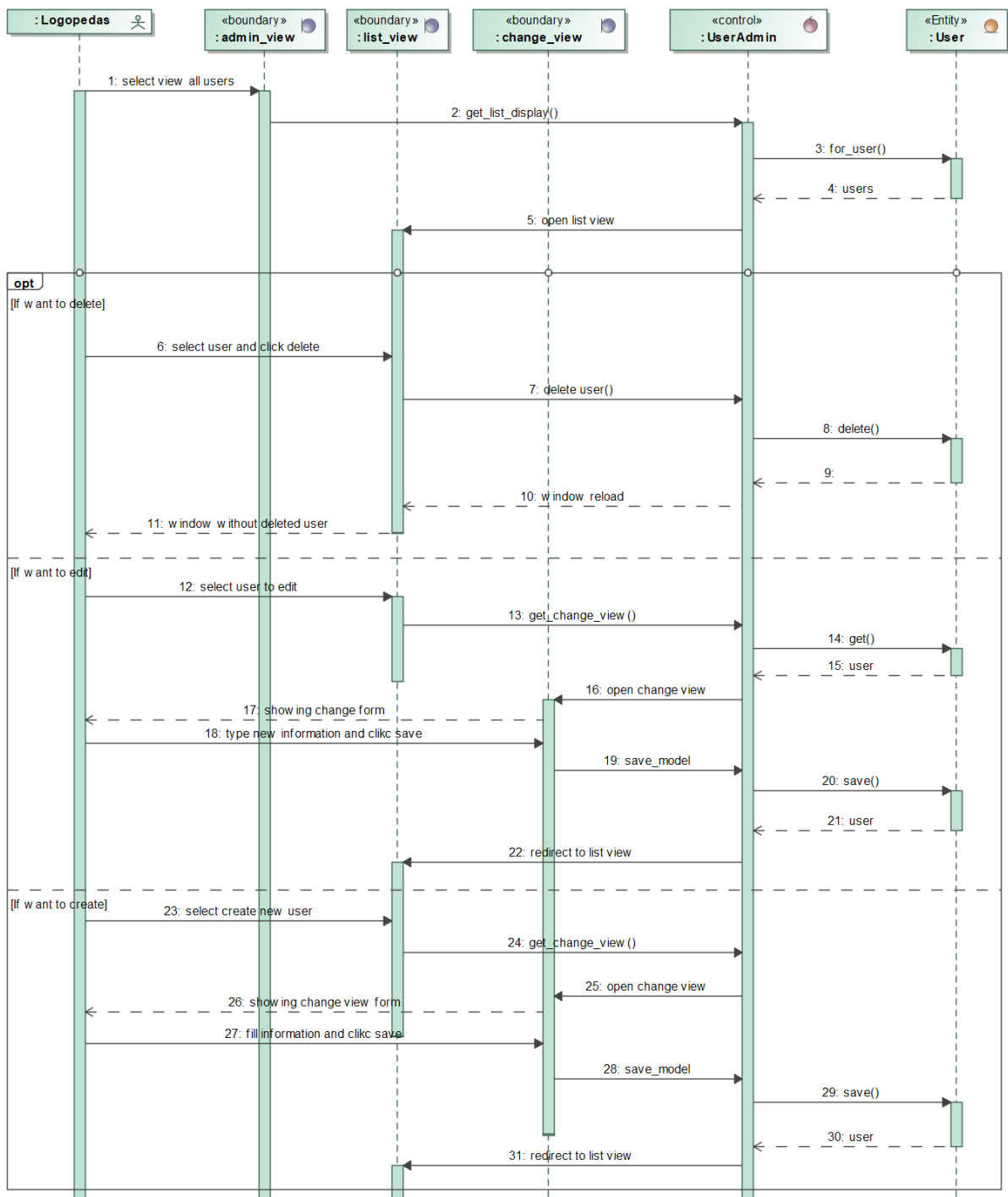
10 paveikslėlyje pavaizduota kokios veiksmų sekos vykdomos, kai naudotojas nori peržiūrėti statistiką.



10 pav. Matyti statistiką sekų diagrama

Naudotojas inicijuoja veiksmą pasirinkdamas tinkamą meniu punktą. Pastebėtas paspaudimas yra apdorojamas „MainActivity“ klasės, kuri sukuria „MySessionActivity“ klasės objektą. „MySessionActivity“ pasirūpina duomenimis, kuriuos reikės atvaizduoti, kreipdamasi į „MySessionsViewModel“, o šis į „MySessionsRepository“, o ši klasė kreipiasi į „Api“ gauti duomenis iš saityno sąsajos. „MySessionActivity“ gautus duomenis atvaizduoja naudotojui „activity_my_sessions“ lange.

11 paveikslėlyje pavaizduota valdyti savo naudotojus sekų diagrama.



11 pav. Valdyti savo naudotojus sekų diagrama

Norint valdyti savo naudotojus logopedas pasirenką matyti naudotojų sąrašą. Šį pasirinkimą aptinka „UserAdmin“ klasė. Ji užsiprašo naudotojų duomenų, kuriuos gali matyti logopedas, iš duomenų bazės ir juos atvaizduoja naudotojų sąrašo lange „list_view“. Iš naudotojų sąrašo logopedas gali pasirinkti sukurti naują naudotoją arba redaguoti esamą arba ištrinti naudotoją.

Norint ištrinti naudotoją ar naudotojus reikia juos pasirinkti ir tada pasirinkti veiksmą ištrinti. „UserAdmin“ aptiks šį veiksmą ir iškvies ištrynimo metodą pasirinktoms naudotojų esybėms.

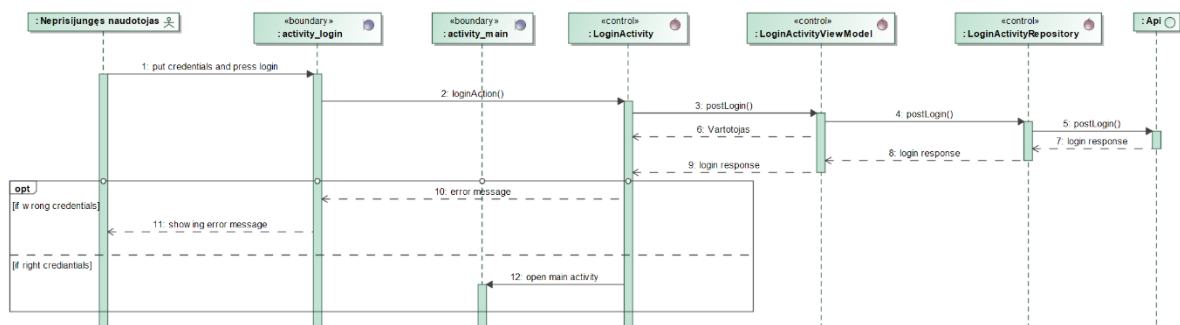
Pasirinkus redaguoti naudotoją „UserAdmin“ aptiks šį veiksmą, ir gavus naudotojo informaciją atidarys redagavimo langą „change_view“. Lange logopedas gali peržiūrėti naudotojo informaciją ir kažką redaguoti. Pasirinkus išsaugoti pakeitimus „UserAdmin“ aptiks šį veiksmą ir išsaugos pakeitimus redaguotam naudotojui, bei atidarys sąrašo langą „list_view“.

Pasirinkus sukurti naudotoją „UserAdmin“ aptiks šį veiksmą ir atidarys redagavimo langą „change_view“. Užpildžius informaciją ir pasirinkus išsaugoti „UserAdmin“ aptiks šį veiksmą ir išsaugos pakeitimus redaguotam naudotojui, bei atidarys sąrašo langą „list_view“.

Veiksmų sekos tokios pačios šiuose panaudojimo atvejuose:

- Valdyti savo naudotojų sesijas.
- Valdyti savo užduotis.
- Valdyti visas užduotis
- Valdyti naudotojus.

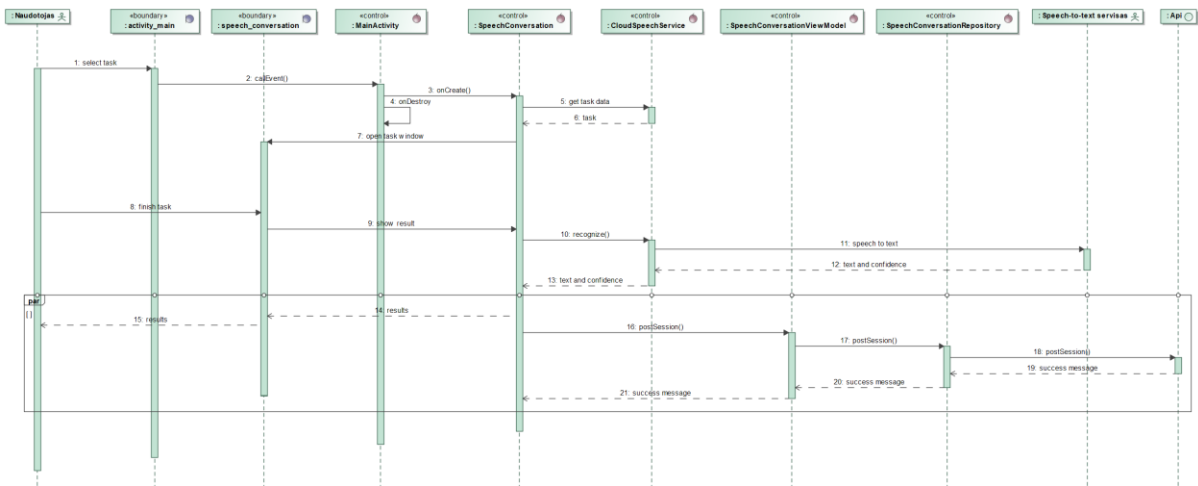
Prisijungimo metu vykdomos sekos pavaizduotos 12 paveikslėlyje.



12 pav. Prisijungti sekų diagrama

Norint prisijungti neprisijungęs naudotojas turi suvesti prisijungimo duomenis ir paspausti prisijungti. „LoginActivity“ aptiks šį veiksmą ir kreipsis į „LoginActivityViewModel“ šis į „LoginActivityRepository“ ši klasė į „Api“ gauti atsakymą ar naudotojo prisijungimo duomenys teisingi. Gavus sėkmės atsaką naudotojas yra prijungimas ir „LoginActivity“ atidaro „activity_main“ langą. Gavus klaidos pranešimą, „LoginActivity“ išveda klaidos pranešimą naudotojui.

Užduočių sprendimo metu vykdomos sekos pavaizduotos 13 paveikslėlyje.

