



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas

Socialiniuose tinkluose migrantų bei ekspatriantų ieškomos ir jiems teikiamos pagalbos identifikavimas

Baigiamasis magistro studijų projektas

Goda Armalytė

Projekto autorė

Prof. dr. Vilmantė Kumpikaitė-Valiūnienė

Vadovė

Doc. dr. Mantas Landauskas

Vadovas

Kaunas, 2022



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas

Socialiniuose tinkluose migrantų bei ekspatriantų ieškomos ir jiems teikiamos pagalbos identifikavimas

Baigiamasis magistro studijų projektas
Didžiųjų verslo duomenų analitika (6213AX001)

Goda Armalytė

Projekto autorė

Prof. dr. Vilmantė Kumpikaitė-Valiūnienė

Vadovė

Doc. dr. Liepa Bikulčienė

Recenzentė

Doc. dr. Mantas Landauskas

Vadovas

Doc. dr. Kęstutis Duoba

Recenzentas

Kaunas, 2022



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas
Goda Armalytė

Socialiniuose tinkluose migrantų bei ekspatriantų ieškomos ir jiems teikiamos pagalbos identifikavimas

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.

Goda Armalytė

Patvirtinta elektroniniu būdu

Armalytė, Goda. Socialiniuose tinkluose migrantų bei ekspatriantų ieškomos ir jiems teikiamos pagalbos identifikavimas. Magistro studijų baigiamasis projektas / vadovė prof. dr. Vilmantė Kumpikaitė-Valiūnienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų kryptių grupė): Taikomoji matematika.

Reikšminiai žodžiai: migrantai, ekspatriantai, socialinė parama, teksto klasifikavimas.

Kaunas, 2022. 60 p.

Santrauka

Vykstant pastoviai migracijai visame pasaulyje, vis svarbesnis uždavinys yra ekspatriantų bei kitų migrantų tinkamos emocinės būklės užtikrinimas. Iki šiol atlikti tyrimai rodo, jog persikeldami į naują gyvenamąją vietą ekspatriantai bei kiti migrantai patiria stresą, susiduria su emocinėmis problemomis, kaip liūdesys ar depresija. Taigi, ekspatriantams ir bendrai migrantams ypatingai svarbu gauti socialinę paramą. Socialinė parama gali būti teikiama migranto šeimos narių, bendradarbių, ar kitų asmenų migranto aplinkoje. Nors įprasta, jog socialinė parama dažniausiai suteikiama pažįstamų asmenų gyvai, vis dažniau migrantai paramos ieško ir ją suteikia internete, įvairiose socialinių tinklų platformose.

Baigiamojo magistro darbo tyrime nagrinėjamos žinutės iš viešo ekspatriantų forumo. Naudojant matematinis duomenų klasifikavimo algoritmus identifikuojami žinutėse išreikšti socialinės paramos tipai, sukuriamas sintetinis duomenų rinkinys tolimesnei analizei atlikti. Duomenų paruošimo bei analizės eigoje taikomi teksto valymo, rekonstravimo, klasterizavimo bei klasifikavimo metodai.

Visi tyrime naudoti klasifikatoriai turėjo stiprybių ir silpnybių, tačiau realioje duomenų aibėje geriausi rezultatai buvo būdingi atraminių vektorių mašinos bei Naive Bayes algoritmams. Tyrimo rezultatai rodo, jog tyrime pasirinktame forume asmenys dažniausiai ieško informacinės paramos. Dažniausiai būtent informacinė pagalba jiems yra suteikiama. Emocinė parama ieškoma gana retai, o instrumentinės paramos ieškojimo atvejai nebuvo užfiksuoti. Statistiškai reikšmingi skirtumai tarp informacinės ir emocinės paramos imčių vertinant paramos suteikimo ir padėkos lygius nebuvo nustatyti. 93,5 proc. atsakymų suteikė paramą, kuri atitiko įrašuose prašomos paramos tipą.

Armalytė, Goda. Identification of the support sought by and provided to migrants and expats on social networks. Master's Final Degree Project / supervisor prof. dr. Vilmantė Kumpikaitė-Valiūnienė; Faculty of Mathematics and Natural Sciences, Kaunas University of Technology.

Study field and area (study field group): Applied Mathematics.

Keywords: migrants, expatriates, social support, text classification.

Kaunas, 2022. 60 p.

Summary

In the context of constant migration around the world, ensuring proper emotional state of migrants is an increasingly important challenge. Studies completed to this date show that expatriates and other migrants experience stress and face emotional problems such as sadness or depression when moving to a new place of residence. Thus, it is particularly important for expatriates and migrants in general to receive social support. Social support can be provided by family members of the migrant, co-workers, or other people in the migrant's environment. Although it is common for social support to be provided by acquaintances live, migrants are increasingly seeking and providing social support online, on a variety of social networking platforms.

The research of the final master's thesis examines messages from a public expatriate forum. Mathematical data classification algorithms are used to identify the types of social support expressed in the messages and to create a synthetic data set for further analysis. During data preparation and analysis, text cleaning, embedding, clustering and classification methods are applied.

All classifiers used in the study had their strengths and weaknesses, but in the real data set, the best results were shown by support vector machine and Naïve Bayes algorithms. The results of the study show that individuals are most likely to seek informational support in the forum selected for the research. Usually, informational support is given to them as well. Emotional support is sought relatively infrequently, while cases of seeking instrumental support have not been identified. No statistically significant differences were found between the informational and emotional support data samples in accessing support and appreciation levels. 93.5% of responses provided support that matched the type of support requested in the initial post.

Turinys

Turinys	6
Paveikslų sąrašas	7
Lentelių sąrašas	8
Įvadas	9
1. Paramos migrantams, teikiamos socialinių tinklų pagalba, literatūros analizė	10
1.1. Migrantų sampratos ir klasifikacijos	10
1.2. Migrantų adaptacijos aktualumas	13
1.3. Parama migrantams	14
1.3.1. Paramos tipai	14
1.3.2. Paramos teikėjai.....	15
1.4. Augantis socialinių tinklų vaidmuo migrantų adaptacijos ir paramos procese	19
1.5. Konceptualus tyrimo modelis	20
2. Matematinų algoritmų literatūros analizė	22
2.1. Duomenų rinkimas	22
2.2. Teksto valymas	22
2.3. Teksto rekonstravimo metodai	22
2.4. Teksto analizės algoritmai.....	24
2.4.1. K-artimiausių kaimynų metodas.....	25
2.4.2. Naive Bayes metodas	25
2.4.3. Atraminių vektorių mašinos metodas.....	26
2.4.4. Palyginimas	27
3. Tyrimo metodologija ir duomenų apžvalga	29
3.1. Duomenų rinkimas	29
3.2. Teksto valymas ir paruošimas	30
3.3. Pirminė teksto analizė.....	31
3.4. Teksto integravimas.....	37
3.5. Duomenų klasifikavimas	38
4. Tyrimo rezultatai	41
4.1. Paramos tipų identifikavimas pasitelkiant matematinius algoritmus	41
4.1.1. Įrašų klasifikavimas pagal prašomos paramos tipą	42
4.1.2. Atsakymų klasifikavimas pagal suteikiamos paramos tipą	43
4.2. Sintetinių duomenų rinkinių analizė.....	47
4.2.1. Sintetinio įrašų duomenų rinkinio analizė	47
4.2.2. Sintetinio atsakymų duomenų rinkinio analizė	52
4.3. Tyrimo apibendrinimas ir diskusija.....	53
Išvados	55
Literatūros sąrašas	56

Paveikslų sąrašas

1 pav. Migrantų (ekspatriantų) sprendimų medis (sudaryta remiantis Andresenu, Bergdoltu, Margenfeldu ir Dickannu [5]).....	12
2 pav. Dalis imigrantų, susiduriančių su emocinėmis problemomis, pagal lytį ir tipą, 2012 m., proc. (sudaryta remiantis Robertu ir Gilkinsonu [56])	13
3 pav. Organizacijos paramos tipai pagal paramos teikėjus (sudaryta remiantis Simosi [61], Hu ir kt. [32] bei Wee Chan ir Pervaizu [71]).	16
4 pav. Imigrantų gaunamos socialinės pagalbos dažnis ir pasitenkinimas gauta pagalba, pagal tipus ir paramos teikėjus, 2019 m. (sudaryta remiantis Hombrados-Mendieta ir kt. [30])	18
5 pav. Tyrimo struktūra.....	21
6 pav. Formulės termino dažnio – atvirkštinio dokumento dažnio apskaičiavimui (Vogleris [70])	23
7 pav. Pastovaus žodžių maišo bei Šokinėjimo modelių architektūra (Mikolovas ir kt. [46])	24
8 pav. Atraminių vektorių mašinos metodo (SVM) su kieta ir minkšta parašėmis palyginimas.....	27
9 pav. Žodžių skaičiaus žinutėse histograma	29
10 pav. Įrašų žodžių debesis	31
11 pav. Dažniausiai įrašuose naudojamų žodžių horizontali stulpelinė diagrama.....	32
12 pav. Įrašų žodžių klasteriai.....	33
13 pav. Atsakymų žodžių debesis	34
14 pav. Dažniausiai atsakymuose naudojamų žodžių horizontali stulpelinė diagrama	35
15 pav. Atsakymų žodžių klasteriai	36
16 pav. Teksto integravimo pavyzdys.....	38
17 pav. Įrašų bei atsakymų duomenų rinkinių pokyčių schema	41
18 pav. Įrašų pasiskirstymas pagal prašomos paramos tipą, proc.	47
19 pav. Įrašų pasiskirstymas remiantis kitų asmenų paliktais komentarais, proc.	48
20 pav. Vidutinis vieno įrašo sulauktų atsakymų skaičius.....	49
21 pav. Įrašų pasiskirstymas pagal įrašų autorių išreikštą padėką, proc.	51
22 pav. Prašomos ir suteikiamos paramos tipų sutapimas	52

Lentelių sąrašas

1 lentelė. Tarptautinių migrantų tipai (sudaryta remiantis Sundari ir Pearson [63])	10
2 lentelė. Algoritmų palyginimas.....	27
3 lentelė. Įrašų terminų dokumentų matrica.....	30
4 lentelė. Dešimties dažniausiai pasikartojančių žodžių palyginimas tarp įrašų ir atsakymų.....	35
5 lentelė. Įrašų žodynas	37
6 lentelė. K-artimiausių kaimynų algoritmo apmokymo imties prognozių tikslumas	39
7 lentelė. Sumaišymo matricos pavyzdys (sudaryta remiantis Zhu, Zengu ir Wangu [75]).....	40
8 lentelė. Algoritmų tikslumo prognozuojant įrašuose prašomos paramos tipą palyginimas.....	42
9 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant įrašuose prašomos paramos tipą palyginimas.....	42
10 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas įrašams	43
11 lentelė. Algoritmų tikslumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas	43
12 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas.....	43
13 lentelė. Algoritmų tikslumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas (sulyginus klasių balansą apmokymo ir testavimo imtyse).....	44
14 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas (sulyginus klasių balansą apmokymo ir testavimo imtyse)	44
15 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas atsakymams.....	44
16 lentelė. Algoritmų tikslumo prognozuojant paramos tipą bendroje imtyje palyginimas	45
17 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant paramos tipą bendroje imtyje palyginimas.....	45
18 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas įrašams ir atsakymams	46
19 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas įrašams ir atsakymams	46
20 lentelė. Įrašų be atsakymo įverčio palyginimas tarp emocinės ir informacinės paramos prašymų	48
21 lentelė. Atsakymų vidurkio palyginimas tarp emocinės ir informacinės paramos prašymų.....	50
22 lentelė. Padėkų vidurkio palyginimas tarp emocinės ir informacinės paramos prašymų	51

Įvadas

Darbo aktualumas:

Nuo pirmosios globalizacijos bangos, įvykusios 19 a. pabaigoje – 20 a. pradžioje, migracijos apimtys vis didėja. Prie to prisideda įvairūs veiksniai, kaip pingantis transportas, prastos sąlygos migrantų gimtojoje šalyje ar labai patrauklios sąlygos migrantą priimančioje valstybėje. 2020 m. tarptautinių migrantų skaičius siekė 281 mln. asmenų, jie sudarė 3,6 proc. pasaulio populiacijos [65].

Persikėlimas į naują šalį migrantams gali sukelti daug streso dėl poreikio prisitaikyti prie naujos kultūros, bendruomenės, darbo ir kitų naujos gyvenamosios vietos aspektų [20]. Taigi, ekspatriantams ir bendrai migrantams ypatingai svarbu gauti socialinę paramą iš juos supančių asmenų, pavyzdžiui, šeimos, bendradarbių arba specialiai paramai migrantams teikti įkurtų organizacijų. Plečiantis interneto teikiamoms galimybėms ir tobulėjant socialiniams tinklams, socialinė pagalba vis dažniau yra suteikiama ir internete. Asmenys socialiniuose tinkluose dalijasi naudinga informacija, savo patirtimis, džiaugsmais ir nusivylimais. Socialinė parama, teikiama socialinių tinklų platformose, gali prisidėti prie savižudybės rizikos mažinimo [15], didinti pasitenkinimą gyvenimu [52], mažinti vienatvės jausmą ir depresiją [60]. Kaip ir kiti gyventojai, migrantai bei ekspatriantai vis dažniau pagalbos ieško būtent socialinių tinklų puslapiuose. Tačiau ieškoma ir gaunama parama gali skirtis priklausomai nuo paramos ieškančio ir paramą teikiančio asmens bei paramos teikimo formos. Taigi, pagrindinis probleminis klausimas, nagrinėjamas šiame magistro darbe, yra kokios paramos migrantai bei ekspatriantai ieško ir kokia parama jiems yra suteikiama socialinių tinklų pagalba.

Darbo tikslas:

Identifikuoti kokios paramos migrantai ieško ir kokia parama yra teikiama socialiniuose tinkluose.

Darbo uždaviniai:

1. Atlikus mokslinės literatūros analizę, nustatyti, kokia parama gali būti teikiama migrantams bei ekspatriantams.
2. Teoriškai atskleisti socialinių tinklų svarbą migrantų adaptacijos procese.
3. Parengti konceptualų modelį skirtą identifikuoti ekspatriantams (migrantams) teikiamą paramą socialiniuose tinkluose.
4. Pasirinkti tinkamus matematinius metodus ir algoritmus, skirtus identifikuoti ekspatriantams (migrantams) teikiamą paramą socialiniuose tinkluose, atsižvelgiant į turimus duomenis.
5. Atlikti socialinių tinklų teksto analizę ekspatriantų (migrantų) ieškomos ir teikiamos paramos identifikavimui.
6. Palyginti keletą matematinių algoritmų tarpusavyje ir pasirinkti geriausią remiantis įvairiais kriterijais, kaip tikslumas, jautrumas ir pan.

1. Paramos migrantams, teikiamos socialinių tinklų pagalba, literatūros analizė

Literatūros analizės dalyje apžvelgiami skirtingi migrantų apibrėžimai bei tipai, parama ir jos tipai, paramos tiekėjai ir socialiniai tinklai bei jų vaidmuo migrantų adaptacijoje. Atlikus literatūros analizę konkretizuojamas pagrindinis tyrimo tikslas ir pateikiamas konceptualus modelis tyrimo atlikimui.