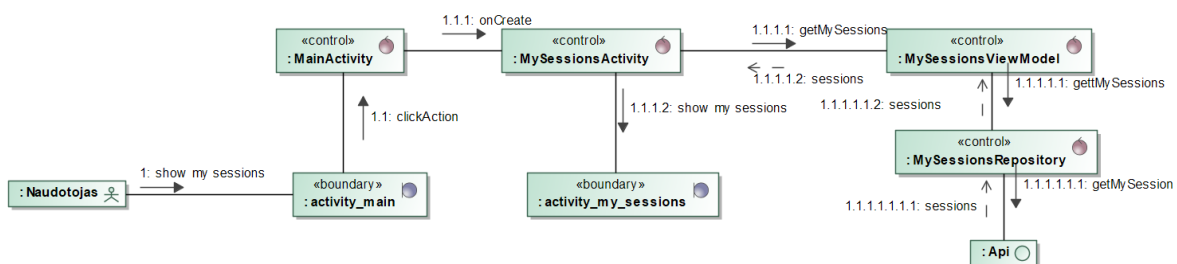


13 pav. Spręsti užduotis sekų diagrama

Norint spręsti užduotį naudotas privalo pasirinkti užduotį it „activity_main“ lange. Pasirinkimą aptinka „MainActivity“ klasė, kuri iškviečia „SpeechConversation“ klasė, o ši atidaro „speech_conversation“ langą. Šiame lange naudotojas spendžia užduotis. Jo sprendimus aptinka „SpeechConversation“ klasė, kuri kreipiasi į „CloudSpeechService“ klasę o ši į „Speech-to-text“ paslaugą gauti atpažintą tekstą ir patikimumo įvertinimą. Gautus rezultatus „SpeechConversation“ klasė išveda naudotojui ir kreipiasi į „SpeechConversationViewModel“ ši klasė į „SpeechConversationRepository“, o ši klasė į „Api“ išsaugoti naudotojo sesijai.

3.2.5.2. Komunikavimo diagrama

Komunikavimo diagramoje (14 paveikslėlis) vaizduojama kaip klasės bendrauja kai naudotojas nori pažiūrėti statistiką.

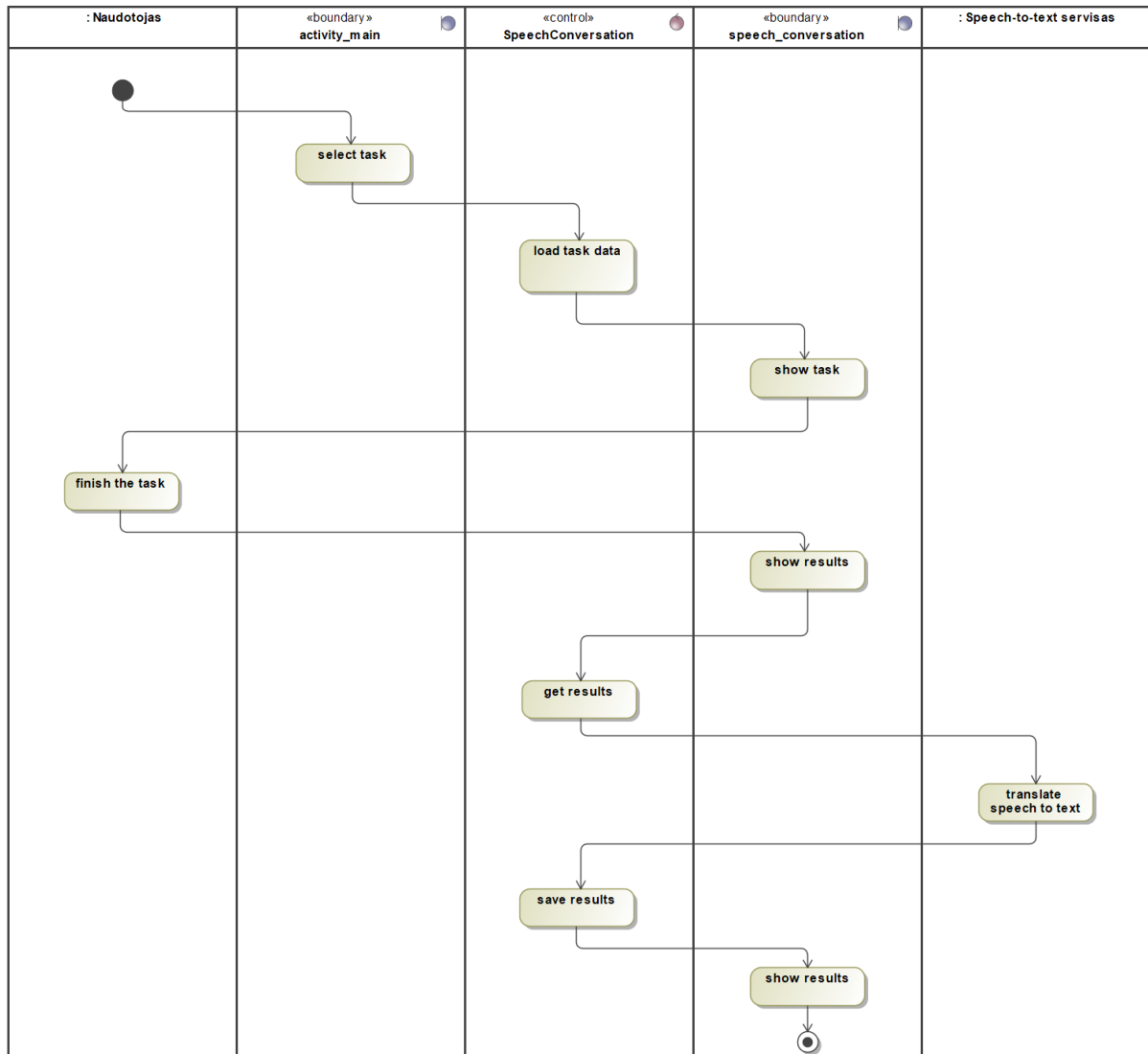


14 pav. Matyti statistiką komunikavimo diagrama

Iš diagramos matyti, kad naudotojas inicijuoja veiksmą „activity_main“ lange, vėliau jam atidaromas „activity_my_sessions“ langas. „MainActivity“ klasė iškviečia „MySessionsActivity“ klasę, kuri yra atsakinga už naudotojo sesijų vaizdavimą. „MySessionsActivity“ gauna duomenis iš „Api“ ir juos atvaizduoja „activity_my_sessions“ lange.

3.2.5.3. Veiklos diagramos

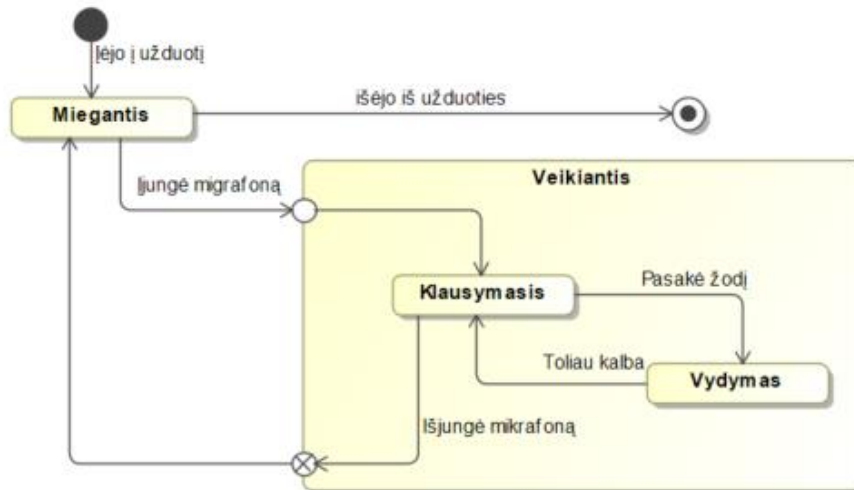
15 paveikslėlyje pateikiama naudotojo spręsti užduotis veiklos diagrama. Veiklos iniciatorius yra naudotojas. Veikloje dalyvauja Speech-to-text servisas, „activity_main“ ir „speech_conversation“ langai, bei „SpeechConversation“ valdiklis.



15 pav. Spręsti užduotis veiklos diagrama

3.2.5.4. Būsenų diagrama

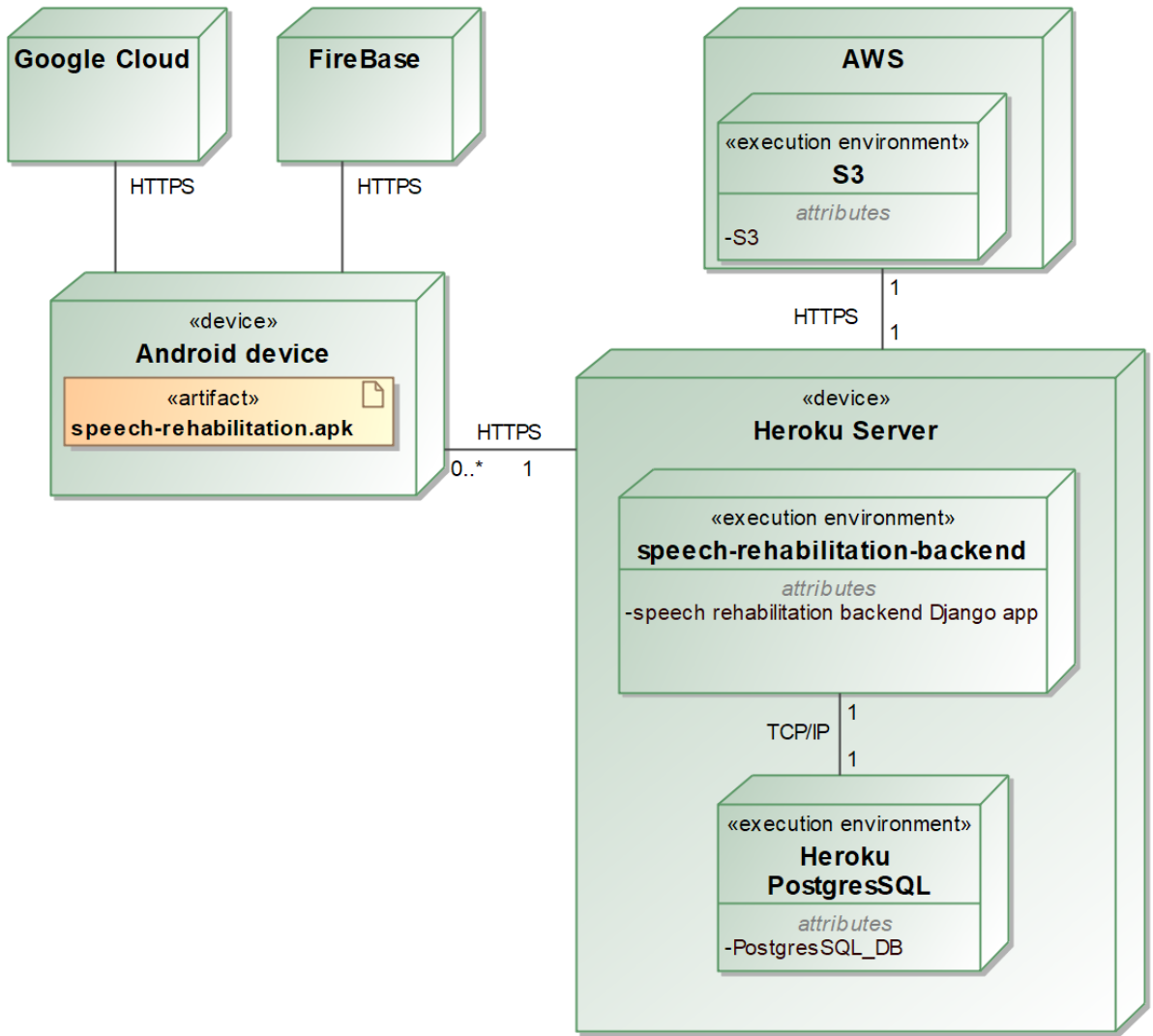
Atliekant užduotis mikrofonas keičia būsenas. Gali miegoti arba būti aktyvus. Aktyvioje būsenoje jis klausosi ir vykdo užklausas į „speech-to-text“ serverį. Šių būsenų diagrama pavaizduota 16 paveikslėlyje.



16 pav. Mikrofono būsenų diagrama

3.2.6. Išdėstymo (deployment) vaizdas

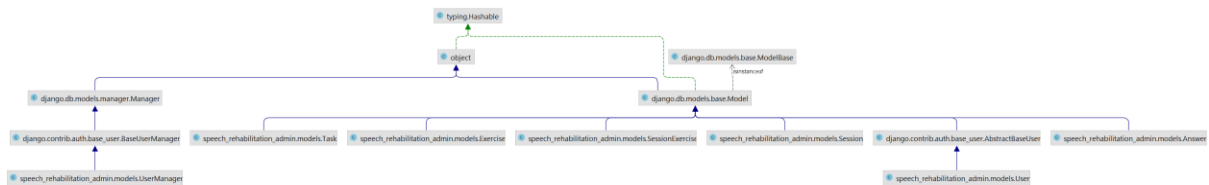
Programėlę veiks tik Android operacinę sistemą turinčiuose išmaniuose įrenginiuose. Programėlė bus diegiama telefonuose, kurių Android versija ≥ 5.0 . Programėlės komunikuos su „Google Cloud“ ir „FireBase“. Saityno sąsaja bus talpinama „Heroku“ serveryje. Duomenų bazei naudojama „Heroku PostgreSQL“. Saityno sąsaja komunikuos su „AWS“ S3 paveikslėlių saugojimui. Visa tai pavaizduota 17 paveikslėlyje.



17 pav. Diegimo diagrama

3.2.7. Duomenų vaizdas

Duomenų bazei bus naudojama „Heroku“ serveryje esanti „PostgreSQL“ duomenų bazė. Esybių ryšių pavyzdys pavaizduotas 18 paveikslėlyje.



18 pav. Esybių ryšių pavyzdys

3.2.8. Kokybė

Visi architektūros stiliai daro įtaką sistemos kokybei, dalinant sistemą į dalis. Programėlei kurti pasirinkta MVVM, o saityno sąsajai MVT architektūra. Šios architektūros turi savų pliusų kurie aptarti žemiau.

MVVM architektūros įtaka sistemai:

- Palaikomumas – Atskyrus skirtingas programos kodo dalis pagal MVVM architektūrą, atsiranda kodo struktūra ir vienodumas. Tada nesunku suprasti, kur viskas turėtų būti arba kur jie greičiausiai bus padėti.
- Išplečiamumas – Verslo logika atsieta nuo „UI“, todėl lengva pakartotinai naudoti komponentus. Verslo logika yra „ViewModels“, o tuo tarpu informacijos atvaizdavimas „Views“ pakete.
- Testuojamumas – Lengva atlikti vienetų testavimą, o kodas priklauso nuo įvykių. MVVM architektūra remiasi įvykiais, todėl duomenų perdavimas vyksta per „stebėjimą“. Pavyzdžiui „views“ klasės stebi savo „viewModel“ klasės duomenų pasikeitimus ir atitinkamai reaguoja į juos.

MVT architektūros įtaka sistemai:

- Palaikomumas – atskyrus skirtingas programos kodo dalis pagal MVT architektūrą, atsiranda kodo struktūra ir vienodumas. Tada nesunku suprasti, kur viskas turėtų būti arba kur jie greičiausiai bus padėti.
- Išplečiamumas – Verslo logika atsieta nuo „UI“. Verslo logika yra „View“, o tuo tarpu informacijos atvaizdavimas „Template“ pakete.

4. Tyrimo dalis

4.1. Tyrimo aprašymas

4.1.1. Kalbos raiškos tyrimas

Skyriuje 2.5 „Atrinkti technologiniai sprendimai tikslui pasiekti“ buvo atrinkta pasirinkti „Google Cloud“ „speech-to-text“ uždavinio tikslui įgyvendinti. Naudojantis šia paslauga buvo siunčiamas naudotojo balsas ir gaunamas pasakytas tekstas bei patikimumo įvertinimas procentais.

Naudojantis gautu naudotojo pasakytu tekstu sistema patikrina ar atpažintas tekstas sutampa su reikiamu tekstu ir įrašo klaidą jeigu nesutampą arba teisingą atsakymą. Su tekstu yra gaunamas ir patikimumo įvertinimas, kuris yra įrašomas kartu su teisingu atsakymu. Klaidingi naudotojų bandymai ir teisingų atsakymų patikimumo įvertinimas yra panaudoti nustatyti ar įmanoma atskirti sveikus ir kalbėjimo sutrikimų turinčius žmones.

4.1.2. Panaudojamumo tyrimas

Sistemos panaudojamumo tyrimui buvo naudojama Sistemos panaudojamumo skalės („System Usability Scale“) metodologija. Ši metodologija yra plačiai naudojama norint nustatyti sistemos panaudojamumą įvertinant sistemos paprastumą, trumpumą ir patikimumą (18). Pasinaudojus sistema kiekvienas respondentas užpildė trumpą klausimyną, kuris paremtas originaliu SUS klausimynu. Klausimyną sudaro 10 klausimų vertinamų penkiabalėje sistemoje, nuo 1 stipriai nesutinku iki 5 stipriai sutinku (19). Nelyginiai klausimai apibūdina teigiamas sistemos savybes, o lyginiai neigiamas. Klausimynas pateiktas 19 paveikslėlyje.

		Srongly Disagree				Srongly Agree
		1	2	3	4	5
1	I think that I would like to use this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2	I found the system unnecessarily complex.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3	I thought the system was easy to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4	I think that I would need the support of a technical person to be able to use this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5	I found the various functions in this system were well integrated.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6	I thought there was too much inconsistency in this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7	I would imagine that most people would learn to use this system very quickly.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8	I found the system very cumbersome to use.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9	I felt very confident using the system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10	I needed to learn a lot of things before I could get going with this system.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

19 pav. Originalus SUS klausimynas (20)

Bendras sistemos įvertinimas apskaičiuojamas pagal formulę pateiktą metodologijoje (20). Rezultatai yra nuo 0 iki 100, tačiau tai nėra procentinė išraiška. Sistemos įvertinimas yra dalinamas į tris dalis:

- 0-64 sistema yra nepriimtina
- 65-84 sistema yra priimtina
- 85-100 sistema yra puiki

4.2. Techninė ir programinė įranga

Tyrimui atlikti buvo naudojama techninė ir programinė įranga aprašyta 11 lentelėje.

11 lentelė Tyrimo įranga

Techninė ir programinė įranga	Versija
„Samsung galaxy“ S9+ telefonas	
Android OS	10
„Google Cloud“ „speech-to-text“ API	v1
„Kotlin“	1.4.20
„Python“	3.9
„Django“	3.2.5

4.3. Tyrimų atlikimo metodika

Tyrimo metu naudotojas turėjo atlikti tris užduotis orientuotas į **r** raidės tarimą ir tris užduotis orientuotas į **š** raidės tarimą. Užduočių eilės tvarka nebuvo nustatyta tačiau užduotys buvo surikiuotos pagal abėcėlę, todėl pirmiau buvo sprendžiamos **r** raidės tarimo užduotys. Kiekvieną užduotį sudaro pratimų rinkinys. Pratimas reikalauja išstarti atitinkamą žodį ar žodžių junginį. Vykdam užduotį yra registruojama:

- Visas užduoties laikas.
- Pratimo bandymo laikas.
- Visi bandymai t. y. teisingi ir neteisingi bandymai išstarti atsakymą.
- Pratimo atsakymo patikimumas gautas iš „Google Cloud“ „speech-to-text“ paslaugos.
- Teisingų bandymų patikimumo suma.

Tyrimo dalyvavo 20 naudotojų iš kurių 15 buvo sveiki žmonės, 3 turintys sutrikimų ir 2 užsieniečiai. Respondentų amžius svyruoja nuo 20 iki 45 metų.

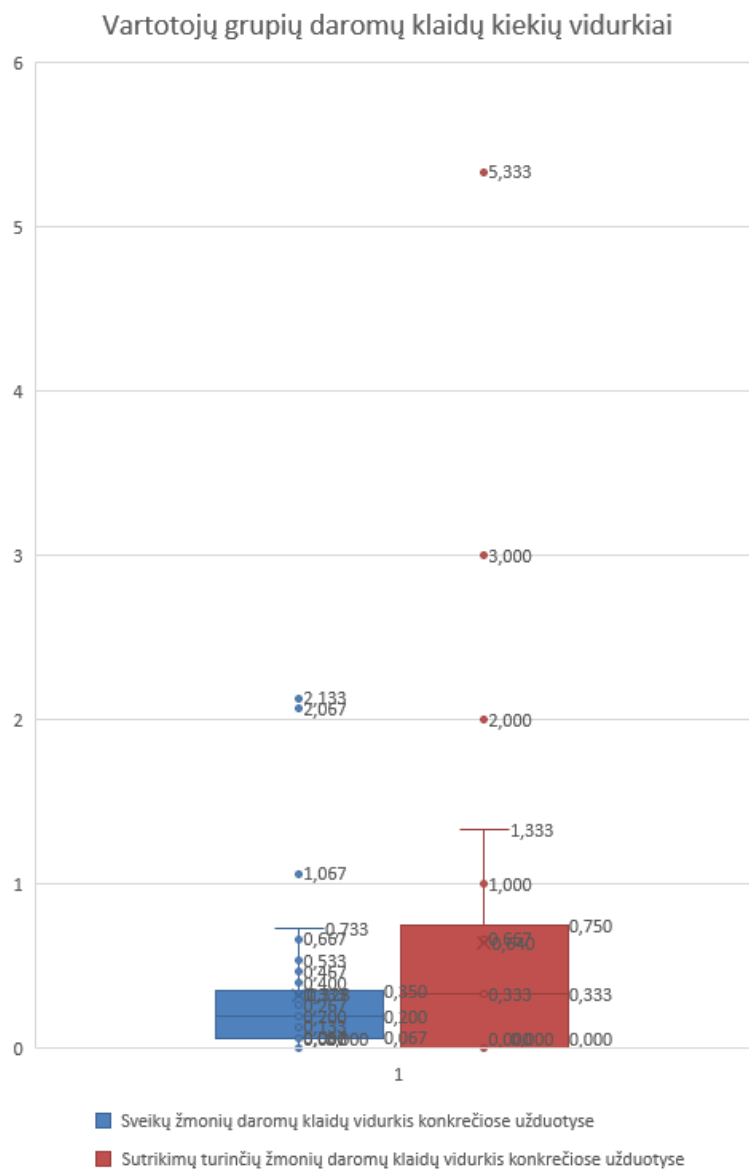
5. Eksperimentinė dalis

5.1. Eksperimentinių tyrimų tikslas

Eksperimentiniu tyrimu siekiama nustatyti ar įmanoma vykdyti žmonių kalbos sutrikimų aptikimą naudojantis „Google Cloud“ „speech-to-text“ paslauga. Sveiki ir sutrikimų turintys naudotojai yra vertinti per daromų klaidų kiekį ir paslaugos patikimumo įvertinimą. Sistemos panaudojamumui iširti yra naudojama sistemos panaudojamumo skalė.