1.1. Migrantų sampratos ir klasifikacijos

Tarptautinė migracijos organizacija terminą migrantai apibrėžia kaip platų terminą, kuris nėra apibrėžtas teisėje, tačiau yra suprantamas kaip žmogus, kuris išsikrausto iš savo įprastos gyvenamosios vietos [64]. Šiame darbe dėmesys bus skiriamas konkrečiai tarptautiniams migrantams, kurie apibrėžiami kaip bet kurie asmenys, kurie keičia savo gyvenamosios vietos šalį [37]. Naujojoje gyvenamojoje valstybėje asmuo turi turėti gyvenamąją vietą ir leisti toje valstybėje laiką ne tik darbo, bet ir poilsio metu, taigi į tarptautinių migrantų kategoriją neįtraukiami asmenys, į užsienio valstybes keliaujantys trumpam laikui rekreaciniais, verslo, gydymosi ar religiniais tikslais (ibid.). Migrantai taip pat gali būti atskiriami pagal jų tipus. Paprastai išskiriami 4 pagrindiniai migrantų tipai: ekonominiai migrantai, politiniai migrantai, migruojantys dėl aplinkos ir migruojantys dėl šeimos suvienijimo. Kiekviena iš šių grupių turi savo apibrėžimą, migranto tipą gali nulemti migracijos motyvacija ir kontekstas (žr. 1 lentelė).

1 lentelė. Tarptautinių migrantų tipai (sudaryta remiantis Sundari ir Pearson [63])

Ekonominiai migrantai	Politiniai migrantai	Aplinkos migrantai	Šeimos suvienijimo migrantai
Ekonominiai migrantai keičia gyvenamosios vietos šalį dėl geresnių ekonominių galimybių priimančiojoje šalyje.	Politiniai migrantai yra priversti keisti gyvenamosios vietos šalį dėl vykstančių karų, persekiojimo baimės arba valstybinės politikos, diskriminuojančios žmones, kurie priešinasi valdantiesiems.	Aplinkos migrantai yra priversti keisti gyvenamosios vietos šalį dėl vietinės aplinkos pokyčių, kurie neigiamai veikia jų gerovę arba pragyvenimo šaltinius.	Šeimos suvienijimo migrantai keičia gyvenamosios vietos šalį tam, kad prisijungtų prie savo artimųjų, gyvenančių priimančiojoje šalyje.

Kiekvienai iš šių grupių dažniausiai aktuali skirtinga parama. Šiame darbe bus nagrinėjamos paramos formos, labiausiai aktualios darbo migrantams, todėl toliau nagrinėjami ekonominių migrantų bei ekspatriantų apibrėžimai ir skirtingos galimos šių grupių charakteristikos.

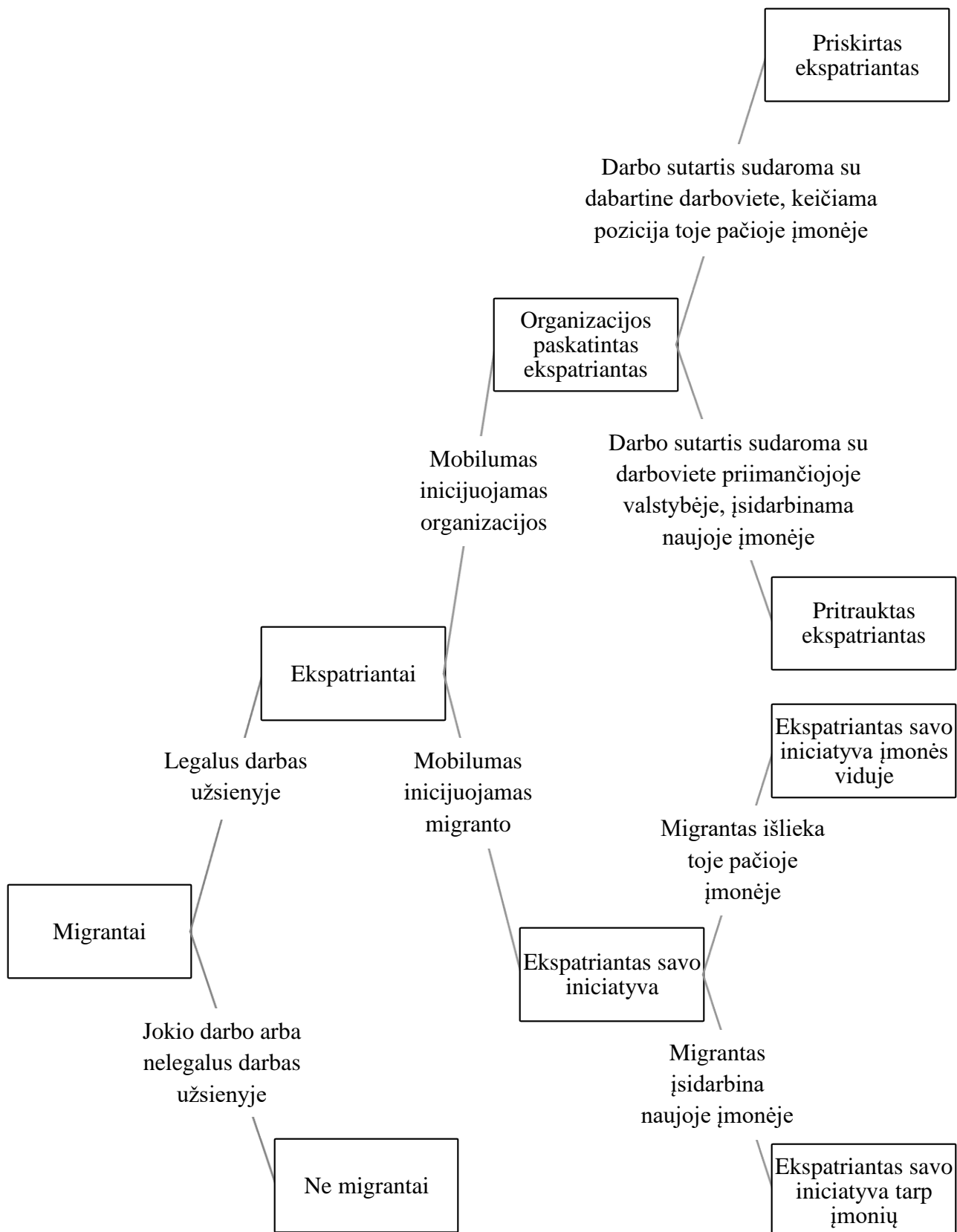
Ekonominiai, arba dar vadinami darbo migrantai, taip pat gali būti skirstomi į kelias skirtingas grupes. Remiantis dualios arba segmentuotos darbo rinkos teorija, darbo rinką galima išskirti į dvi dalis: pirminė darbo rinka, kurios darbuotojai yra gerai apmokami, plačios kompetencijos, dažniausiai yra stabilūs, ir antrinė darbo rinka, kurioje darbuotojai yra prastai apmokami, jiems būdingi žemi įgūdžiai, dažniausiai yra žemo statuso visuomenėje [6]. Dėl hierarchinės infliacijos (fenomeno, kai pakėlus atlyginimus žemų įgūdžių darbuotojams, turi kilti visi atlyginimai, kad būtų išlaikoma visuomenės struktūra), išsivysčiusios šalys siekia pritraukti žemų įgūdžių migrantus, kuriems atlyginimo kelti nereikėtų. Nors antrinėje darbo rinkoje darbai yra menkai apmokami ir neturi prestižo, migrantams, atkeliaujantiems iš mažiau išsivysčiusių šalių, priimančiojoje šalyje gaunamas atlyginimas gali reikšmingai pagerinti jų ir kai kuriais atvejais jų šeimos narių padėtį jų gimtojoje šalyje. Tuo tarpu pirminė darbo rinka pagal dualios darbo rinkos teoriją dažniausiai pritraukia vietinius gyventojus. Tačiau dualios darbo rinkos teorija nepaaiškina aukštos kvalifikacijos ekonominių migrantų

egzistavimo. Tai geriau paaiškina neoklasikinė migracijos teorija, kuri teigia, kad asmenys turi visą informaciją apie atlyginimus ir įsidarbinimo galimybes priimančiose šalyse ir priima sprendimus pasverdami savo asmenines sąnaudas ir naudas priimant vieną ar kitą sprendimą dėl migracijos [45]. Pagal tokį modelį galėtų vykti ir žemos, ir aukštos kvalifikacijos asmenų migracija, atsižvelgus ne tik į atlyginimų skirtumus, bet ir į asmenų tikimybę rasti darbą gimtojoje ir priimančioje valstybėse, migracijos kainą ir pan. Tuo tarpu nauja ekonominė migracijos teorija atsižvelgia ne tik į atlyginimus, bet ir į rizikos valdymą, taigi ji irgi galėtų paaiškinti ir žemos ir aukštos kvalifikacijos asmenų migraciją į valstybes, kuriose jie įgyja didesnę finansinę saugumą (dėl esamų nedarbingumo išmokų, senatvės pensijos, stabilų kredito rinkų ir pan.).

Dar vienas svarbus terminas darbo kontekste yra ekspatriantai. Ekspatriantais vadinami laikinai, bet dažniausiai ne trumpai užsienyje dirbantys asmenys, paprastai tarptautinių įmonių darbuotojai, dirbantys ne gimtojoje šalyje. Dažniausiai ekspatriantų persikėlimo procesą sudaro 4 etapai: sprendimo priėmimas, planavimas ir pasiruošimas, persikėlimas ir įsikūrimas priimančioje valstybėje. Nors tais atvejais, kai ekspatriantas yra tarptautinės įmonės darbuotojas, jam ar jai dažniausiai padedama su planavimu ir persikėlimu (dokumentų tvarkymu, gyvenamosios vietos radimu ar finansavimu ir pan.), tačiau neretai ekspatriantams nesuteikiama pakankamai pagalbos prisitaikyti prie naujo gyvenimo ir bendruomenės jau persikėlus.

Ekspatriantai irgi gali būti klasifikuojami į skirtingas grupes. Sonesh ir Denisis išskiria 2 ekspatriantų kategorijas, į kurias juos suskirsto priimančios valstybės vietiniai gyventojai [19]. Ekspatriantai, kurie atvykę bando save kategorizuoti panašioje socialinėje kategorijoje kaip vietiniai ir siekia palaikyti teigiamą įvaizdį, sumažina vietinių neuztikrintumą, todėl dažniausiai būna priskiriami „priklausančių grupei“ kategorijai (angl. *in-group*) (ibid.). Tuo tarpu likusieji gali būti priskiriami kategorijai „už grupės“ (angl. *out-group*) – su šiais ekspatriantais vietiniai linkę bendrauti rečiau, vietiniai gali jų netgi vengti, nes laiko juos kitokiais, o tai indikuoja, kad bendravimas su jais reikalauja daugiau pastangų, nei su panašiais į vietinius „grupės“ ekspatriantais (ibid.).

Ekspatriantai taip pat gali būti skirstomi remiantis didesne kriterijų gausa ir jų pačių veiksmais, o ne vietinių požiūriu. Vienas iš tokių skirstymų yra „pačių inicijuota ekspatriacija“ (angl. *self-initiated expatriation or SIE*) ir „priskirta ekspatriacija“ (angl. *assigned expatriation or AE*) [5]. Migrantais laikomi tik tie asmenys, keičiantys pagrindinę rezidenciją, kurie išvyksta į kitą šalį neturėdami darbo arba turėdami nelegalų darbą, o visais atvejais, kai išvykimas siejamas su darbo santykiais, asmenys laikomi ekspatriantais (ibid.). Jeigu išvykimas inicijuojamas organizacijos, ekspatriantas gali būti arba „priskirtas“, arba „pritrauktas“ (pavyzdžiui, organizacijos pritraukia aukšto lygio vadovus), o jeigu išvykimas inicijuojamas ekspatrianto, jie skirstomi į *intra-SIE* (asmenys, kurie teikia paraišką į jų dabartinės darbovietės laisvas darbo vietas užsienio šalyse) ir *inter-SIE* (asmenys, kurie įsidarbina kitokioje organizacijoje) (ibid.). Aprašyta migrantų skirstymo logika pavaizduota 1 pav.

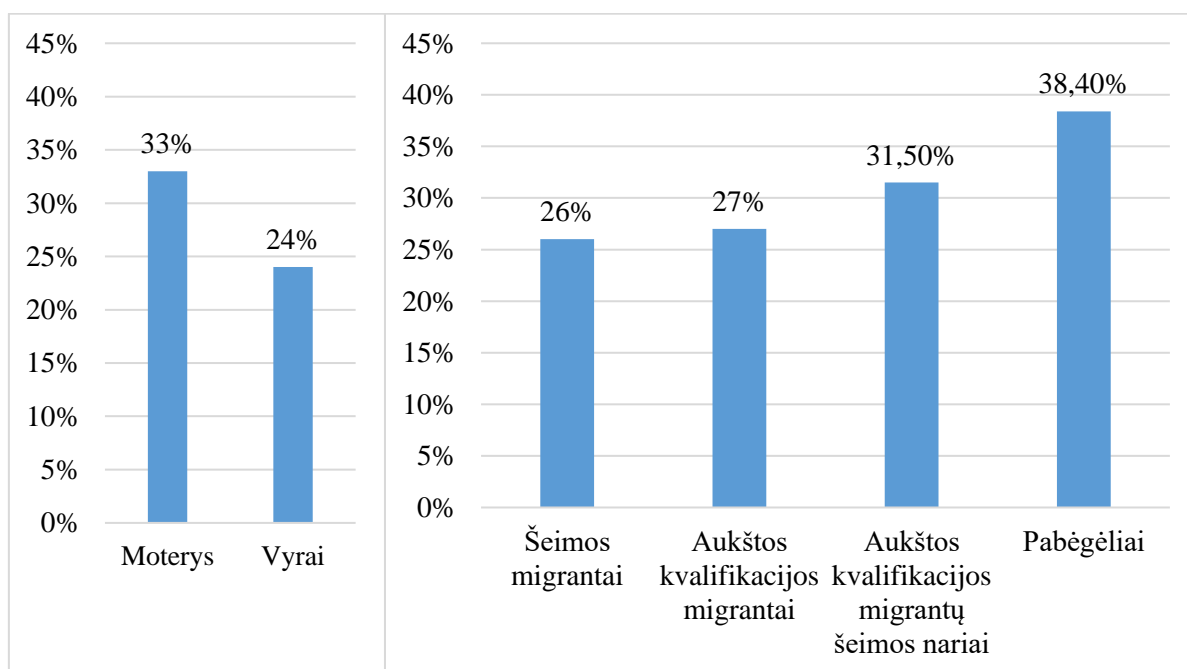


1 pav. Migrantų (ekspatriantų) sprendimų medis (sudaryta remiantis Andresenu, Bergdoltu, Margenfeldu ir Dickannu [5]).

Tokį atskyrimą identifikuoti lengviau nei anksčiau nagrinėtą ekspatriantų skirstymą remiantis vietinių požiūriu, tačiau nagrinėjant komunikaciją socialiniuose tinkluose, gali būti sunku atskirti migrantus nuo ekspatriantų ir įvardyti skirtingus ekspatriantų tipus. Remiantis 1 pav. pavaizduota Andreseno ir kolegų parengta migrantų suskirstymo logika, visi tarptautinių migrantų tipai, aprašyti 1 lentelėje, galėtų būti priskiriami ekspatriantams, jei nagrinėjami migrantai yra legaliai įsidarbinę. Tačiau, schema remiasi sprendimų priėmimo procesu, inicijuotu dar gyvenant gimtojoje šalyje. Didžioji dalis šeimos suvienijimo, aplinkos ir politinių migrantų sprendimus dėl darbo priimtų tik atvykę į priimančiąją valstybę, taigi, tikriausiai būtų priskiriami savo iniciatyva išvykusių ekspatriantų, pakeitusių įmonę, grupę. Toks scenarijus tikėtinas ir vertinant dalį ekonominių migrantų, kurie iš pradžių priima sprendimą pakeisti gyvenamąją vietą, o tik persikėlę pradeda ieškoti darbo.