5.2. Daromų klaidų tyrimas

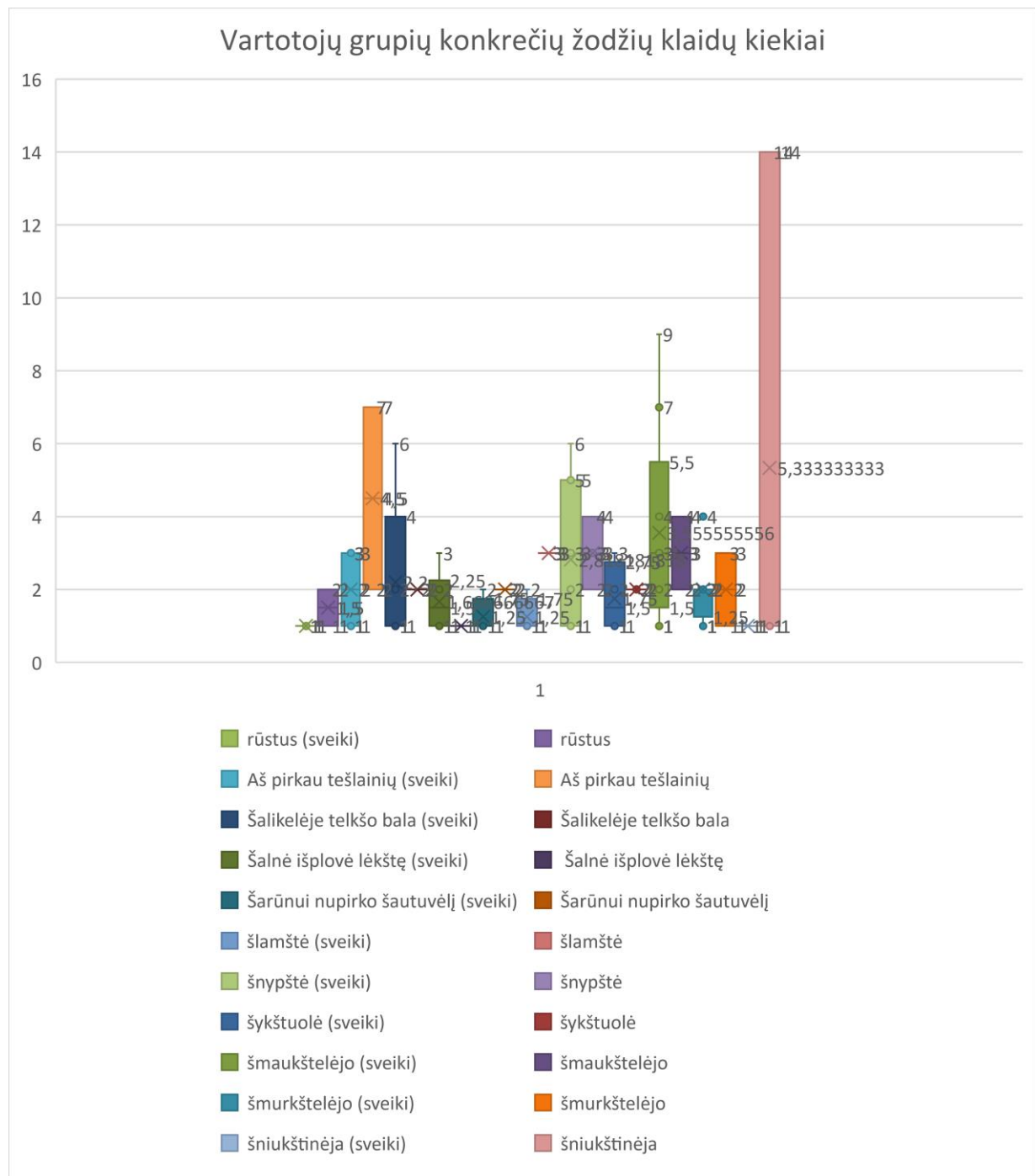
Eksperimento metu sprendžiant užduotis buvo registruojami naudotojų neteisingi atsakymai. Neteisingi atsakymai tai sistemos atpažintas balso tekstas, kuris nesutapo su atsakymu. Klaidingų atsakymų kiekių pasiskirstymas konkrečiame pratime pavaizduotas 20 paveikslėlio diagramoje.



20 pav. Naudotojų grupių daromų klaidų kiekių vidurkiai

Diagramoje pavaizduotos reikšmės yra vidutiniai klaidų kiekiai konkrečiuose užduoties pratimuose. Iš diagramos galime matyti, kad sveiki žmonės daro vidutiniškai mažiau klaidų, nes reikšmės neviršina 2,133 ir yra pasiskirsčiusios tarp 0 ir 0,733. Bendras sveikų žmonių klaidų vidurkis yra 0,328. Aukščiausia žmonių su sutrikimais reikšmė yra 5,333 klaidos per pratimą, o reikšmės pasiskirstę tarp 0 ir 1,333. Ši sritis yra beveik dviguba lyginant su sveikų žmonių sritimi. Sutrikimų turinčių žmonių klaidų vidurkis yra 0,640.

Labiau tyrinėjant daromas klaidas galima panagrinėti konkrečius žodžius. Konkrečiuose žodžiuose daromų klaidų kiekių pasiskirstymas pavaizduotas 21 paveikslėlio diagramoje.



21 pav. Naudotojų grupių konkrečių žodžių klaidų kiekiai

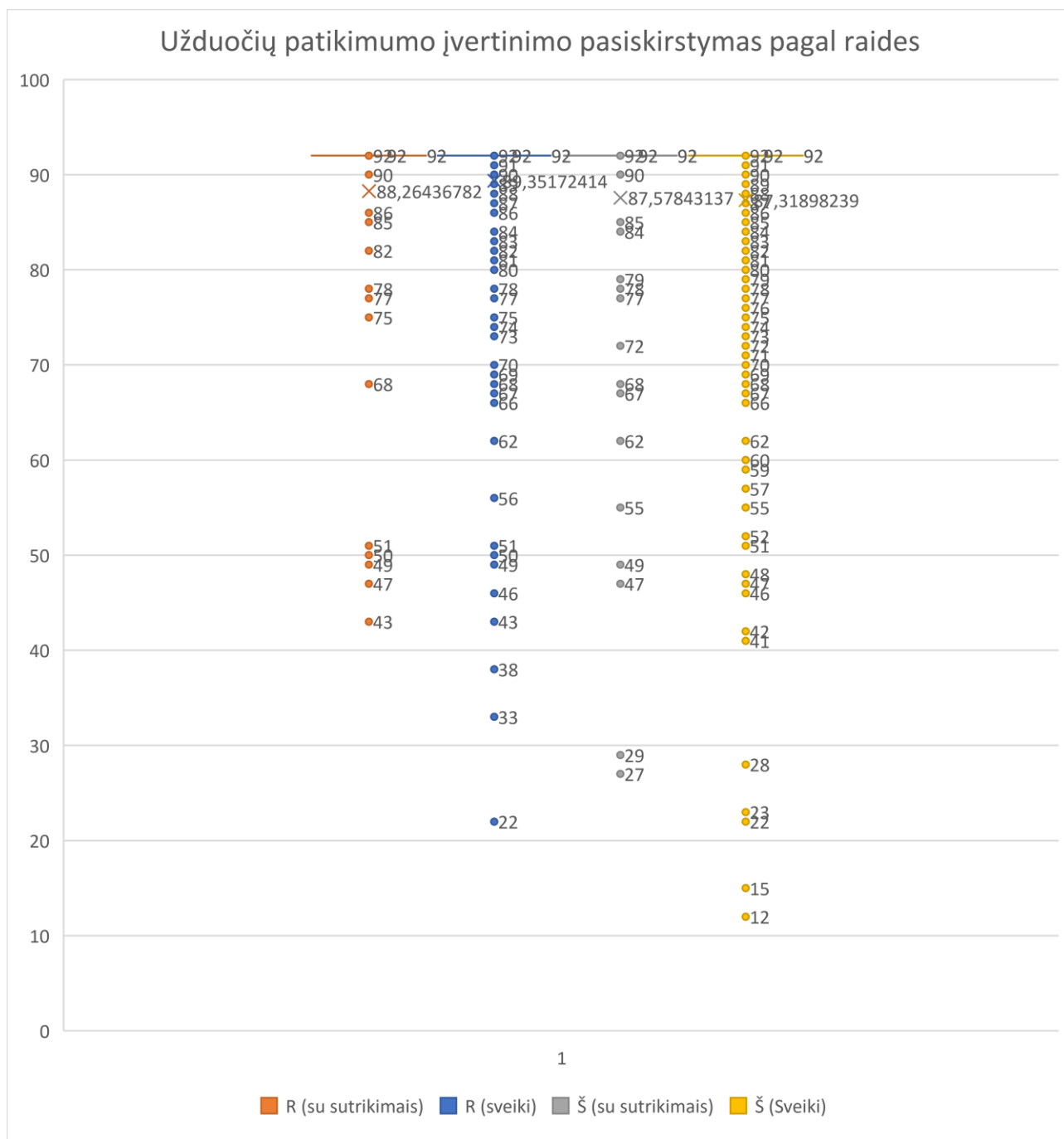
Diagramoje aiškiai išsiskiria stulpeliai su žodžiais „Aš pirkau tešlainių“ ir „šniukštinėja“. Žmonių su sutrikimais daromų klaidų kiekiai kelis kartus viršija žmonių be sutrikimų. Lyginant klaidų vidurkius „Aš pirkau tešlainių“ sveiki žmonės vidutiniškai padaro 2 klaidas, o turintys sutrikimų 4,5 klaidas. „Šniukštinėja“ sveiki žmonės vidutiniškai padaro 1 klaidą, o turintys sutrikimų 5,(3) klaidos.

„Šalikelėje telkšo bala“, „šnypštė“ ir „šmaukštelėjo“ yra pratimai, kuriuose daugiau klaidų padaro sveiki žmonės. Vidurkių palyginimas 2,2 ir 2, 2,(81) ir 3, 3,(5) ir 3 parodo kad vidutiniškai reikšmės nėra toli nutolusios. Panašūs vidurkiai parodo, kad pratimai buvo vienodai sunkūs sveikiems ir sutrikimų turintiems žmonėms.

5.3. Aptiktų žodžių patikimumo tyrimas

Naudotojų užduočių bandymų patikimumas yra vertinamas „Google Cloud“ „speech-to-text“ paslaugos. Jis parodo kaip labai paslauga pasitiki savimi, kad naudotojas pasakė tai ką paslauga atpažino.

Eksperimento metu sprendžiant užduotis buvo registruojamas naudotojų pasakytų atsakymų patikimumas. Atrinkus tik teisingų atsakymus, galima palyginti gautus patikimumo įvertinimus, sugrupavus juos pagal raidės užduotis. 22 paveikslėlyje pavaizduota užduočių teisingų atsakymų patikimumo įvertinimų pasiskirstymas pagal raides.

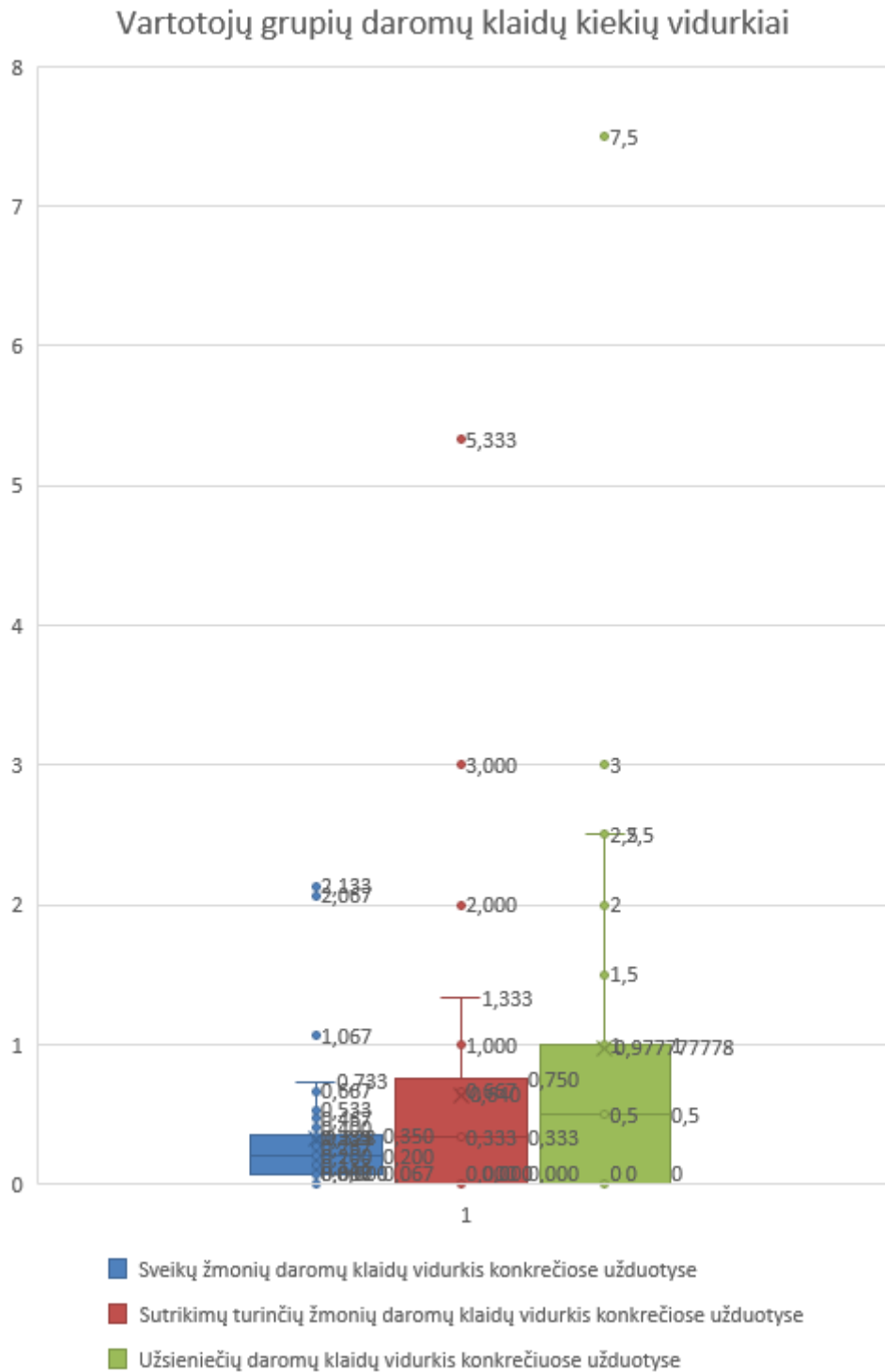


22 pav. Užduočių patikimumo įvertinimo pasiskirstymas pagal raides

Diagramoje aiškiai matyti, kad nei sveiki nei sutrikimų turintys žmonės neviršijo 92 procentų patikimumo ir daugiausia reikšmių yra ties 92 procentų riba. Tikėtina, kad aukštesnių vertinimų paslauga negali suteikti. Stebint patikimumo reikšmes matosi, kad sveikų žmonių įverčiai yra plačiau pasiskirstę, negu žmonių su sutrikimais. Liginant aibių vidurkius matomas panašumas. R raidės sveikų žmonių vidurkis yra 89,352, o žmonių su sutrikimais 88,264, tai yra 1,087 procentu mažiau negu sveikų žmonių. Š raidės sveikų žmonių vidurkis yra 87,396, o žmonių su sutrikimais 87,478, tai yra 0,182 procento daugiau negu sveikų žmonių. Vertinant vidutinius įverčius galima teigti, kad tiek sutrikimų turinčius, tiek sveikus žmones „Google Cloud“ „speech-to-text“ paslauga vertina panašiai, t. y. suteikia panašų patikimumo įvertinimą.

5.4. Užsieniečių respondentų tyrimas

Užsieniečiai respondentai dalyvavo tyrime ir atlikinėjo tokias pačias užduotis, kaip ir kiti respondentai. Sprendžiant užduotis buvo registruojamos jų klaidos kaip ir kitų naudotojų. Neteisingi atsakymai tai sistemos atpažintą balso tekstą, kuris nesutapo su atsakymu. Klaidingų atsakymų kiekių pasiskirstymas konkrečiame pratime, ir jų palyginimas su kitais respondентаis, pavaizduotas 23 paveikslėlio diagramoje.



23 pav. Naudotojų grupių daromų klaidų kiekių vidurkiai

Diagramoje pavaizduotos reikšmės yra vidutiniai klaidų kiekiai konkrečiuose užduoties pratimuose. Iš diagramos aiškiai matyti kad užsieniečių žalias stulpelis lenkia sutrikimų turinčių žmonių stulpelį du kartus, o sveikų žmonių net tris su puse karto. Užsieniečių padaromų klaidų vidurkis yra 0,978 klaidos per užduotį, kai sveiki žmonės padaro 0,328 klaidos, žmonės turintys sutrikimų 0,640 klaidos. Užsieniečiams ištarti lietuviškus žodžius buvo labai sunku.

5.5. Sistemos panaudojamumo tyrimas

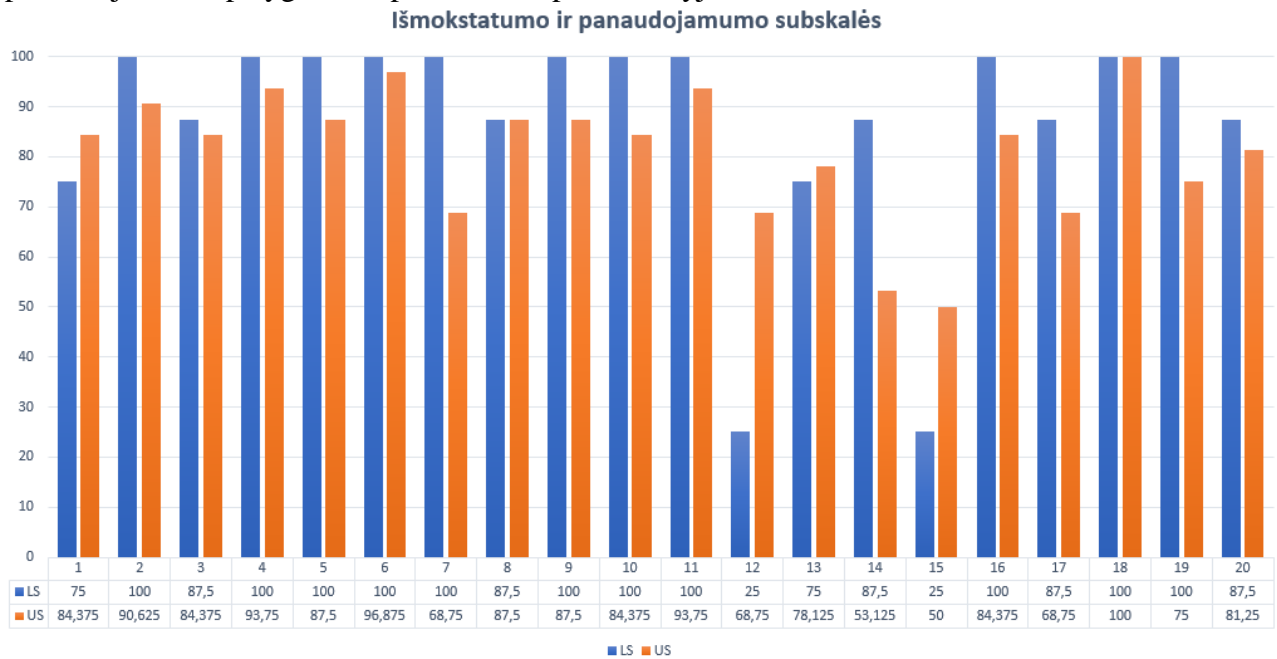
5.5.1. Bendras panaudojamumas

Dvidešimt respondentų užpildė sistemos panaudojamumo skalės klausimyną, kurio rezultatai pateikti lentelė 24 paveikslėlyje. Raudona spalva žymimi neigiami atsiliepimai, geltona neutralūs ir žalia teigiami. Bendras respondento įvertinimas pateikiamas dešinėje lentelės pusėje, o apačioje vidutinis klausimo įvertinimas. Bendras sistemos įvertinimas yra įvertinimų vidurkis ir jis yra 82.125 iš 100. Tai reiškia kad sistema yra priimtina, tačiau dar yra vietų kurias būtų galima patobulinti.