1.2. Migrantų adaptacijos aktualumas

Migrantams atvykus į naują valstybę jie turi prisitaikyti prie naujos, dažnai nepažįstamos kultūros, susikurti naują paramos tinklą. Šis procesas vadinamas tarpkultūriniu prisitaikymu (angl. *cross-cultural adjustment*) [50]. Paprastai prisitaikymo procesą sudaro 3 dalys: stresas, adaptacija ir vystymasis (angl. *stress-adaptation-growth*) [39]. Ypač daug streso patiria nereguliarūs migrantai, t. y. asmenys, į priimančiąją valstybę atvykę nelegaliai. Tačiau net ir šiame darbe nagrinėjami ekspatriantai ir kiti ekonominiai migrantai, kurie, tikėtina, atvyko į priimančiąją šalį legaliai, gali patirti stresą jaudindamiesi dėl savo teisinio statuso ir teisių [33], be to, visų tipų migrantai gali jausti stresą dėl persikraustymo, akulturacijos, ir, kai kuriais atvejais, išankstinių priimančiosios visuomenės nusistatymų prieš migrantus. Kanadoje atliktas tyrimas atskleidė, kad apie 30 proc. neseniai atvykusių imigrantų susiduria su emocinėmis problemomis kaip liūdesys, depresija, vienišumo jausmas [56]. Be to, tyrimai parodė (žr. 2 pav.), kad neseniai atvykusių imigrantų, susiduriančių su emocinėmis problemomis, dalis buvo didesnė tarp moterų, pabėgėlių bei aukštos kvalifikacijos darbuotojų šeimos narių (ibid.).



2 pav. Dalis imigrantų, susiduriančių su emocinėmis problemomis, pagal lytį ir tipą, 2012 m., proc. (sudaryta remiantis Robertu ir Gilkinsonu [56])

Tyrimas rėmėsi išilginio Kanados imigrantų tyrimo duomenimis, tyrimo imtis siekė 7,7 tūkst. apklaustųjų. Vertinant iš skirtingų regionų atvykusius imigrantus, didžiausia su emocijinėmis problemomis susiduriančių imigrantų dalis buvo užfiksuota tarp atvykėlių iš Afrikos ir Vidurio rytų valstybių (ibid.). Tyrime buvo vertintas ir imigrantų streso lygis. Maždaug 15 proc. imigrantų atsakė, kad didžioji dalis dienų yra labai įtemptos (angl. *very stressful*) arba ypatingai įtemptos (angl. *extremely stressful*) (ibid.).

Taigi, pirmajame adaptacijos etape, kai migrantas patiria daugiausiai streso, susiduria su kultūriniu šoku, ir ilgalaikiams migrantams ir ekspatriantams labai svarbu gauti paramą, kad jų pereinamasis laikotarpis būtų kuo lengvesnis. Migrantams labai svarbūs nauji sukuriami ryšiai, kurie padeda suprasti priimančiosios šalies kultūros normas ir jų laikytis, aktyviau įsitraukti į priimančiąją bendruomenę, pradėti akulturacijos procesus, kurie leistų ilguoju laikotarpiu asimiliuotis [39]. Todėl tolimesniuose skyreliuose bus apžvelgiami galimi paramos tipai, jų teikėjai, ir paramos gavimo įtaka migrantams.

1.3. Parama migrantams

Persikėlimas į naują valstybę sukelia daug streso [4], migrantams svarbu sugebėti prisitaikyti prie kultūrinių ir socialinių pokyčių, todėl labai svarbi bet kokia parama, kuri gali padėti prisitaikyti lengviau. Šiame poskyryje apžvelgiami skirtingi paramos tipai, paramos teikėjai ir ankstesnės literatūroje pateiktos išvados dėl skirtingų paramos tipų bei teikėjų reikšmės migrantų gerovei.

1.3.1. Paramos tipai

Įvairūs eksperimentai ir tyrimai su gyvūnais ir žmonėmis atskleidė, kad socialinė parama gali sumažinti sergamumą ir mirtingumą [31]. Be to, tarpkultūrinis prisitaikymas daro didelę įtaką sėkmei darbe [28]. P. Caligiuri ir M. Lazarova, nagrinėdamos kokius veiksniai padeda moterims migrantėms lengviau kultūriškai prisitaikyti priimančiojoje valstybėje, išskyrė tris socialinės paramos tipus, kurie gali būti aktualūs visiems migrantams: *emocinė parama, informacinė parama ir instrumentinė parama* [11].

Emocinė parama

Pirmasis socialinės paramos tipas yra emocinė parama. Bendravimas su šeima, kolegomis ar kitais asmenimis gali pagerinti migrantų savijautą. Remiantis atliktais tyrimais, neseniai atvykę migrantai jaučiasi vienišesni nei vietiniai gyventojai [66, 40]. Tačiau jei migrantams pavyksta susirasti draugų, patenkinti žmogišką bendravimo poreikį, vienišumo jausmas gali sumažėti. Nagrinėjant moterų ekspatriančių patirtis buvo nustatyta, kad bendravimas padidina moterų savivertę ir psichologinį saugumą, suteikia joms priklausymo priimančiai visuomenei ar naujai aplinkai jausmą [11]. Emocinės paramos pavyzdys galėtų būti kitas migrantas ar ekspatriantas, kuris išklauso, pasidalina savo problemomis ir patirtimi. Tai galėtų padėti sumažinti vienišumo jausmą, o išsakant problemas, diskusijos metu, galėtų kilti ir idėjos problemų sprendimui, kas irgi palengvintų adaptacijos periodą. Tačiau problemų sprendimas aktualesnis antrajame socialinės paramos tipe, kuris apžvelgiamas tolimesniame paragrafe.

Informacinė parama

Antrasis socialinės paramos tipas yra informacinė parama. M. Caligiuri ir M. Lazarova apibrėžia informacinę paramą, kaip bet kokią informaciją, kuri sumažina ekspatrianto neužtikrintumą ir sumišimą, pavyzdžiui, patarimai kur naujoje gyvenvietėje geriausia apsipirkti ar pavakarieniauti arba kaip interpretuoti vietinių gyventojų veiksmus ar kūno kalbą [11]. Tai palengvina migrantų kasdienybę bei gali padėti geriau sutarti su vietiniais (pavyzdžiui, migrantams atvykus į Japoniją svarbu žinoti, kad Japonijoje gyventojai pasisveikindami vienas prieš kitą nusilenkia, o ne spaudžia ranką, kaip populiariu Vakarų valstybėse, kad išvengtų nemalonių situacijų ir nepasirodytų lyg negerbiantys priimančiosios šalies kultūros). Remiantis neapibrėžtumo mažinimo teorija [8], žmogaus prigimtis lemia norą žinoti kuo daugiau apie kitą asmenį dar prieš pradėdant bendrauti, tam, kad sumažintų savo neužtikrintumą, ir galėtų geriau nuspėti kito žmogaus veiksmus. Prie to irgi prisideda informacinė parama – jei naujas kolega migrantui praneša, kad vadovas būna suirzęs rytais, kol neišgeria puodelio kavos, migrantas gali į tai atsižvelgti ir planuoti susitikimus su vadovu praėjus kelioms valandoms po darbo pradžios, taip išvengiant nereikalingos neigiamos reakcijos, nulemtos ne darbo kokybės, bet kitų aplinkybių.

Instrumentinė parama

Paskutinis socialinės paramos tipas yra instrumentinė parama. Ši paramos forma yra materialiausia iš vertinamų formų, nes ekspatriantams ne tik suteikiamos žinios, bet reikalingi resursai. M. Caligiuri ir M. Lazarova pateikia įvairius instrumentinės paramos pavyzdžius, kaip pinigai, auklės paslaugos ar nemokamos priimančios šalies kalbos pamokos [11]. Neretai instrumentinė parama, gaunama iš artimo aplinkos rato, gali būti tuo pačiu metu suvokiama ir priimama kaip emocinė parama [59], be to, instrumentinė parama gali prisidėti prie kitokios paramos išsivystymo [11].

Tyrimuose dažniausiai išskiriama emocinės pagalbos svarba ir įtaka asmenų gerovei [48, 49] bei emocinės pagalbos trūkumas emigrantų tarpe [27] arba socialinė parama vertinama bendrai, nevertinant emocinės, instrumentinės ir informacinės paramos atskirai [43]. Svarbu paminėti, jog įtaką migrantų savijautai daro ne tik jiems suteikiama parama ir jos tipas, bet ir suteikiamos paramos kiekis, jos pakankumas migrantui bei migranto paramos suvokimas [22].

1.3.2. Paramos teikėjai

Kaip jau galima suprasti apžvelgus skirtingus paramos tipus, paramą gali suteikti įvairūs asmenys. Tai gali būti *migranto šeimos nariai, bendradarbiai, kaimynai, nauji vietiniai draugai, kiti migrantai ar specialių su migrantais dirbančių organizacijų atstovai*. Toliau skirtingi paramos teikėjai apžvelgiami detaliau.

Šeima

Migranto šeimos nariai, ir kartu persikėlę į naują šalį, ir likę gimtojoje valstybėje, dažniausiai suteikia migrantams emocinę paramą, kuri gali padėti sumažinti stresą ir įtampą. Šeimos nariai, likę gimtojoje valstybėje, dažniausiai negali suteikti informacinės paramos, nes turės mažiau žinių apie priimančiąją šalį nei pats migrantas, arba suteikti tokios instrumentinės paramos, kaip asmens palydėjimas pas daktarą ar vaikų priežiūra, tačiau jų suteikiama emocinė parama vis tiek gali reikšmingai prisidėti prie geresnės migranto savijautos. Be to, šeimos nariai gali padėti ekspatriantams ir migrantams pasitikėti savo jėgomis ir gebėjimu prisitaikyti prie naujos kultūros [11].

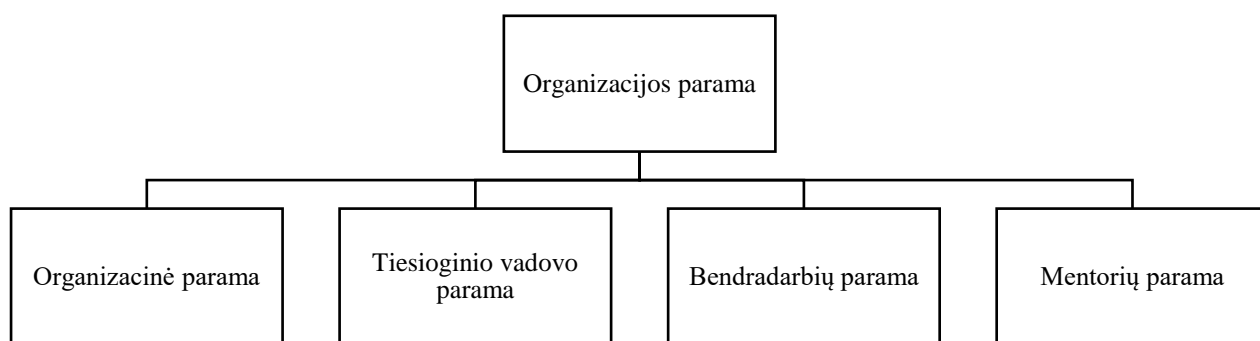
Darbovietė

Organizacijos parama gali būti įvairių lygių. Wee Chan ir Pervaizas išskyrė du organizacijos paramos lygius [71]:

- asmeninį lygį (angl. *person-level*);
- organizacijos lygį (angl. *firm-level*).

Asmeniniame lygyje socialinę paramą darbuotojams teikia tiesioginiai vadovai ir bendradarbiai (pavyzdžiui, grįžtamasis ryšys), o organizacijos lygyje paramą teikia organizacijoje įdiegtos ir vystomos politikos bei programos, padedančios darbuotojams atitinkamose srityse (pavyzdžiui, darbo ir laisvalaikio balansas). Taigi, pasak Simosi [61] bei Wee Chan ir Pervaizo [71], galima išskirti šiuos paramos tipus, pagal paramos teikėjus (žr. 3 pav.):

- organizacinė parama (angl. *organizational support*);
- vadovo parama (angl. *superior support*);
- bendradarbių parama (angl. *peer support*);
- mentorių parama (angl. *mentor support*)



3 pav. Organizacijos paramos tipai pagal paramos teikėjus (sudaryta remiantis Simosi [61], Hu ir kt. [32] bei Wee Chan ir Pervaizu [71]).

Organizacinė parama. Didžioji dalis didelių darbuočių atėjus visiems naujiems darbuotojams stengiasi sukurti pereinamąjį laikotarpį, supažindinti su įmonės tvarka, kolegomis, priklausomai nuo darbo specifikos kartais rengiami ir įvadiniai mokymai. Atvejais, kai darbuotojas atvyksta iš užsienio, organizacija gali padėti prisitaikyti ne tik naujoje įmonėje ar kitame jos padalinyje (tarptautinių įmonių atveju), bet ir priimančiojoje valstybėje. Įmonės pagalbos egzistavimas gali prisidėti prie asmens sprendimo pakeisti gyvenamąją vietą ir sukurti tvirtesnį pagrindą persikraustymui [57]. Įmonės sukurtos programos dažniausiai skirtos suteikti informacinę ir instrumentinę pagalbą: suteikti informacijos apie priimančiąją valstybę, padėti gauti visus reikalingus leidimus, padėti surasti arba tiesiog suteikti apgyvendinimą, suteikti lėšų persikėlimui, padėti užtikrinti, kad migrantui bus prieinamos vaikų priežiūros ar švietimo paslaugos, praveisti arba finansuoti kalbos kursus ir kt. (ibid.).

Tiesioginio vadovo parama. Tyrimai rodo, kad tiesioginio vadovo parama prisideda prie darbuotojo pasitenkinimo darbu ir didesnio įsitraukimo į darbą [1], o tiesioginio vadovo paramos trūkumas gali prisidėti ne tik prie prastėjančių rezultatų darbe, bet ir sveikatos problemų išsivystymo [26]. Vadovai dažniausiai gali suteikti darbuotojams informacinę arba emocinę paramą, pavyzdžiui,

supažindindami darbuotojus su informacija apie darbovietę ir klientus, suteikdami grįžtamąjį ryšį ir patarimus tobulėjimui ar tiesiog išklaudydami ar pagirdami [53].

Bendradarbių parama. Migranto kolegos, ypač vietiniai, gali padėti pagerinti pritapimo jausmą, be to, gali suteikti visų trijų tipų pagalbą: išklausę migrantą gali suteikti emocinę pagalbą, supažindinę su įmonės tvarka, papasakoję apie kolegas ar kur geriausia pietauti prie darbovietės gali suteikti informacinę pagalbą, o padėdami migrantui su darbo atsakomybėmis gali suteikti ir instrumentinę pagalbą. Žemas socialinės paramos iš bendradarbių lygis tyrimuose siejamas su didesniu streso lygiu [21] bei mažesne darbuotojų autonomija [62].

Mentorių parama. Dalyje įmonių taip pat veikia mentorystės programos, pagal kurias ilgiau įmonėje dirbantys kolegos padeda naujiems kolegoms prisitaikyti, ir visu darbo laikotarpiu tampa asmeniu, į kurį naujas darbuotojas gali kreiptis iškilus klausimams. Naujo darbuotojo, „mentorius globotinio“ jaučiama ir suvokiama organizacinė parama stipriai priklauso nuo mentorius suvokimo apie organizacinę paramą, mentorius asmenybės bruožų ir mentoriumi organizacijoje priskiriamų funkcijų [32]. Atvejais, kai šie kintamieji yra teigiami, mentorius turėjimas padidina šansus, kad naujas migrantas darbuotojas sėkmingai įsilies į bendruomenę, pasieks gerus rezultatus bei išdirbs įmonėje ilgesnį laiką [11].