Respondentas	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	Taškai
1	3	1	4	2	4	1	5	1	4	2	82,5
2	4	1	5	1	5	1	4	1	4	1	92,5
3	4	2	3	2	5	1	5	1	4	1	85
4	5	1	4	1	5	1	4	1	5	1	95
5	3	1	5	1	4	1	5	1	4	1	90
6	5	2	5	1	5	1	5	1	5	1	97,5
7	2	3	4	1	4	2	5	3	5	1	75
8	3	1	5	2	4	2	5	1	5	1	87,5
9	3	1	5	1	4	2	5	1	5	1	90
10	3	1	4	1	4	1	4	1	5	1	87,5
11	3	1	5	1	5	1	5	1	5	1	95
12	4	2	3	4	5	3	4	2	3	4	60
13	3	1	5	1	3	1	4	1	3	3	77,5
14	2	2	3	1	3	4	4	2	3	2	60
15	3	4	3	4	3	4	4	3	4	4	45
16	3	1	5	1	4	2	5	1	4	1	87,5
17	3	2	4	1	3	3	5	2	4	2	72,5
18	5	1	5	1	5	1	5	1	5	1	100
19	4	2	4	1	4	3	4	2	5	1	80
20	3	1	5	2	4	3	5	1	4	1	82,5
Taškai	60	86,25	82,5	87,5	78,75	77,5	90	90	82,5	86,25	
										Vidurkis	82,125
Įvertinimo rėžiai											
Nepriimtina	0-64										
Priimtina	65-84										
Tobula	85-100										
Atsakymų rėžiai											
Pozityvus	Sutinka arba stipriai sutinka teigiamiems klausimams. Nesutinka arba stipriai nesutinka neigiamiems klausimams										
Neutralus	Nei sutinka nei nesutinka										
Negatyvus	Sutinka arba stipriai sutinka neigiamiems klausimams. Nesutinka arba stipriai nesutinka teigiamiems klausimams										

24 pav. SUS apklausos atsakymų lentelė

Vertinant sistemos išmokstamumą reikia žiūrėti į 4 ir 10 klausimus, o vertinant panaudojamumą reikia žiūrėti į likusius aštuonis klausimus. Sistemos išmokstamumo ir panaudojamumo palyginimas pateiktas 25 paveikslėlyje.



25 pav. Išmokstamumo ir panaudojamumo palyginimas

Respondentų išmokstamumo taškai dauguma atvejų didesni iš panaudojamumo. Vidutinis išmokstamumo taškų skaičius (86,875), jis taip pat yra didesnis už panaudojamumo taškų vidurkį (80,838). Tai rodo, kad sistema nesunku išmokti naudotis, tačiau panaudojamumas gali būti pagerintas.

5.5.2. Sveikų respondentų sistemos panaudojamumas

Sveikų respondentų sistemos panaudojamumo klausimyno rezultatai pateikti lentelė 26 paveikslėlyje.

Respondentas	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10	Taškai
1	3	1	4	2	4	1	5	1	4	2	82,5
5	3	1	5	1	4	1	5	1	4	1	90
7	2	3	4	1	4	2	5	3	5	1	75
8	3	1	5	2	4	2	5	1	5	1	87,5
9	3	1	5	1	4	2	5	1	5	1	90
10	3	1	4	1	4	1	4	1	5	1	87,5
11	3	1	5	1	5	1	5	1	5	1	95
12	4	2	3	4	5	3	4	2	3	4	60
14	2	2	3	1	3	4	4	2	3	2	60
15	3	4	3	4	3	4	4	3	4	4	45
16	3	1	5	1	4	2	5	1	4	1	87,5
17	3	2	4	1	3	3	5	2	4	2	72,5
18	5	1	5	1	5	1	5	1	5	1	100
19	4	2	4	1	4	3	4	2	5	1	80
20	3	1	5	2	4	3	5	1	4	1	82,5
Taškai	53,33333	85	81,66667	85	75	70	91,66667	86,66667	83,33333	85	
										Vidurkis	79,66667

26 pav. SUS apklausos sveikų respondentų rezultatai

Sveikų respondentų bendras sistemos įvertimas yra 75,(6), o tai yra maždaug 7 procentais mažiau už bendrą įvertinimą. Šie respondentai sistemą įvertino prasčiau tačiau sistema vis dar priimtinos sistemos režiuose.

5.5.3. Respondentų su sutrikimais sistemos panaudojumas

Sutrikimų turinčių respondentų sistemos panaudojamumo klausimyno rezultatai pateikti lentelė 27 paveikslėlyje.

4. Tyrinėjant konkrečių pratimų klaidas, iš pratimų aibės buvo rasti sunkiausi žodžiai sveikiems žmonėms („Šalikelėje telkšo bala“, „šnypštė“ ir „šmaukštelėjo“). Sveikų ir sutrikimų turinčių respondentų vidurkių palyginimas 2,2 ir 2,2,(81) ir 3,3,(5) ir 3.
5. Tyrinėjant „Google Cloud“ „speech-to-text“ paslaugos teikiamą patikimumą nustatyta, kad paslauga neteikia didesnio negu 92 procentų patikimumo ir tiek sveikus, tiek sutrikimų turinčius respondentus vertina panašiai. **R** raidės sveikų žmonių vidurkis yra 89,352, o žmonių su sutrikimais 88,264. **Š** raidės sveikų žmonių vidurkis yra 87,396, o žmonių su sutrikimais 87,478.
6. Tyrinėjant sistemos panaudojamumą buvo nustatyta, kad bendras sistemos įvertinimas yra 82,125. Šis įvertinimas reiškia, kad sistema yra priimtina, tačiau dar yra vietų kurias būtų galima patobulinti.
7. Tyrinėjant sistemos panaudojamumą buvo ištirtas sistemos išmokstamumas ir nustatyta, kad sistemą nesunku išmokti, tačiau panaudojamumas gali būti pagerintas. Išmokstamumo įvertinimas 86,875, o panaudojamumo 80,838.

6. Išvados

1. Skyriuje 2.5 „Atrinkti technologiniai sprendimai tikslui pasiekti“ buvo nustatyta, kad „Google Cloud“ „speech-to-text“ paslauga tinka uždavinio tikslui siekti, nes teikia lietuvių kalbos atpažinimą, bei paprastą integraciją.
2. Programėlei realizuoti buvo pasirinkta MVVM architektūra, nes tai suteiks sistemos kodui struktūrą ir vienodumą, pagerins išplečiamumą ir testuojamumą. Saityno sąsajai kurti pasirinkta MVT architektūra, nes ją teikia naudojamas „Django“ karkasas ir tai pagerina sistemos išplečiamumą bei palaikomumą
3. Naudotojų pratimų atsakymų patikrinimui naudotas Levenšteino atstumas padėjo užskaityti panašius atsakymus į teisingus, pasirinkus lūžio tašku atstumą, kuris mažesnis negu ketvirtadalis teisingo atsakymo teksto ilgio. Tai buvo reikalinga norint užtikrinti, kad naudotojų turinčių sutrikimų atsakymai būtų užskaityti.
4. Tyrinėjant spręstų užduočių klaidas buvo nustatyta, kad sutrikimų turintys žmonės vidutiniškai daro daugiau klaidų negu sveiki žmonės. Sutrikimų turinčių žmonių vidurkis yra 0,640, o sveikų žmonių 0,328, beveik dvigubai mažesnis.
5. Tyrinėjant konkrečių pratimų klaidas buvo rasti sunkiausi žodžiai sutrikimų turintiems žmonėms („Aš pirkau tešlainių“ ir „šniukštinėja“). „Aš pirkau tešlainių“ sveiki žmonės vidutiniškai padaro 2 klaidas, o turintys sutrikimų 4,5 klaidas. „Šniukštinėja“ sveiki žmonės vidutiniškai padaro 1 klaidą, o turintys sutrikimų 5,(3) klaidos.
6. Tyrinėjant konkrečių pratimų klaidas, iš pratimų aibės buvo rasti sunkiausi žodžiai sveikiems žmonėms („Šalikelėje telkšo bala“, „šnypštė“ ir „šmaukštelėjo“). Sveikų ir sutrikimų turinčių respondentų vidurkių palyginimas 2,2 ir 2, 2,(81) ir 3, 3,(5) ir 3.
7. Tyrinėjant „Google Cloud“ „speech-to-text“ paslaugos teikiamą patikimumą nustatyta, kad paslauga neteikia didesnio negu 92 procentų patikimumo ir tiek sveikus, tiek sutrikimų turinčius respondentus vertina panašiai. **R** raidės sveikų žmonių vidurkis yra 89,352, o žmonių su sutrikimais 88,264. **Š** raidės sveikų žmonių vidurkis yra 87,396, o žmonių su sutrikimais 87,478.
8. Tyrinėjant sistemos panaudojamumą buvo nustatyta, kad bendras sistemos įvertinimas yra 82,125. Šis įvertinimas reiškia, kad sistema yra priimtina, tačiau dar yra vietų kurias būtų galima patobulinti.
9. Tyrinėjant sistemos panaudojamumą buvo ištirtas sistemos išmokstamumas ir nustatyta, kad sistemą nesunku išmokti, tačiau panaudojamumas gali būti pagerintas. Išmokstamumo įvertinimas 86,875, o panaudojamumo 80,838.

Aprobavimas:

- Sukurta sistema buvo pristatyta tarptautinėje „Alta 2021“ (ALTA'21) konferencijoje.
- Straipsnis tyrimo tema išspausdintas recenzuojamoje Lietuvos tarptautinės konferencijos ALTA'21 medžiagoje.