Priimančiosios valstybės vietiniai gyventojai

Kaip ir vietiniai migranto ar ekspatrianto kolegos, kiti vietiniai, su kuriais migrantas užmezga ryšius, gali suteikti įvairaus tipo socialinę paramą ir palengvinti tarpkultūrinio prisitaikymo procesą [11]. Bendravimas su vietiniais gali prisidėti prie abipusio ryšio stiprinimo (t. y. pagerinti ir migranto požiūrį į vietinius, ir vietinių gyventojų požiūrį į migrantą) bei prie migranto gaunamos įvairių tipų socialinės paramos gavimo (ibid.). Dalis interakcijų su vietiniais paprastai būna funkcinės, pavyzdžiui, kai migrantas iškviečia vietinį santechniką, lankosi pas vietinę kosmetologę ar paprašo papildomos prekės parduotuvėje. Tokie kontaktai taip pat sudaro galimybę kai kuriais atvejais gauti emocinę, informacinę ar net instrumentinę paramą, nors ir nesukuria ilgalaikių santykių.

Kiti migrantai

Kitų ekspatriantų arba nedirbančių migrantų socialinė pagalba gali būti ypatingai svarbi ir naudinga, nes jų patirtis neretai atitinka pagalbos ieškančio migranto ar ekspatrianto patirtį, tad jie gali suteikti labai aktualios informacijos ir patarimų [2]. Be to, jeigu migrantas, su kuriuo bendraujama, yra tarpkultūriškai prisitaikęs prie naujos vietovės, sėkmingai dirba ir turi socialinę gyvenimą priimančiojoje šalyje, jis gali tapti įkvėpimo šaltiniu ir siekiamybe naujiems migrantams, kurie jaučiasi geriau žinodami, kad tokiems kaip jie gali pavykti įsitvirtinti naujoje aplinkoje. Atvejais, kai kiti migrantai yra atvykę iš tos pačios šalies, jie gali lengviau pateikti informaciją apie priimančiąją šalį, suteikdami palyginimų tarp abiejų valstybių ir kultūrų. Taigi, bendradarbiaudami migrantai gali padėti vieni kitiems ir suteikti visų trijų tipų socialinę pagalbą [11].

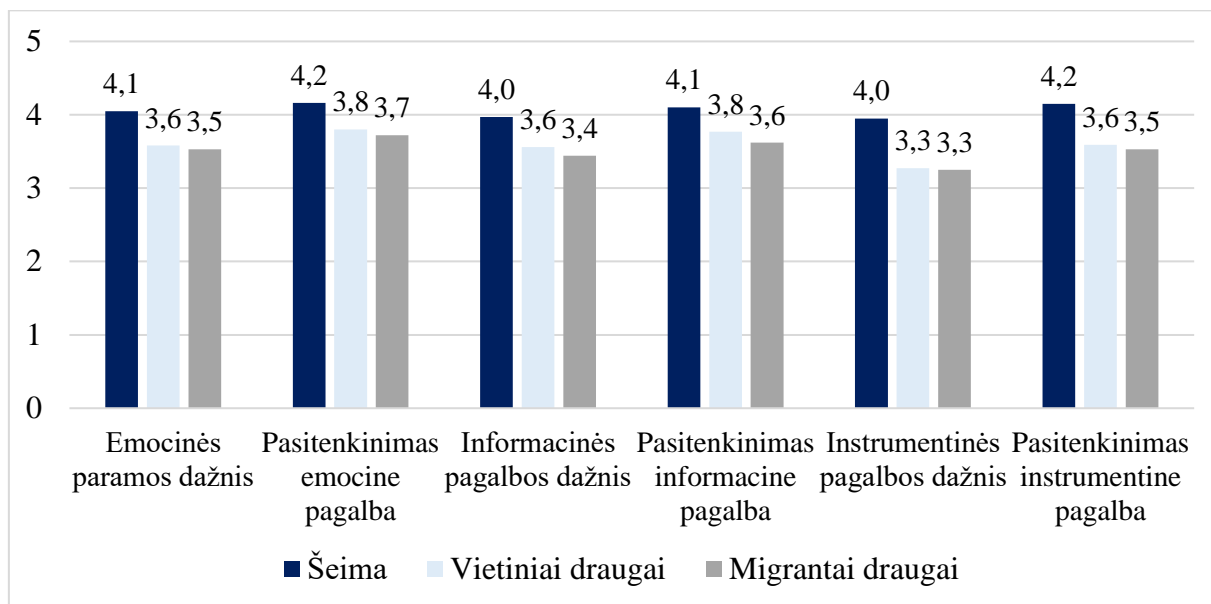
Migrantų paramos organizacijos

Organizacijos, kurios suteikia pagalbą migrantams, dažniausiai prisideda prie instrumentinės ir informacinės paramos suteikimo. Neretai jos padeda migrantams surasti darbą, gyvenamąją vietą, siūlo priimančiosios šalies kalbos kursus, padeda migrantams susitvarkyti reikiamus dokumentus ir pan. [28]. Nors migrantams skirtų organizacijų veikla pastaraisiais dešimtmečiais sparčiai plečiasi

priimančioiose valstybėse, tyrimai rodo, kad šių organizacijų resursai yra menkai naudojami, o pirminis migrantų pagalbos šaltinis yra jų pačių socialinis tinklas (ibid.). Taigi, migrantų organizacijų teikiama socialinė parama nebus nagrinėjama šiame tyrime.

Paramos teikėjų palyginimas

2019 m. publikuotame tyrime buvo vertinami skirtingi paramos tipai, suteikiami skirtingų paramos teikėjų, jų dažnis ir migrantų pasitenkinimas gauta pagalba [30]. Gaunamos pagalbos dažnis buvo vertinamas skalėje nuo 1 iki 5, kur 1 reiškė retai, o 5 – visada, o pasitenkinimas gauta pagalba buvo vertinamas skalėje nuo 1 iki 5, kur 1 reiškė nepatenkintas, o 5 reiškė labai patenkintas.



4 pav. Imigrantų gaunamos socialinės pagalbos dažnis ir pasitenkinimas gauta pagalba, pagal tipus ir paramos teikėjus, 2019 m. (sudaryta remiantis Hombrados-Mendieta ir kt. [30])

Geriausiai imigrantai įvertino savo šeimos narių suteikiamą pagalbą, dažnio ir pasitenkinimo vertinimuose ji visais atvejais aplenkė vietinių ar imigrantų draugų suteiktą socialinę pagalbą. Remiantis tyrimo rezultatais, imigrantams dažniausiai suteikiama emocinė pagalba (vidutiniškai įvertinta 3,7 balais iš 5), pasitenkinimas ja taip pat įvertintas aukščiausiais balais (vidutiniškai 3,9 balai iš 5).

Kadangi šiame darbe bus nagrinėjamos socialinių tinklų migrantų ir ekspatriantų grupės, tikėtina, kad didžioji dauguma nagrinėjamų bendravimo atvejų bus su nepažįstamais žmonėmis internete, galimai su kitais migrantais ir ekspatriantais įvairiuose forumuose. Atsižvelgus į tai, galima įvardinti atskirą pagalbos teikėjų tipą: socialinių tinklų nariai. Nors socialinių tinklų nariai gali priklausyti vienam iš jau nagrinėtų paramos teikėjų tipų, priklausomai nuo socialinės platformos, socialinių tinklų nariai gali tiesiogiai nepažinti migranto, kuriam teikia pagalbą. Asmenims prisijungus prie to paties forumo sukuriama latentiniai ryšiai – asmenys yra susieti dalyvavimo vienos platformos veikloje [18]. Vėliau, jei asmenys tiesiogiai bendrauja platformoje (pavyzdžiui, susirašinėja arba pakomentuoja vienas kito įrašus), latentiniai ryšiai yra aktyvuojami ir tampa silpnais socialiniais ryšiais, turinčiais potencialo ateityje sustiprėti (ibid.).

1.4. Augantis socialinių tinklų vaidmuo migrantų adaptacijos ir paramos procese

Interneto ir socialinių tinklų populiarėjimas suteikia galimybes migrantams ir ekspatriantams išplėsti paramos teikėjų ratą: jei anksčiau ekspatriantai dažniausiai kreipdavosi dėl pagalbos darbovietėje ar ieškodavo pažįstamų žmonių priimančiojoje valstybėje, kurie galėtų padėti, tai šiuo metu pasitelkę internetą ekspatriantai gali kreiptis į milijonus žmonių visame pasaulyje. Įvairios interneto platformos, pavyzdžiui, AlloExpat, InterNations ar paprasčiausios Facebook grupės („berlin EXPATS“, „Expats in London“ ar „LIETUVIAI OSLO“) leidžia asmenims, persikrausčiusiems į tą pačią vietovę, bendrauti vieniems su kitais, pasidalinti patarimais, užduoti klausimus, padėti surasti darbą ir pan. Ekspatriantai Nyderlanduose yra linkę prisijungti prie virtualių bendruomenių ne tik Facebook, bet ir LinkedIn platformoje [67], kuri yra labiau orientuota į dirbančius asmenis, neretai verslo atstovus. Žmonės yra linkę ieškoti emocinės pagalbos tarp žmonių, su kuriais dalijasi panašiomis patirtimis bei patiria panašų stresą [16], o ekspatriantų ir migrantų grupės socialiniuose tinkluose sukuria tinkamą aplinką, padedančią surasti panašius asmenis.

Tyrimai rodo, kad internetiniai resursai padeda asmenims gauti socialinę paramą: nustatytas teigiamas tiesioginis ryšys tarp naudojimosi Facebook ir socialinės paramos priėmimo [38], tinklaraščiai įvertinti kaip sukuriantys naujus paramos išteklius lengvesniam ekspatriantų prisitaikymui [50], įvertinta, kad kompiuterinė socialinė parama kai kuriais atvejais gali atstoti gyvą socialinę paramą ir yra ypač svarbi pradinėse migracijos stadijose [14], be to, nustatyta, kad socialiniai tinklai leidžia lengviau užmegzti naujus ryšius priimančiojoje šalyje ir gauti socialinę pagalbą iš vietinių [29]. Ekspatriantų grupės socialiniuose tinkluose neretai naudojamos ne tik informacijai iš daugiau patyrusių ekspatriantų gauti, bet ir diskusijoms verslo temomis ar net migranto produktų ar paslaugų reklamai [3], o tai gali tapti netiesiogine instrumentine parama reklamą skleidžiančiam ekspatriantui. Tačiau socialinių tinklų naudojimas gali turėti ir neigiamų pasekmių, jeigu migrantų tinklai skatina tik bendravimą etninėse grupėse ir leidžia migrantams lengviau išvengti akulturacijos poreikio bei integracijos į priimančios šalies visuomenę [47].

2020 m. publikuotas ekspatriantams suteikiamos per virtualias platformas socialinės paramos tyrimas (apklausa) parodė, kad dažniausiai internete ekspatriantai gauna pagalbą iš kitų ekspatriantų [12]. Ekspatriantai, kurie ieško informacinės pagalbos internete, dažniausiai ją ir gauna, o emocinė ir instrumentinė pagalba internete yra gaunama rečiau (ibid.). Emocinę bei instrumentinę paramą ekspatriantai dažniau gauna, jei bendrauja ne tik internete, bet ir gyvai susitinka su kitais ekspatriantais, pavyzdžiui, dalyvauja ekspatriantams organizuojamuose tinklų kūrimo (angl. *networking*) renginiuose (ibid.).

Virtualiose platformose socialinė pagalba gali būti suteikiama ne tik atvejais, kai grupės ar forumo nariai atsako vieni į kitų klausimus, bendrauja ar susitinka, bet ir tais atvejais, kai paprasčiausiai paspaudžiamas „like“ ar „upvote“ mygtukas [25]. Tačiau mygtuko paspaudimas skirtinguose socialiniuose tinkluose turėjo skirtingą reikšmę (pavyzdžiui, Facebook'o jaustukai buvo efektyvesni suteikiant socialinę pagalbą nei jaustukai kituose tinkluose), situacija visada priklauso nuo ieškomos socialinės pagalbos tipo, o kai kurie socialiniai tinklai buvo įvertinti kaip netinkami socialinės pagalbos prašymams (ibid.):

- LinkedIn, Snapchat ir Instagram buvo įvardinti kaip netinkamos vietos kreiptis dėl emocinės pagalbos;

- Snapchat ir Instagram buvo įvardinti kaip tinklai, kuriuose informacinės pagalbos ieškojimas būtų nevaisingas;
- visi socialiniai tinklai buvo įvardinti kaip netinkama vieta viešai prašyti instrumentinės pagalbos.

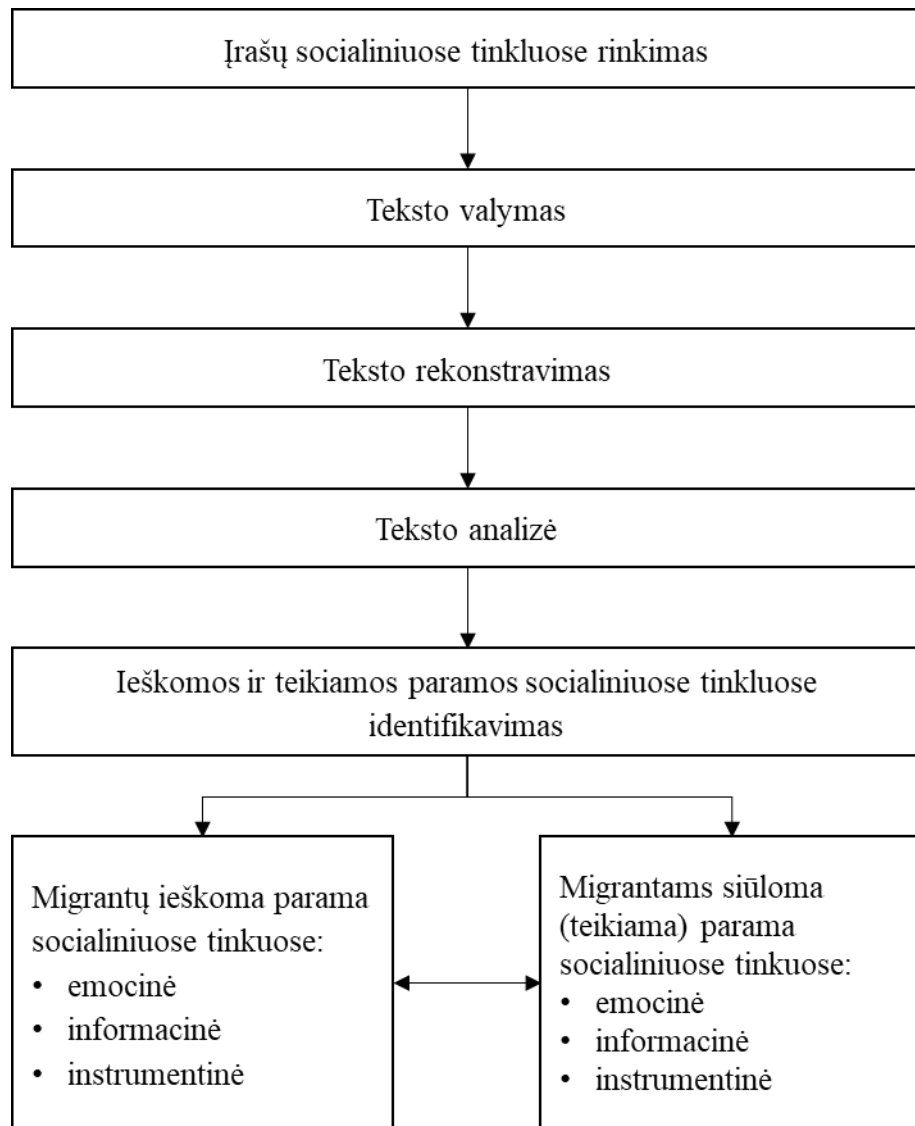
Asmenys, kurie pastoviai ieško emocinės pagalbos socialinėse platformose buvo pavadinti „lėtiniais besiskundžiančiais“ (ibid.) – tai indikuoja, kad emocinės pagalbos paieškos socialiniuose tinkluose gali ne tik teigiamai, bet ir neigiamai paveikti migranto įvaizdį. Be to, jeigu migrantas ieškos pagalbos per dažnai, gaunama nauda galimai palaipsniui mažės.

Įvertinus išanalizuotus tyrimus, galima teigti, kad socialiniai tinklai tapo svarbiu kanalu, per kurį migrantai gali gauti ir teikti kitiems socialinę paramą, tačiau nors gana sėkmingai galima gauti informacinę pagalbą, socialiniuose tinkluose gali būti sunku sulaukti instrumentinės, o kartais ir emocinės socialinės pagalbos.

Atsižvelgus į atliktą literatūros analizę, naudingiausia būtų tirti paramos tipus, kurių ekspatriantai ieško Facebook socialiniame tinkle arba migrantų forumuose. Didelė dalis migrantų komunikacijos Facebook platformoje vyksta privačiose grupėse, kuriose informacijos rinkimą riboja duomenų politika, taigi, tyrime dėmesys skiriamas migrantų forumams specialiose platformose.

1.5. Konceptualus tyrimo modelis

Atsižvelgus į atliktą literatūros analizę išskiriamas pagrindinis darbo tikslas: identifikuoti kokios paramos migrantai ieško ir kokia parama yra teikiama socialiniuose tinkluose. Remiantis darbo tikslu sudarytas konceptualus problemos sprendimo modelis, kuris bus nagrinėjamas šiame tyrime.



5 pav. Tyrimo struktūra

Pasitelkus matematinius algoritmus numatoma išsiaiškinti ar iš tiesų socialiniuose tinkluose migrantai dažniausiai ieško būtent informacinės pagalbos, kaip teigiama apžvelgtoje literatūroje, ir ar migrantų išreiškiamas poreikis yra atliepiamas (t. y. ar jiems suteikiama socialinė pagalba).

Atsižvelgiant į literatūros analizės metu identifikuotus socialinės paramos tipus bei tyrimo struktūrą, iškeliami uždaviniai likusioms tyrimo dalims:

1. Pasirinkti tinkamus matematinius metodus ir algoritmus, skirtus identifikuoti ekspatriantams (migrantams) teikiamą paramą socialiniuose tinkluose, atsižvelgiant į turimus duomenis.
2. Atlikti socialinių tinklų teksto analizę ekspatriantų (migrantų) ieškomos ir teikiamos paramos identifikavimui.
3. Palyginti keletą matematinių algoritmų tarpusavyje ir pasirinkti geriausią remiantis įvairiais kriterijais, kaip tikslumas, jautrumas ir pan.

2. Matematinų algoritmų literatūros analizė

Matematinų algoritmų literatūros analizės dalyje apžvelgiami skirtingi matematiniai algoritmai, kurie galėtų būti naudojami tyrime, atliekamas jų palyginimas, aprašomi kiti svarbūs duomenų paruošimo etapai bei metodai.

2.1. Duomenų rinkimas

Tyrimui reikalingi teksto tipo duomenys yra renkami pasitelkiant tinklo duomenų gavybos (angl. *web scraping*) metodą. *Web scraping* apima įvairias technikas duomenims surinkti – asmenys, renkantys duomenis, gali tai atlikti rankiniu būdu (pavyzdžiui, nukopijuojant tekstą ir įklijuojant į Word ar Excel dokumentus) arba pasitelkdamis specialias programas, kurios automatizuoja procesą [44].

Gauti duomenys visų pirma turėtų būti išvalomi bei transformuojami, o vėliau analizuojami pasitelkiant klasterizavimo / klasifikavimo algoritmus.

2.2. Teksto valymas

Duomenys iš interneto nėra iškart parengti darbui, taigi, visų pirma svarbu pasitelkti įvairias funkcijas teksto valymui. Jomis galima pašalinti skaičius, skyrybos ženklus, nereikalingus tarpus ar neaktualius, dažnai pasikartojančius žodžius, kurie nepadedą išspręsti tyrimo problemos ir tiesiog apsunkina teksto analizės algoritmų veikimą. Tai dažniausiai yra įvairūs jungtukai, išiktukai, jaustukai ar įvardžiai, kurie suteikia nedaug informacijos apie teksto esmę [58]. Literatūroje tokie žodžiai vadinami nereikšmingais žodžiais (angl. *stopwords*). Juos pašalinti analizuojant tekstus anglų kalba yra lengva, nes egzistuoja iš anksto parengti nereikšmingų žodžių sąrašai (juose įtraukiami tokie žodžiai *and, for, thus, to, are, is* ir pan.) bei specialios kodo funkcijos (ibid.). Suskirsčius žodžius pagal šaknis taip pat galima pašalinti retai pasikartojančius žodžius.

2.3. Teksto rekonstravimo metodai

Pirminė teksto analizė (dažniausiai pasikartojančių žodžių apžvalga, pirminių asociacijų identifikavimas) gali būti atliekama naudojant iš netransformuotų žodžių sudarytą terminų ir dokumentų matricą. Tačiau tolimesni žingsniai (klasifikavimas, klasterizavimas), reikalauja duomenų transformacijos. Atsižvelgus į tai, jog analizuojami duomenys yra teksto tipo, siekiant atlikti sėkmingą analizę reikalingi teksto rekonstravimo (angl. *text embedding*) metodai. Teksto arba žodžių rekonstravimo metodai leidžia atskirti žodžius, turinčius panašią reikšmę. Naudojant teksto rekonstravimo metodus, tekstas yra transformuojamas į požymių vektorius. Kiekvienas žodis yra susiejamas su vienu vektoriumi, o vektorių reikšmių mokymasis primena neuroninį tinklą, taigi, teksto rekonstravimo metodas gali būti priskiriamas gilaus mokymosi sričiai [9]. Remiantis antrinių šaltinių analize, populiariausi teksto rekonstravimo metodai yra dvejetainis kodavimas (angl. *binary encoding*), termino dažnio kodavimas (angl. *TF encoding*), termino dažnio – atvirkštinio dokumento dažnio kodavimas (angl. *TF-IDF encoding*), latentinės semantinės analizės kodavimas (angl. *Latent Semantic Analysis Encoding*) bei žodžių į vektorius rekonstravimas (angl. *Word2Vec embedding*). Toliau nagrinėjami 3 iš šių metodų: termino dažnio kodavimas, termino dažnio – atvirkštinio dokumento dažnio kodavimas bei žodžių į vektorius rekonstravimas.

Termino dažnio kodavimas

TF santrumpa anglų kalba gali būti iššifruojama kaip *Term Frequency* – termino dažnis. Termino dažnis atspindi kiek kartų viename dokumente pasikartoja konkretus terminas, kokią dalį dokumento jis užima (termino dažnis apskaičiuojamas padalinant žodžio pasikartojimų skaičių iš bendro žodžių skaičiaus nagrinėjamame dokumente). Tačiau kai kuriais atvejais termino dažnio metodas nepagrįstai padidina dalies mažiau reikšmingų žodžių svarbą. Taigi, dažniau praktikoje taikomas toliau apžvelgiamas termino dažnio – atvirkštinio dokumento dažnio kodavimo metodas.

Termino dažnio – atvirkštinio dokumento dažnio kodavimas

Termino dažnio – atvirkštinio dokumento dažnio kodavimo metodą sudaro jau aptarta termino dažnio (angl. *TF*) dalis bei *IDF* dalis. *IDF* anglų kalba yra termino *inverse document frequency*, kuris reiškia atvirkštinį dokumento dažnį, santrumpa. Atvirkštinis dokumento dažnis apskaičiuojamas dokumentų, kuriuose yra tam tikras terminas, skaičiaus logaritmą padalinant iš bendro dokumentų skaičiaus ir vieneto sumos. Termino dažnio – atvirkštinio dokumento dažnio kodavimas bendrai yra statistinis metodas, leidžiantis įvertinti atskirų žodžių svarbą dokumentuose [69]. Termino dažnio – atvirkštinio dokumento dažnio kodavimo formulė pateikiama 6 pav.

$$\mathbf{tf}(t, d) = \frac{f_d(t)}{\max_{w \in d} f_d(w)}$$

$$\mathbf{idf}(t, D) = \ln \left(\frac{|D|}{|\{d \in D : t \in d\}|} \right)$$

$$\mathbf{tfidf}(t, d, D) = \mathbf{tf}(t, d) \cdot \mathbf{idf}(t, D)$$

$$\mathbf{tfidf}'(t, d, D) = \frac{\mathbf{idf}(t, D)}{|D|} + \mathbf{tfidf}(t, d, D)$$

$f_d(t) :=$ frequency of term t in document d

$D :=$ corpus of documents

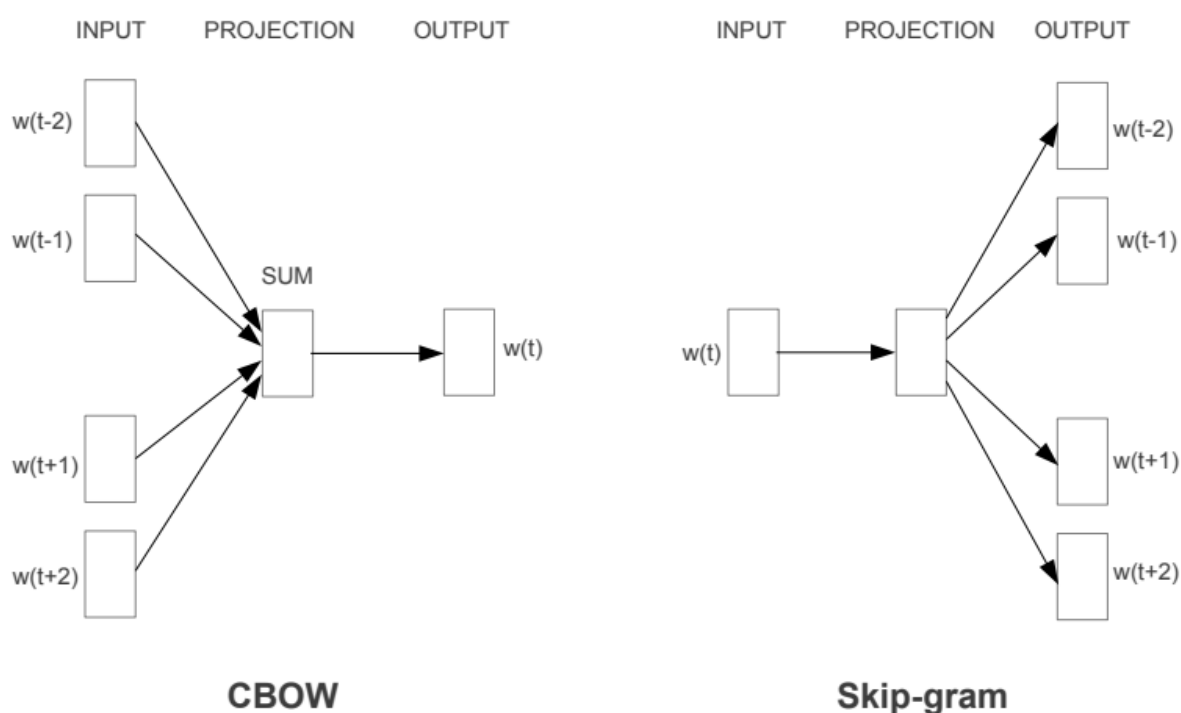
6 pav. Formulės termino dažnio – atvirkštinio dokumento dažnio apskaičiavimui (Vogleris [70])

Atvirkštinio dokumento dažnio dalis taikant termino dažnio – atvirkštinio dokumento dažnio metodą leidžia išspręsti termino dažnio metodo trūkumą, nes atvirkštinis dokumento dažnis sumažina dažniausiai pasikartojančių žodžių svarbą [24]. R programinėje įrangoje galima patiems susikurti specifinį kodą ir funkcijas termino dažnio – atvirkštinio dokumento dažnio apskaičiavimui arba pasinaudoti jau egzistuojančiais paketais, suteikiančiais galimybę apskaičiuoti skirtingų žodžių termino dažnio – atvirkštinio dokumento dažnio įverčius. Kuo įvertis didesnis, tuo aktualesnis žodis yra nagrinėjamame dokumente. Tuo tarpu *Python* programavimo kalboje galima naudoti

„*TfidfVectorizer*“ funkciją iš „*Sklearn*“ nemokamos mašininio mokymosi bibliotekos submodulio „*feature_extraction.text*“.

Žodžių į vektorių rekonstravimas

Žodžių į vektorių rekonstravimo (angl. *Word2Vec*) neuroninis tinklas leidžia paversti žodžius vektoriais. Šis metodas buvo pristatytas T. Mikolovo bei jo kolegų ir gali būti paremtas vienu iš dviejų modelių – Pastovaus žodžių maišo (angl. *Continuous bag-of-words*) bei Šokinėjimo (angl. *Skip-gram*) [46]. Pastovaus žodžių maišo modelyje įvestis yra žodžiai, sudarantys norimo identifikuoti termino kontekstą, o išvestis yra norimo identifikuoti termino prognozė [42]. Tuo tarpu Šokinėjimo modelio įvestis yra konkretus terminas, o išvestis yra žodžių, supančių pirmąjį terminą, prognozės (ibid.). Veidrodinis ryšys tarp dviejų modelių puikiai atsispindi 7 pav. schemeje.



7 pav. Pastovaus žodžių maišo bei Šokinėjimo modelių architektūra (Mikolovas ir kt. [46])

Pasitelkiant vieną iš modelių ir pakartojant 7 pav. pavaizduotus žingsnius visame turimame žodžių rinkinyje, sudaromas žodžių klasifikavimo įrankis. Remiantis pirminėmis Mikolovo ir jo kolegų įžvalgomis, abu modeliai įvertinti kaip geresni ir greitesni nei anksčiau naudoti, tačiau Šokinėjimo modeliui būdingi geresni tikslumo rezultatai (Šokinėjimo modelio tikslumas siekė 65,6 proc., o Pastovaus žodžių maišo – 63,7 proc.) [46]. Tačiau Landthalerio ir jo kolegų eksperimentuose Pastovaus žodžių maišo modelio naudojimas buvo greitesnis (skirtumas reikšmingai didėja augant iteracijų skaičiui bei vektorių dydžiui) ir jį pasitelkus gauti geresnės kokybės rezultatai [42].

2.4. Teksto analizės algoritmai

Žodžius pavertus į vektorių turėtų sekti analizės žingsnis, leisiantis identifikuoti pagrindines temas nagrinėjamuose tekstuose bei suskirstyti duomenis į klases. Toliau apžvelgiami pagrindiniai duomenų klasifikavimui naudojami matematiniai algoritmai, pateikiamas jų palyginimas.