7. Literatūra

1. Microsoft Research Blog. *Microsoft*. [Tinkle] 2013 m. 5 14 d. <https://www.microsoft.com/en-us/research/blog/technology-can-bridge-language-gaps/>.
2. *Speech recognition in schools: An update from the field*. Follensbee, Bob; McCloskey-Dale, Susan. s.l. : Technology And Persons With Disabilities Conference, 2000.
3. MassMATCH. Overcoming Communication Barriers. *MassMATCH*. [Tinkle] 2010 m. 3 18 d. [Cituota: 2020 m. 11 15 d.] <https://www.massmatch.org/aboutus/listserv/2010/2010-03-31.html>.
4. *SPEECH RECOGNITION TECHNOLOGY*. Tang, K., R. Kamoua, Victor Sutan, Omer Farooq, G. Eng, Wei Chu and Guofeng Hou. 2, s.l. : J. EDUCATIONAL TECHNOLOGY SYSTEMS, 2004 m., T. 33.
5. *The Infinite Hidden Markov Model*. Beal, MJ. and Ghahramani, Z. and Rasmussen, CE. Masačusetsas : MIT Press, 2002 m.
6. *Phoneme recognition using time-delay neural networks*. A. Waibel, T. Hanazawa, G. Hinton, K. Shikano and K. J. Lang,. 3, s.l. : IEEE Transactions on Acoustics, Speech, and Signal Processing, 1989 m., T. 37.
7. *Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups*. Geoffrey Hinton, Li Deng, Dong Yu, George E. Dahl, Abdelrahman Mohamed, Navdeep Jaitly,. 6, s.l. : IEEE Signal Processing Magazine, 2012 m., T. 29.
8. Li Deng Dong Yu. *Deep Learning: Methods and Applications*. s.l. : Microsoft, 2014.
9. *Large-Scale Visual Speech Recognition*. Brendan Shillingford, Yannis Assael, Matthew W. Hoffman, Thomas Paine, Cían Hughes, Utsav Prabhu, Hank Liao, Hasim Sak, Kanishka Rao, Lorraine Bennett, Marie Mulville, Ben Coppin, Ben Laurie, Andrew Senior, Nando de Freitas. s.l. : DeepMind & Google, 2018.
10. Speech-to-Text. *Google cloud*. [Tinkle] Google. [Cituota: 2020 m. 11 15 d.] <https://cloud.google.com/speech-to-text#section-1>.
11. Speech to Text. *Azure Microsoft*. [Tinkle] Microsoft. [Cituota: 2020 m. 11 15 d.] <https://azure.microsoft.com/en-us/services/cognitive-services/speech-to-text/>.
12. Watson Speech to Text. *IBM*. [Tinkle] IBM. [Cituota: 2020 m. 11 15 d.] <https://www.ibm.com/cloud/watson-speech-to-text>.
13. OneDollar. English Speak Practice. *Google Play*. [Tinkle] Google. [Cituota: 2020 m. 11 15 d.] <https://play.google.com/store/apps/details?id=com.novateam.englishspeak>.
14. takacs.albert. Ello English Learning. *Google Play*. [Tinkle] Google. [Cituota: 2020 m. 11 15 d.] <https://play.google.com/store/apps/details?id=com.takacsalbert.ello>.
15. LLC, SANVADA. Nemokamas Kalbos ugdymas 4-7 m. *Google Play*. [Tinkle] Google. [Cituota: 2020 m. 11 15 d.] <https://play.google.com/store/apps/details?id=com.hedgehogacademy.sequencesfree>.
16. Tamošiūnienė, Aistė. Pradžia. *Logopediniai pratimai*. [Tinkle] Aistė Tamošiūnienė. [Cituota: 2020 m. 11 15 d.] <http://logopediniaipratimai.lt/index.php>.

17. Galubickas. *MIT ir BSD licencijų lietuviški vertimai*. [Tinkle] [Cituota: 2022 m. 01 14 d.] https://galubickas.lt/mit_bsd_lt/.
18. *A Comparison of Questionnaires for Assessing Website Usability*. Tullis, Thomas & Stetson, Jacqueline. 2006 m.
19. Brooke, John. SUS -- a quick and dirty usability scale. [knygos aut.] B. Thomas, Ian Lyall McClelland, Bernard Weerdmeester Patrick W. Jordan. *Usability Evaluation in Industry*. s.l. : Taylor & Francis, 1996.
20. James R. (Jim) Lewis, Ph.D, Jeff Sauro, PhD. Item Benchmarks for the System Usability Scale. *Journal of Usability Studies*. vol 13, 2018 m., T. issue 3.

8. Priedai

1. Originalus straipsnio tekstas

Balso atpažinimo technologijų taikymo tyrimas kalbos sutrikimų lavinimui

Arnastauskas D., Blažauskas T.

Kauno Technologijos Universitetas, Lietuva

Santrauka. Balso atpažinimo technologija turi begale daug pritaikymo sričių. Plačiausiai visiems žinoma yra prietaisų valdymo balsu sritis, tačiau balso atpažinimas gali būti pritaikomas ir edukacinėje veikloje. Ši technologija gali būti panaudota mokantis išarti tam tikrus balsus, ar tiesiog tobulinant kalbos gebėjimus. Tam buvo sukurta ši sistema, kuri pateiks kalbėjimo užduotis ir naudojantis kalbos atpažinimo technologijomis įvertins kalbą. Šiame darbe yra pristatoma sistema ir jos veikimas.

Raktažodžiai: balso atpažinimas, balso vertimas į tekstą, kalbos rehabilitacija.

Įvadas

Kalbos atpažinimas informatikos pasaulyje yra žinomas kaip automatinis kalbos atpažinimas (ASR), kompiuterio kalbos atpažinimas arba kalba į tekstą (STT). Jis yra informatikos ir kompiuterinės lingvistikos pogrupis, kuriame kuriami naujausi algoritmai ir technologijos išversti sakinį į tekstą. Ši sritis apima žinias ir mokslinius tyrimus kompiuterių mokslo, kalbotyros ir kompiuterių inžinerijos srityse.

Kalbos atpažinimo technologijos sparčiai patobulėjo su keliomis naujovių bangomis. Gilaus mašininio mokymosi, didžiųjų duomenų pasiekimai buvo pritaikyti tobulinant kalbos atpažinimą. Apie tai byloja šioje srityje paskelbtų akademinų darbų gausa ir kompanijos, kurios pritaikė įvairius gilaus mokymosi metodus diegiant kalbos atpažinimo sistemas.

Kalbos rehabilitacijos programinė įranga dažniausiai yra skirta žmonėms su klausos sutrikimais, tobulinti jų kalbėjimo gebėjimus. Pradedama nuo paprastų raidžių garsų toliau skiemenų ir pereinant prie sudėtingų spontaniškų žodžių ir frazių. Tokia rehabilitacija yra atliekama su terapeuto pagalba, tačiau programinė įranga leistų savarankišką mokymąsi ir padėtų žmonėms, kurie negali susimokėti už tokias sesijas su terapeutu. Mokantis savarankiškai nereikėtų kontaktuoti su žmonėmis to pasėkoje sumažėtų rizika užsikrėsti virusu.

Kuriama programėlė pateiks užduotis su atitinkamais pratimais ir galimais pratimo atsakymais. Naudotojo pratimo teisingumas bus vertinamas pagal balso atpažinimo patikimumo procentą. Patikimumo procentas yra lygus gaunamų taškų skaičiui už pratimą. Vartotojas galės pažiūrėti savo statistiką, ten matyti kiek kartų bandė užduotį ir jos taškų vidurkį. Administratoriai ir logopedai galės sukurti užduotis ir valdyti vartotojus.

Susiję darbai

Akustikos ir kalbos modeliavimai yra svarbūs šiuolaikinių statistiškai pagrįstų kalbos atpažinimo algoritmų dalys. Kalbos modeliavimas naudojamas daugelyje natūralios kalbos apdorojimo programų, tokių kaip dokumentų klasifikavimas ar statistinis mašininis vertimas. Paslėpti Markovo modeliai (HMM) yra plačiai naudojami daugelyje sistemų. Priežastys, dėl kurių šis metodas tapo toks populiarus, yra būdinga statistinė sistema, mokymo algoritmų, skirtų modelių parametrų iš baigtinių kalbos duomenų mokomųjų rinkinių apskaičiavimo, paprastumas ir prieinamumas. Naudojantis šiuo modeliu gaunama lanksti atpažinimo sistema, kurioje galima lengvai pakeisti modelių dydį, tipą ar architektūrą, kad jie atitiktų tam tikrus žodžius, garsus ir pan. Juang ir Rabiner savo aiškinamajame straipsnyje aptaria statistinių metodų vaidmenį šioje technologijoje, taikomą kalbos atpažinimui ir daugybę teorinių bei praktinių klausimų, kurie dar neišspręsti (1).

Giluminis neuroninis tinklas (DNN) yra dirbtinis neuroninis tinklas, turintis kelis paslėptus vienetų sluoksnius tarp įvesties ir išvesties sluoksnių. DNN architektūros sukuria kompozicinius modelius, kur papildomi sluoksniai suteikia galimybę komponuoti žemesnių sluoksnių ypatybes, suteikdami didžiulį mokymosi pajėgumą ir tokiu būdu potencialą modeliuoti sudėtingus kalbos duomenų modelius. Galimybė modeliuoti sudėtingus duomenų modelius šį metodą padaro daug kartų pranašesnę negu paprastus neuroninius tinklus. Dong Yu ir Li Deng savo knygoje „Automatic Speech Recognition A Deep Learning Approach“ pateikia išsamią naujausios pažangos automatinio kalbos atpažinimo srityje apžvalgą, daugiausia dėmesio skiriant giluminio mokymosi modeliams, įskaitant giluminius neuroninius tinklus ir daugelį jų variantų (2). Tai pirmoji automatinio kalbos atpažinimo knyga, skirta gilus mokymosi metodui. Be griežto matematinio dalyko traktavimo, knygoje taip pat pateikiamos labai sėkmingų giluminio mokymosi modelių įžvalgos ir teorinis pagrindas.