2.4.1. K-artimiausių kaimynų metodas

K-artimiausių kaimynų (angl. *K-nearest neighbors*) algoritmas buvo pristatytas dar 20-ojo amžiaus viduryje [23], vėliau buvo tobulinamas iki šiuo metu naudojamo varianto. Algoritmas gali būti naudojamas detekcijos, klasifikavimo bei regresijos uždaviniams spręsti. K-artimiausių kaimynų algoritmas dažniausiai remiasi standartinio Euklido atstumo tarp dviejų taškų skaičiavimu [36]. Euklido atstumo apskaičiavimo formulė pavaizduota 1 lygtyje pateiktoje žemiau.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

1 lygtis. Euklido atstumo formulė

Euklido atstumas vertina tiesiausią kelią tarp dviejų taškų. Naudojant K-artimiausių kaimynų algoritmą naudojama 1 lygtyje pavaizduota formulė, tačiau teksto analizės atveju vietoj įprastų taškų būtų vertinami požymių vektoriai.

K-artimiausių kaimynų algoritmas taip pat gali naudoti kitokius atstumo apskaičiavimo metodus (pavyzdžiui, Manheteno, Minkovskio, Chi-kvadrato ir kt.), nuo konkretaus metodo pasirinkimo priklauso ir K-artimiausių kaimynų algoritmo rezultatų kokybė [72]. Rezultatus reikšmingai nulemia ir k (artimiausių kaimynų skaičiaus) pasirinkimas. Didesnis skaičius gali padėti sumažinti dispersiją, kurią lemia duomenų triukšmas ir atsitiktinės klaidos, tačiau taip pat gali nulemti mažesnių šablonų praleidimą [74]. Per didelis k skaičius gali nulemti nepakankamo tikslumo (angl. *underfitting*) problemą, o per mažas – persimokymo (angl. *overfitting*).

R programavimo kalboje paketas „*class*“ turi naudojimui parengtą „*knn()*“ funkciją, kurioje naudotojas gali pasirinkti apmokymo bei testavimo duomenų rinkinius, teisingos klasifikacijos faktorių, artimiausių kaimynų skaičių, minimalų balsų skaičių sprendimui priimti bei nurodyti, ar pateikiant rezultatą reikėtų įtraukti jo tikimybę (dalį balsų, kurie nulėmė pasirinktą klasę).

Python programavimo kalboje K-artimiausių kaimynų algoritmo pritaikymui galima naudoti „*Sklearn*“ nemokamos mašininio mokymosi bibliotekos „*sklearn.neighbors*“ modulio funkciją „*KNeighborsClassifier*“. Kaip ir R programavimo kalboje, funkcija remiasi parametrais, kuriuos naudotojas gali keisti. Siekiant panaudoti anksčiau paminėtą Euklido atstumą, metrikos parametro reikšmė turi būti „*minkowski*“, o galios parametras „*p*“ turi būti lygus 2.

Atvejais, kai tikrosios klasės yra žinomos, galima įvertinti ir algoritmo tikslumą, sudaryti sumaišymo matricą, kuri leidžia įvertinti parengtų prognozių jautrumą bei specifiskumą. Plačiau šios charakteristikos aptariamos 2.4.4. skyrelyje.

2.4.2. Naive Bayes metodas

Naive Bayes algoritmas remiasi 18 a. britų statistiko Thomo Bayeso teorema, kuri nustato įvykio tikimybę remiantis tikėtinomis įvykio sąlygomis [7]. Bayeso teorema gali būti pavaizduota ir kaip formulė.

$$P(A | B) = \frac{P(B | A) \cdot P(A)}{P(B)}$$

2 lygtis. Bayeso teoremos formulė

2 lygtyje pavaizduotoje formulėje atsispindi Bayeso teorema – $P(A | B)$ parodo įvykio A tikimybę darant prielaidą, kad galioja tam tikros sąlygos B. Būtent šia formule remiasi Naive Bayes klasifikatorius. Klasifikatorius remiasi prielaida, jog visos skirtingos klasės turi skirtingus bruožus [55]. Nors ši prielaida dažniausiai nėra teisinga praktikoje ir gali lemti netikslias tikimybių prognozes, klasifikavimo uždavinys dažniausiai atliekamas pakankamai tiksliai (ibid.). Naive Bayes rezultatai gali būti pagerinti pasitelkiant papildomus sprendimus kaip vietinis mokymasis, didesnio duomenų kiekio naudojimas, pagalbinių savybių metodai, *Laplace* kalibravimas ir kt. [13, 35, 73].

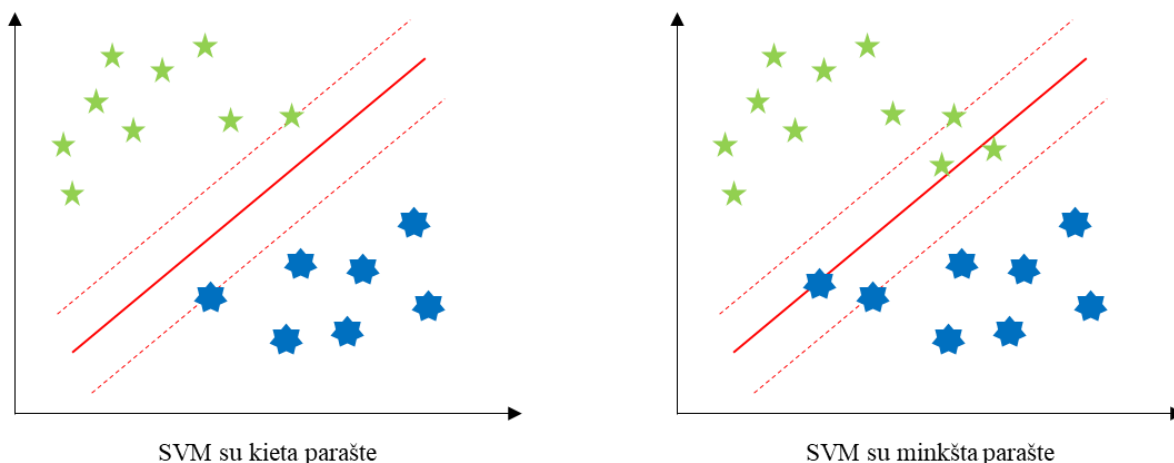
Naive Bayes algoritmas sėkmingai naudojamas įvairiose srityse kaip medicina (klasifikavimas, ligų prognozavimas ir pan.) ar finansai (paskolų rizikos prognozės, sukčiavimo vertinimas ir kt.). Algoritmas taip pat plačiai taikomas teksto klasifikavimo uždavinių sprendimui.

R programavimo kalboje Naive Bayes algoritmą galima surasti „*klAR*“ paketo funkcijose „*NaiveBayes()*“, „*predict.NaiveBayes()*“ jį taip pat iškviečia funkcija „*naiveBayes()*“ pakete „e1701“. Norint pasinaudoti Naive Bayes funkcijomis reikia nurodyti duomenis, apibrėžti formulę, pasirinkti ar norima, jog būtų taikomas *Laplace* išlyginimas ir kt.

Kaip ir K-artimiausių kaimynų algoritmo atveju, *Python* programavimo kalboje Naive Bayes algoritmo funkcija („*MultinomialNB*“) gali būti randama „*Sklearn*“ nemokamoje mašininio mokymosi bibliotekoje, specifiskai jos modulyje „*naive_bayes*“.

2.4.3. Atraminių vektorių mašinos metodas

Atraminių vektorių mašinos metodas (angl. *support vector machine*) yra dar vienas klasifikavimui naudojamas algoritmas. Panaudojęs apmokymo imtį atraminių vektorių mašinos algoritmas suranda tinkamiausią hiperplokštumą skirtingų klasių atskyrimui. Kadangi siekiama, jog hiperplokštuma kuo labiau atskirtų klases, atraminių vektorių mašinos metodą galima vadinti didžiausios paraštės klasifikatoriumi. Atraminių vektorių mašinos algoritmas gali veikti skirtingai, priklausomai nuo nustatymų: gali būti pasirenkama kieta arba minkšta paraštė (atitinkamai į paraštę negali arba gali patekti kažkurie stebėjimai, kai kuriais atvejais atsirandantys arčiau kitos grupės krašto nei savosios), skirtingos funkcijos (tiesinė, kvadratinė, tiesinio ar polinominio branduolio funkcija ir pan.).



8 pav. Atraminių vektorių mašinos metodo (SVM) su kieta ir minkšta paraštėmis palyginimas

Minkštos paraštės pasirinkimas ypač aktualus, kai grupės turi sąlytį tarpusavyje ir nėra galimybės tobulai jas atskirti hiperplokštuma.

Atraminių vektorių mašinos funkcijas galima rasti įvairiuose R programinės kalbos paketuose: pakete „e1071“ yra funkcija „*svm()*“, pakete „*fdm2id*“ yra funkcija „*SVM()*“, pakete „*svmplus*“ yra funkcijos „*SVMP()*“ bei „*svmplus()*“ ir pan. Atraminių vektorių mašinos algoritmo funkcijose naudotojas dažniausiai gali pasirinkti tinkamą funkciją, nurodyti kurie duomenys turi būti naudojami, pasirinkti kokias paskirčiai naudojamas atraminių vektorių mašinos algoritmas (klasifikavimui, regresijai ar išskirčių nustatymui), nurodyti branduolį (angl. *kernel*) ir jo laipsnį, ir, pagal poreikį, nustatyti kitus parametrus.

Python programavimo kalboje atraminių vektorių mašinos funkciją galima rasti „*Sklearn*“ nemokamos mašininio mokymosi bibliotekos „*svm*“ modulyje.

2.4.4. Palyginimas

Kiekvienas algoritmas turi savas silpnybes bei stiprybes, skirtingą veikimo laiką ir specifines charakteristikas.

Mokslinėje literatūroje pateikiamos išvalgos apie K-artimiausių kaimynų, Naive Bayes bei atraminių vektorių mašinos algoritmus, jų stipriąsias ir silpnąsias puses, apibendrinamos 2 lentelė. Algoritmų palyginimas

2 lentelė. Algoritmų palyginimas

	K-artimiausių kaimynų metodo trūkumai	Naive Bayes trūkumai	Atraminių vektorių mašinos metodo trūkumai
K-artimiausių kaimynų metodo privalumai		K-artimiausių kaimynų metodas, kitaip nei Naive Bayes, nereikia apmokymo	K-artimiausių kaimynų metodas gali surasti labiau komplikuočius šablonus [10]

	K-artimiausių kaimynų metodo trūkumai	Naive Bayes trūkumai	Atraminių vektorių mašinos metodo trūkumai
Naive Bayes privalumai	Naive Bayes veikia greičiau nei K-artimiausių kaimynų metodas [34] Naive Bayes yra tikslesnis nei K-artimiausių kaimynų metodas [17, 34]		Naive Bayes yra paprastesnis ir veikia greičiau nei Atraminių vektorių mašinos metodas Kai kuriais atvejais Naive Bayes rezultatai geresni nei Atraminių vektorių mašinos metodo [54], ypač kai yra mažai mokomųjų atvejų [51]
Atraminių vektorių mašinos metodo privalumai	Atraminių vektorių mašinos metodas geriau identifikuoja išskirtis nei K-artimiausių kaimynų metodas Atraminių vektorių mašinos metodas reikalauja mažiau resursų ir yra lengviau interpretuojamas [10]	Atraminių vektorių mašinos metodas nedaro prielaidos, kad stebėjimai yra nesusiję Atraminių vektorių mašinos metodas kai kuriais yra tikslesnis [41], tačiau rezultatai priklauso nuo duomenų charakteristikų	

Remiantis atliktu algoritmų palyginimu, galima teigti, jog tinkamiausias algoritmas priklausys nuo turimų duomenų ir apmokymo galimybių, taigi, kiekvieno tyrimo atveju naudinga išbandyti kelis skirtingus algoritmus.

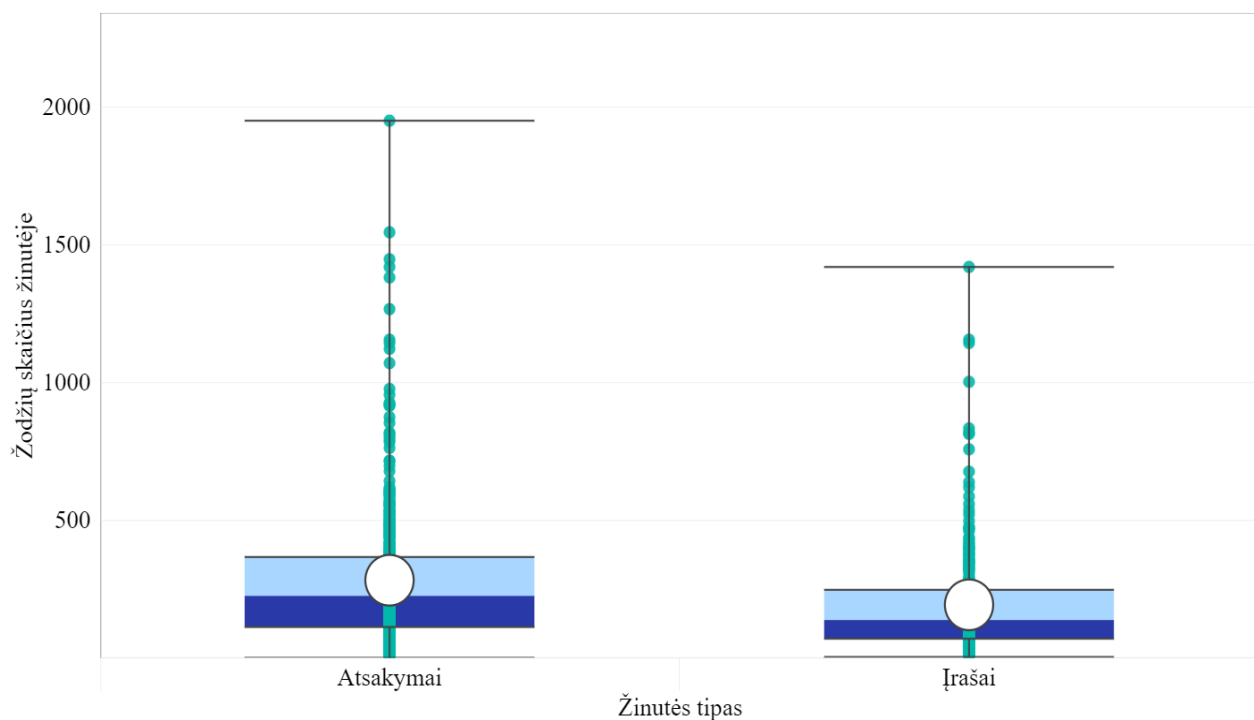
3. Tyrimo metodologija ir duomenų apžvalga

Šiame skyriuje aprašomi pasirinkti ir tyrimo eigoje naudoti metodai, atliekama tyrimo naudojamų duomenų apžvalga ir pirminė analizė.

3.1. Duomenų rinkimas

Visų pirma, siekiant įvertinti kokios paramos migrantai bei ekspatriantai prašo ir kokią gauna socialiniuose tinkluose, buvo pasirinktas platformos *Expatriate.com* forumas. Tinklapyje lankytojai skatinami „užduoti klausimus ir dalintis patirtimi apie gyvenimą užsienyje“ (angl. *ask your questions and share your experience on life abroad*). Šio duomenų šaltinio pasirinkimas leidžia įvertinti paramos teikimą artimai arba visai nepažįstamų migrantų tarpe. Duomenų surinkimui buvo naudojama speciali *web scraping* programa, surenkanti forumo narių įrašus ir atsakymus remiantis įrašų nuorodomis. Iš viso buvo atrinkta daugiau nei 1 tūkst. įrašų ir daugiau nei 6 tūkst. atsakymų. Tačiau buvo atliktas papildomas duomenų valymas, siekiant atrinkti tik tuos įrašus, kuriuose forumo nariai ieškojo pagalbos.

Siekiant suprasti ar įrašuose buvo ieškoma paramos, buvo pasitelkti klausiamieji žodžiai ar frazės, indikuojančios, jog įrašų autoriai kreipėsi į kitus forumo narius siekdami gauti informaciją, pagalbą. Išfiltruoti įrašai, kurie turėjo kažkurį iš šių žodžių ar žodžių junginių: *why which, where, what, searching for, advice, looking for, looking to, want to know, can I, wondering, wonder, does anyone, how, suggestions, advice, help, is there, will I*. Taip pat buvo atrinkti ir įrašai, kuriuose buvo naudojamas klaustukas. Šių kriterijų neatitiko maždaug 5 proc. įrašų, kurie buvo pašalinti. Atlikus duomenų filtravimo žingsnį, galutinis nagrinėjamų įrašų skaičius siekė 941. Tuo tarpu galutinis atsakymų skaičius siekė 6093. Įrašuose bei atsakymuose naudojamų žodžių skaičiaus statistinė analizė pavaizduojama 9 pav.



9 pav. Žodžių skaičiaus žinutėse histograma

Vieną įrašą sudarė nuo 4 iki 1420 žodžių, vidutinis žodžių skaičius įrašė siekė 194. Žodžių skaičiaus įrašuose mediana siekė 139. Tai, kad mediana yra žemesnė, nei vidurkis, parodo, jog duomenų pasiskirstymas yra iškreiptas į dešinę pusę.

Vieną atsakymą sudarė nuo 1 iki 1951 žodžio, vidutinis žodžių skaičius įrašė siekė 283. Žodžių skaičiaus atsakymuose mediana siekė 227. Tai, kad mediana yra žemesnė, nei vidurkis, parodo, jog, kaip ir įrašų atveju, duomenų pasiskirstymas yra iškreiptas į dešinę pusę.

Trečdalis visų vertintų įrašų autorių atsakymuose padėjo kitiems forumo nariams, tad galima daryti prielaidą, jog jie sulaukė paramos, kurios ieškojo.

3.2. Teksto valymas ir paruošimas

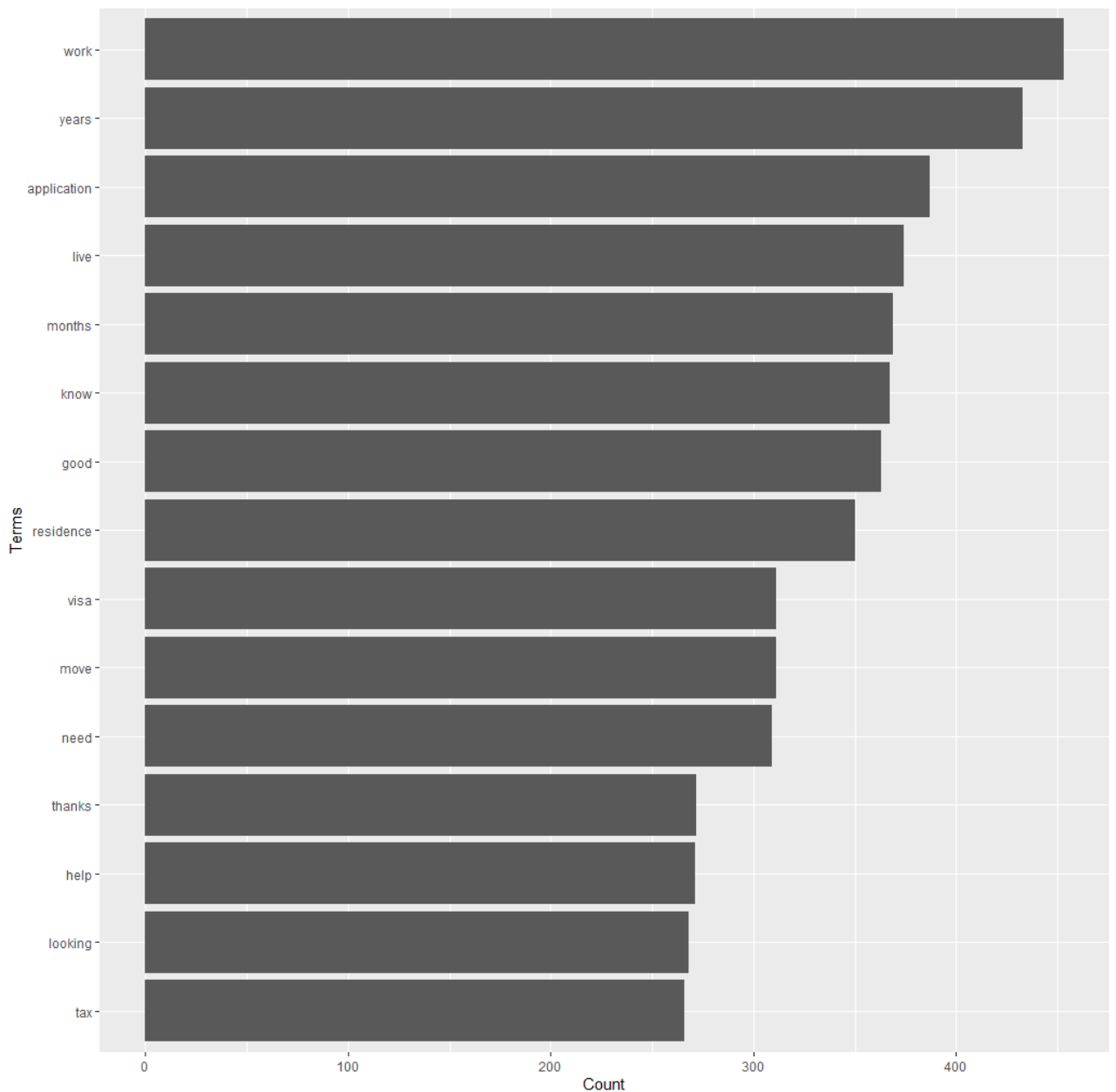
Prieš atliekant pirminę teksto analizę, tekstas buvo valomas R programoje. Kaip numatyta 2.2 poskyryje, iš įrašų ir iš atsakymų buvo pašalinami skyrybos ženklai, nereikalingi tarpai, nereikšmingi žodžiai (įprastas R programoje naudojamas anglų k. nereikšmingų žodžių sąrašas buvo papildytas angliškais terminais *it, everything, will, people, end, as, said, since, now, says, hi, we, would, something, also, already, totally, anyone, can, hello, get, like, make, think, us, match*), ne anglų kalbos raidės ar kitokie simboliai. Taip pat, visos raidės buvo paverstos mažosiomis. Iš žinučių buvo pašalinamos nuorodos, kurių nebūtų galima lengvai klasterizuoti bei interpretuoti. Įrašų bei atsakymų duomenų rinkiniai buvo apdorojami atskirai.

Išvalius duomenų rinkinių teksto įrašus, visų žodžių galūnės buvo pašalinamos, paliekant tik žodžio šaknį. Iš naujo žodžiams galūnės buvo suteikiamos remiantis dažniausiai vartojamu tą šaknį turinčiu žodžiu žinutėse. Programa interpretuotų skirtingus to paties žodžio linksnius kaip skirtingus žodžius, taigi, suvienodinus žodžius pagal šaknį buvo išvengiamas žodžių dubliavimasis, sudaroma galimybė svorį suteikti žodžio reikšmei, o ne konkrečiai gramatinei išraiškai.

Paskutinis žingsnis prieš pradėdant pirminę teksto analizę buvo terminų dokumentų matricos sukūrimas. Terminų dokumentų matricoje eilutės nurodo terminą, o stulpeliai nurodo dokumentą. Įrašų terminų dokumentų matricos dalis pavaizduota 3 lentelėje.

3 lentelė. Įrašų terminų dokumentų matrica

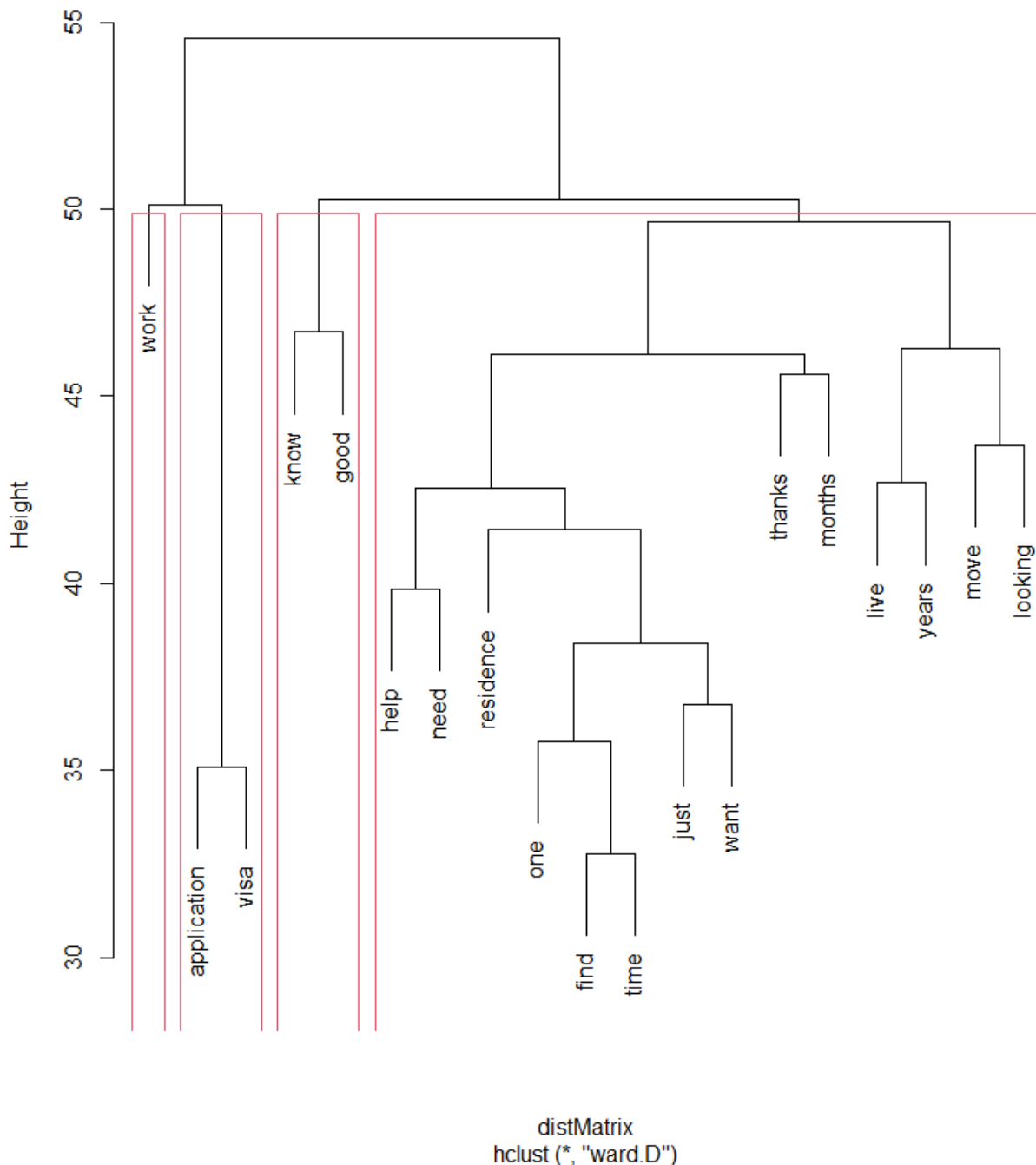
Terminai	Dokumentai									
	103	104	176	322	346	389	412	518	62	663
Paraiška	0	5	0	9	6	0	24	0	1	2
Geras	3	0	1	0	0	0	0	0	6	1
Metai	7	2	5	4	0	1	9	6	1	0
Žinoti	3	0	0	0	0	0	1	0	0	0
Gyventi	1	0	1	3	1	0	13	1	0	3
Mėnesiai	10	1	6	5	1	0	4	0	0	1
Persikraustyti	0	0	0	0	0	0	0	0	0	0
Gyvenamoji vieta	0	2	4	13	0	0	44	0	2	8
Viza	0	1	1	0	0	0	0	0	0	0
Darbas	2	1	1	0	0	1	2	0	0	2



11 pav. Dažniausiai įrašuose naudojamų žodžių horizontali stulpelinė diagrama

Dažniausiai naudojimas terminas buvo darbas (angl. *work*). Jis įrašuose pasikartojo kiek daugiau nei 450 kartų. Daugiau nei 400 kartų buvo minimas ir terminas metai (angl. *years*). Tarp 351 ir 400 kartų buvo vartojami terminai paraiška (angl. *application*), gyventi (angl. *live*), mėnesiai (angl. *months*), žinoti (angl. *know*), geras (angl. *good*).

Siekiant suskirstyti žodžius į klases, kurias sudarytų panašūs terminai, išbandyti septyni skirtingi hierarchinio klasterizavimo metodai: *single*, *complete*, *Ward.D*, *Ward.D2*, *average*, *median*, *centroid*. Hierarchinio klasterizavimo algoritmas leidžia sugrupuoti panašius objektus į grupes (klasterius). Nagrinėjamu atveju, *Ward.D* hierarchinio klasterizavimo metodas leido gražiausiai atskleisti bei atskirti klasterius tarpusavyje, todėl būtent jis buvo pasirinktas. Įrašų klasifikavimo rezultatai pavaizduoti 12 pav.