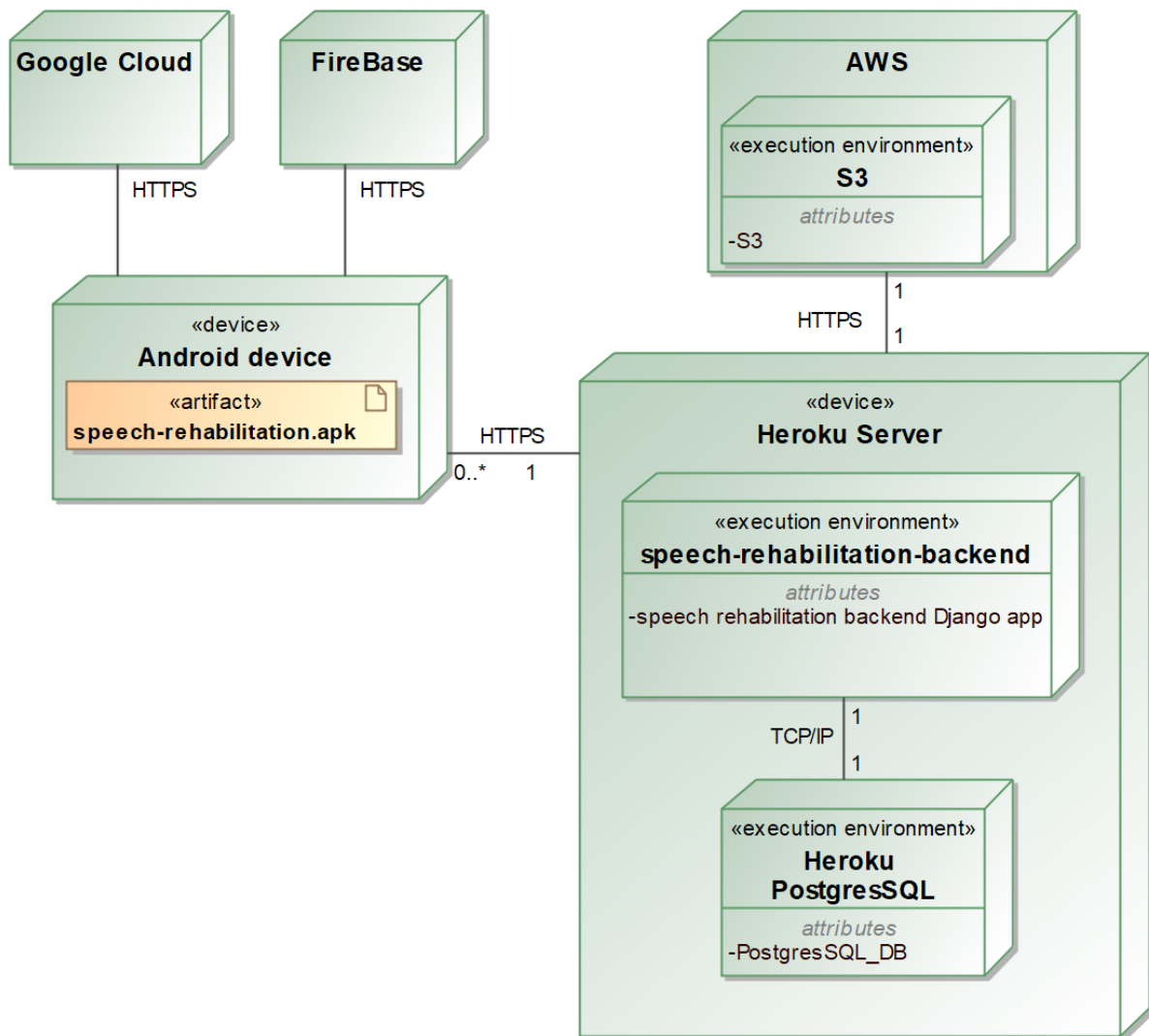
Šiais laikais daug dėmesio yra skiriama giluminiu mokymosi (DL) pagrįstais semantinės komunikacijos metodams, skirtiems efektyviam vaizdų, teksto ir kalbos perdavimui. Skirtingai nuo tradicinių belaidžio ryšio metodų, kuriuose pagrindinis dėmesys skiriamas abstrakčių simbolių perdavimui, semantinio ryšio metodais siekiama geresnio perdavimo efektyvumo siunčiant tik semantinę šaltinio duomenų informaciją. „Semantic-aware Speech to Text Transmission with Redundancy Removal“ straipsnyje nagrinėjamas į semantiką orientuotas kalbos vertimas į tekstą (3). Šis metodas išskiria tik su tekstu susijusias semantines ypatybes, taip dar labiau sumažindamas semantinio pertekliaus kiekį, palyginti su esamais metodais. Modeliavimo rezultatai rodo, kad siūlomas metodas pranoksta dabartinius metodus gauto teksto tikslumu ir perdavimo efektyvumu. Be to, siūlomas metodas taip pat turi mažesnę modelio dydį ir trumpesnę veikimo laiką.

Balso atpažinimo technologijų taikymo kalbos sutrikimų lavinimui

Sistemą sudaro Android programėlė ir saityno sąsaja. Android programėlės sukurta naudojantis Kotlin programavimo kalba, kuri leidžia sumažinti nereikalingo kodo kiekį palyginus su Java kalba. Žemiausia palaikoma Android versija pasirinkta penkta, tai užtikrins, kad programėlė

galės veikti beveik visuose telefonuose. Saityno sąsajai buvo sutelkta Python programavimo kalba ir Django karkasas. Šis karkasas labai palengvina standartinių sukūrimo, redagavimo, ištrynimo ir peržiūros operacijų kūrimą.

Sistema turi daug integracijų. Android programėlė komunikuoja su Google Cloud speech-to-text servisu gauti balso išversto į tekstą rezultatams. Su FireBase komunikuoja saugoti nenumatytoms klaidoms ir vartotojų naudojimosi informacijai. Su saityno sąsaja komunikuoja duomenims gauti ir saugoti. Saityno sąsaja komunikuoja su PostgrSQL duomenų baze duomenims gauti ir saugoti. Taip pat komunikuoja su AWS S3 servisu paveikslėliams talpinti. Visa tai pavaizduota paveikslėlyje Nr. 1.



28 pav. Sistemos išdėstymo diagrama

Prisijungus prie programėlės bus matomas pagrindinis langas su užduotimis pavaizduotas paveikslėlyje Nr. 2. Iš pagrindinio lango galima patekti į sesijų statistiką arba pasirinkus užduotį patekti į užduoties langą. Norint peržiūrėti sesijų informaciją reikia paspausti mygtuką „Mano Sesijos“. Norint pradėti spręsti užduotį reikia paspausti ant vienos užduoties iš sąrašo.



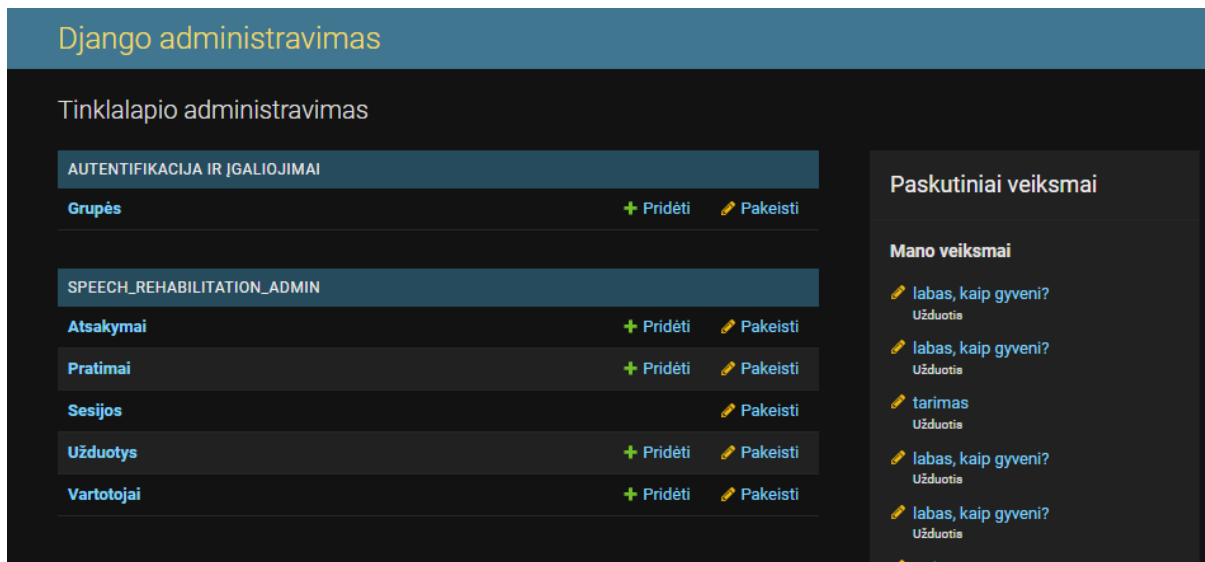
29 pav. Programėlės pagrindinis langas

Užduoties lange pateikiama užduotis ir galimi atsakymai. Iš užduoties lango galima patekti tik į pagrindinį langą. Norint pradėti užduotį reikia įsijungti mikrofoną ir šnekėti atsakymus t. y. paspausti ant mikrofono ikonėlės ir kalbėti. Mikrofono ikonėlė pasidarys aktyvi ir paryškės tekstas „Klausome...“.

Sesijų statistikos lange matoma vidutinė užduočių statistika. Pateikiamas spęstų užduočių sąrašas su informacija kiek kartų buvo bandyta užduotis ir koks yra taškų vidurkis.

Užduočių sprendimas ir sesijų statistikos peržiūros yra svarbiausios sistemos funkcijos ir jos atliekamos tik iš programėlės.

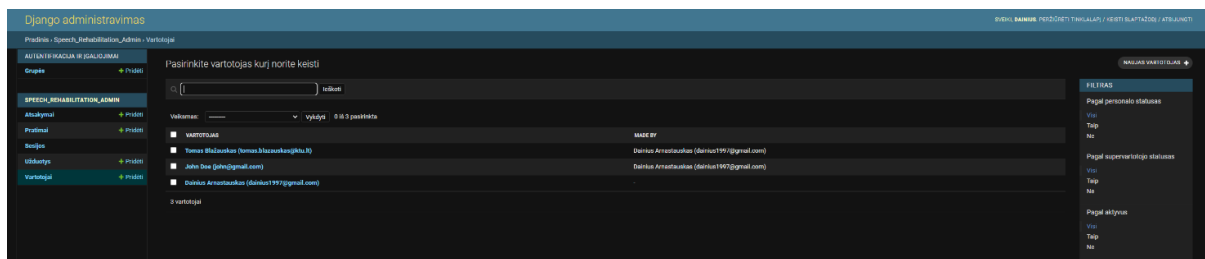
Norint valdyti vartotojus, užduotis ir sesijas reikia naudotis saityno sąsaja pavaizduota paveikslėlyje Nr 3.



30 pav. Saityno sąsaja

Administravimo pagrindiniame lange galima matyti paskutinius atliktus veiksmus ir su vartotojais bei užduotimis susijusius objektų tipus. Šis lango meniu suteikia galimybę, sukurti naują vartotoją ar užduotį ir peržiūrėti visus vartotojus ar užduotis.

Objektų sąrašas pateikia tam tikro tipo objektus puslapiais po 50 elementų. Jei yra numatyta sąrašas turi paiešką filtrus ir veiksmus. Iš sąrašo galima pasirinkti norimus ištrinti objektus arba išsirinkti norimą redaguoti objektą. Esant sąrašui lieka ir administratoriaus lango sumažintas meniu. Filtrai talpinami dešinėje o paieška viršuje. Susižymėjus elementus galima išsiskleisti veiksmų sąrašą, pasirinkti veiksmą ir jį įvykdyti. Objektų sąrašo pavyzdys pavaizduotas paveikslėlyje Nr. 4.



31 pav. Objektų sąrašas

Išvados

Balso atpažinimo technologijos turi daug pritaikymo galimybių ir yra nuolat tobulinamos. Nepaisant to ši technologija nėra plačiai taikoma mokantis išstarti tam tikrus balsus, ar tiesiog tobulinant kalbos gebėjimus. Šiauriant pasauliniai pandemijai svarbu kuo mažiau turėti kontakto su žmonėmis. Sukurta programinė įranga leidžia savarankišką mokymąsi ir padėtų žmonėms, kurie negali susimokėti už tokias sesijas su terapeutu. Mokantis savarankiškai nereikėtų kontaktuoti su žmonėmis to pasėkoje sumažėtų rizika užsikrėsti virusu.

Nuorodos

1. **B. H. Juang & L. R. Rabiner.** Hidden Markov Models for Speech Recognition. *Technometrics*. 33, 1991 m., T. 3.
2. **Dong Yu, Li Deng.** *Automatic Speech Recognition: A Deep Learning Approach*. London : Springer, 2016. 9781447169673.
3. *Semantic-aware Speech to Text Transmission with Redundancy Removal.* **Tianxiao Han, Qianqian Yang, Zhiguo Shi, Shibo He, Zhaoyang Zhang.** Hangzhou : Zhejiang University, 2022 m.