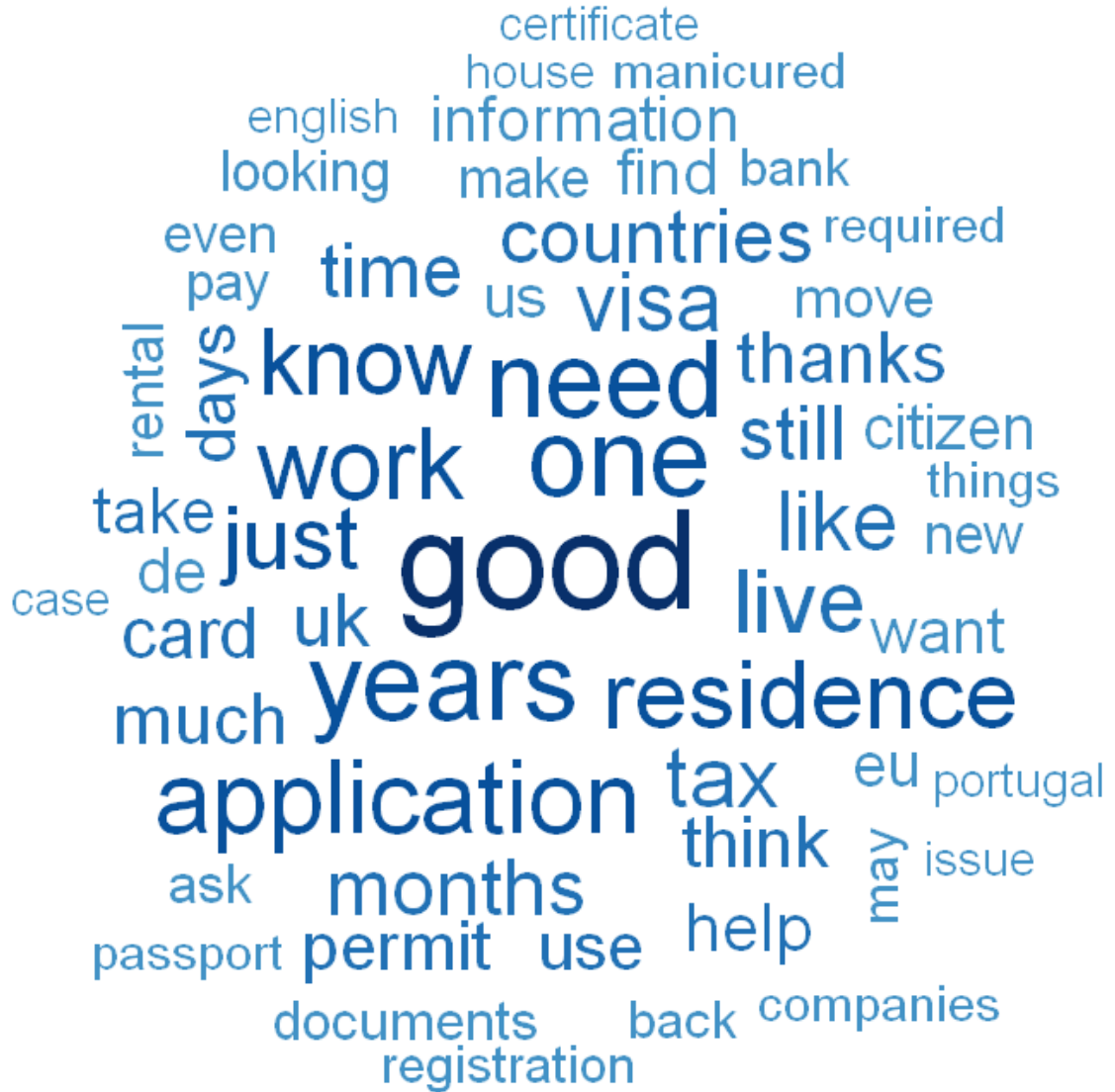
12 pav. Įrašų žodžių klasteriai

Remiantis sudarytais klasteriais, darbo tema atskiriama nuo kitų nagrinėjamų tematikų. Aiškiai išsiskiria ir dokumentų tematika, į kurią patenka paraiškų ir vizų terminai. Trečiajam klasteriui buvo priskirti žodžiai žinoti ir geras. Suvokiant *Expatriation.com* platformos forumo paskirtį ir įprastą įrašų pobūdį, galima daryti subjektyvią prielaidą, jog įrašuose, įtraukiančiuose šiuos žodžius, migrantai bei ekspatriantai klausė kitų forumo narių, ar jie žino gerą produktą, vietą, paslaugą arba kitą jiems rūpimą dalyką (pavyzdžiui, ar žino gerą advokatą, gerą vietą persikėlimui, gerą mokymosi įstaigą vaikams ir pan.).

Žodžių sąsajos, įvertintos naudojant „*findAssocs*“ funkciją, atskleidė, jog su ekspatriacijos terminu dažniausiai buvo siejami terminai neįprastas (angl. *unusual*), kelionės tikslas (angl. *destination*), veiksnys (angl. *factor*), įvairiopas (angl. *diverse*). Remiantis atlikta literatūros analize ir ekspatriantų

forumo kontekstu, galima teigti, jog ekspatriantai uždavė klausimus, susijusius su tinkamos valstybės pasirinkimu (kelionės tikslas), domėjosi veiksniais, kurie lemia sėkmingą persikėlimą ir pan.

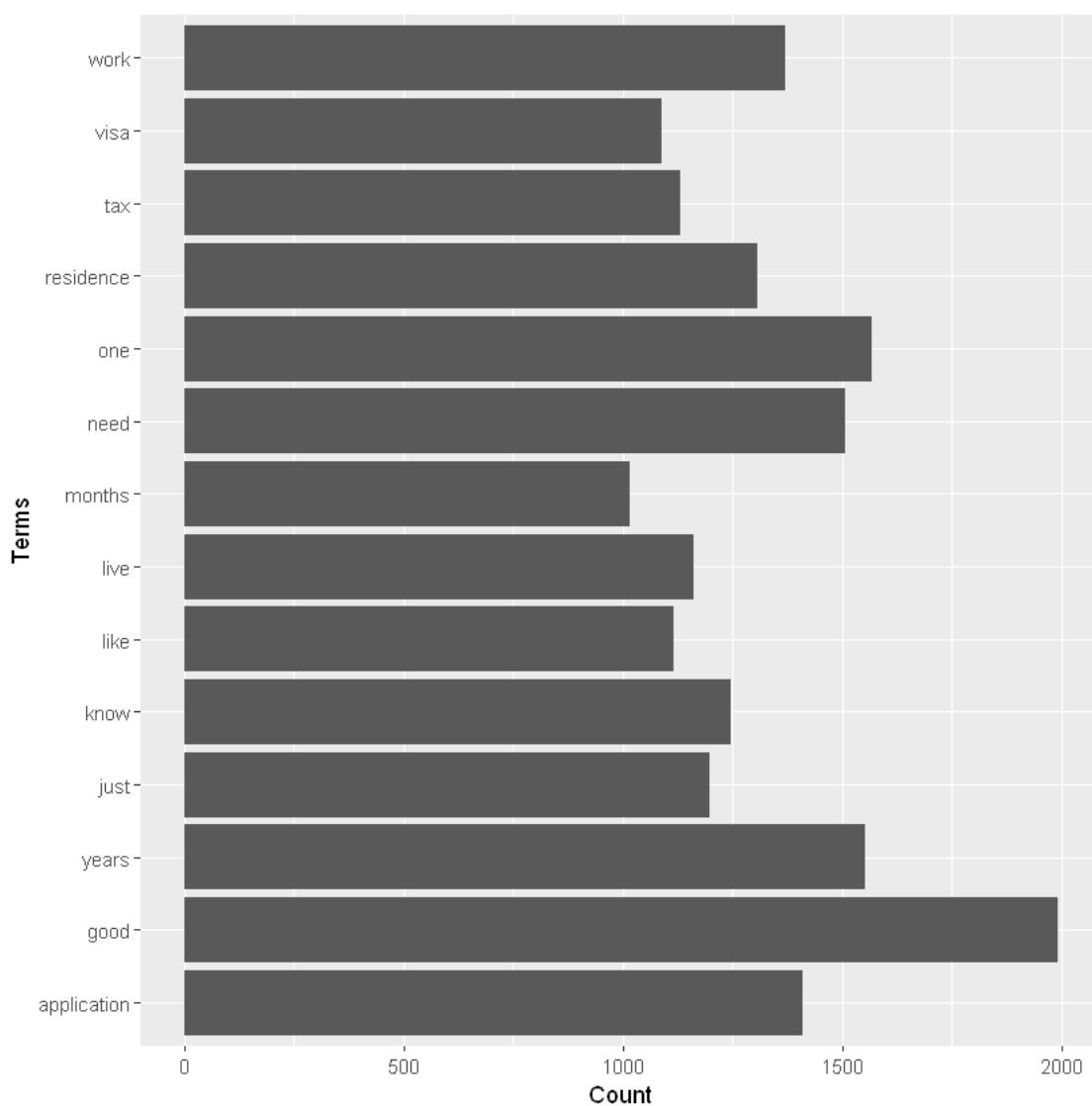
Nagrinėjant atsakymus į jau apžvelgtus ekspatriantų bei migrantų klausimus taip pat buvo sudarytas žodžių debesis. Atsakymų žodžių debesis pateikiamas 13 pav. Dalis žodžių debesyje sutampa su žodžiais, matytais įrašų žodžių debesyje (žr. 10 pav.).



13 pav. Atsakymų žodžių debesis

Remiantis atsakymų žodžių debesių, dažniausiai atsakymuose vyravo darbo, dokumentų, gyvenamosios vietos pasirinkimo temos. Atsižvelgiant į tai, kad įrašuose bei atsakymuose vyraujančios temos sutampa, galima daryti prielaidą, kad įrašuose klausimus uždavusiems asmenims buvo suteikiama jiems aktuali informacija, atitinkanti klausimo temą.

Populiariausi atsakymuose naudoti žodžiai taip pat buvo pavaizduoti diagramoje, siekiant aiškiau pavaizduoti kiekvieno žodžio pasikartojimo dažnį bei suteikti galimybę lengviau palyginti žodžių dažnį tarpusavyje. Diagrama pateikiama toliau esančiame 14 pav.



14 pav. Dažniausiai atsakymuose naudojamų žodžių horizontali stulpelinė diagrama

Beveik 2 tūkst. kartų atsakymuose buvo panaudotas žodis geras (angl. *good*), daugiau nei 1,5 tūkst. kartų buvo vartoti žodžiai vienas (angl. *one*), metai (angl. *years*) bei reikia (angl. *need*). Dešimties dažniausiai įrašuose ir atsakymuose pasinaudojančių žodžių palyginimas pateikiamas 4 lentelėje. Žodžiai, pasikartojantys abiejuose stulpeliuose, pavaizduoti palinkusiu šriftu.

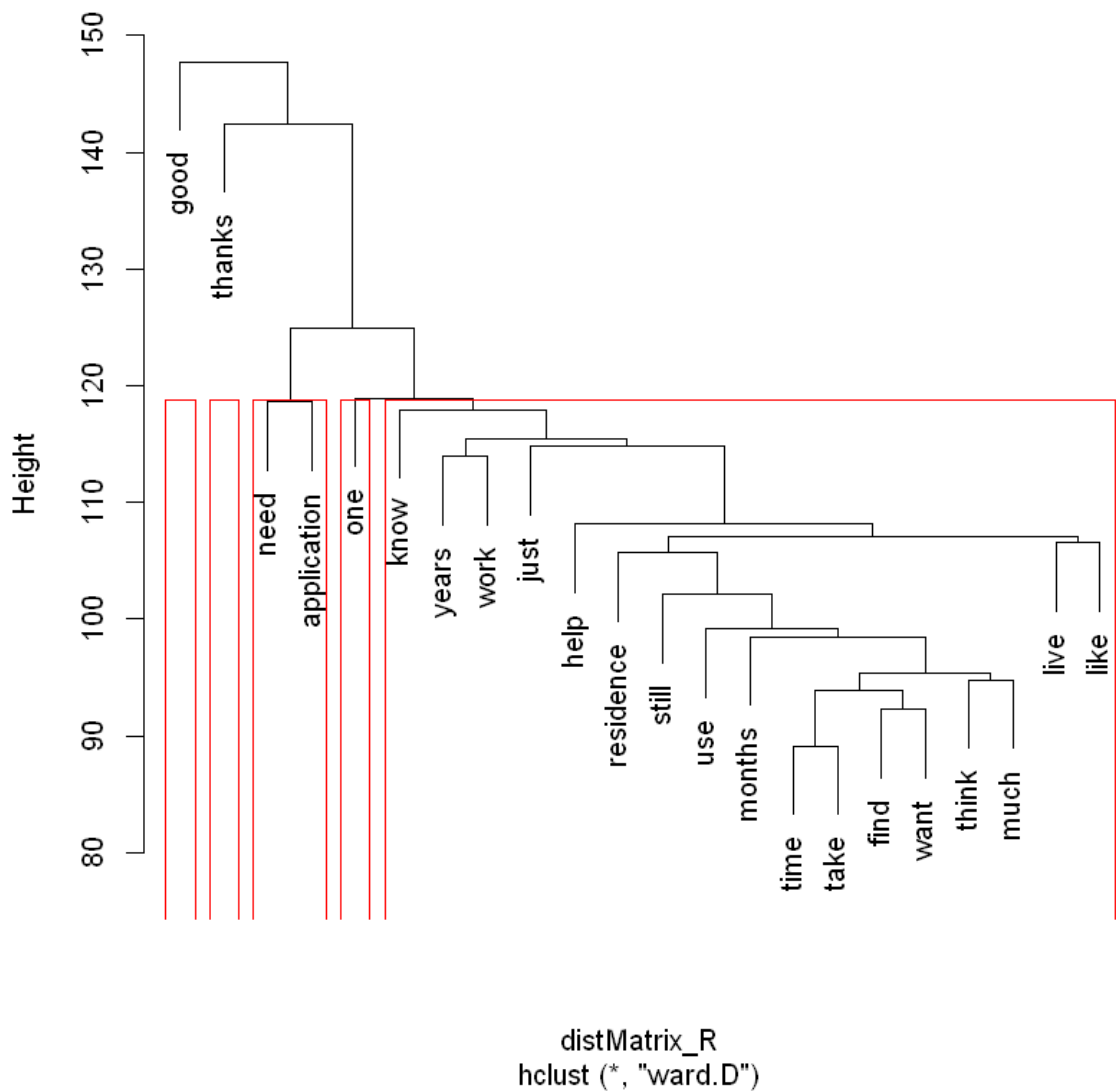
4 lentelė. Dešimties dažniausiai pasikartojančių žodžių palyginimas tarp įrašų ir atsakymų

	Įrašai	Atsakymai
1.	<i>Darbas</i>	<i>Geras</i>
2.	<i>Metai</i>	Vienas
3.	<i>Paraiška</i>	<i>Metai</i>
4.	<i>Gyventi</i>	Reikia
5.	Mėnesiai	<i>Paraiška</i>
6.	<i>Žinoti</i>	<i>Darbas</i>

	Įrašai	Atsakymai
7.	<i>Geras</i>	<i>Gyvenamoji vieta</i>
8.	<i>Gyvenamoji vieta</i>	<i>Žinoti</i>
9.	Viza	Tiesiog
10.	Persikėlimas	<i>Gyventi</i>

Nors vertinant įrašus ir atsakymus nesutapo populiariausių žodžių eiliškumas, net 70 proc. terminų imtyse sutapo.

Rengiant atsakymų žodžių klasterį vėl buvo išbandyti 7 skirtingi metodai. *Ward.D2* bei *Ward.D* metodų rezultatai buvo gana panašūs, taigi, kaip ir įrašų nagrinėjimo atveju, pasirinktas *Ward.D* metodas klasteriams identifikuoti. *Ward.D* metodu į klasterius suskirstyti dažniausiai atsakymuose naudojami žodžiai pavaizduoti 15 pav.



15 pav. Atsakymų žodžių klasteriai

Pavieniai žodžiai gerai (angl. *good*), ačiū (angl. *thanks*) ir vienas (angl. *one*) buvo priskirti atskiriems individualiems klasteriams. Žodžio ačiū išskyrimas gali parodyti, jog atsakymuose vyravo dėkojimo

tematika – galimai padėka buvo išreiškiama ne tik įrašų autorių, bet ir kitų forumo narių. Remiantis faktu, jog buvo nagrinėjami atsakymai į klausimus ir buvo naudojamas žodis gerai (arba geras), galima daryti prielaidą, kad atsakymuose buvo rekomenduojami geri daiktai, vietos, paslaugos ar kt.

3.4. Teksto integravimas

Teksto integravimui buvo pasirinktas termino dažnio – atvirkštinio dokumento dažnio kodavimas modelis, apžvelgtas Teksto rekonstravimo metodai^{2.3} poskyryje. Nors pirminė duomenų analizė buvo atliekama naudojant R programą, teksto integravimo žingsnis buvo atliekamas naudojant *Python* programavimo kalbą *JupyterLab* internetinėje interaktyvioje kūrimo aplinkoje. Šį sprendimą lėmė *Python* kalbos patogumas integruojant tekstą (R programoje dauguma teksto integravimo metodų yra paremti *Python*, taigi, reikia naudoti specialius paketus, leidžiančius sukurti sąsają tarp R ir *Python*) bei didesni *JupyterLab* skaičiavimo pajėgumai.

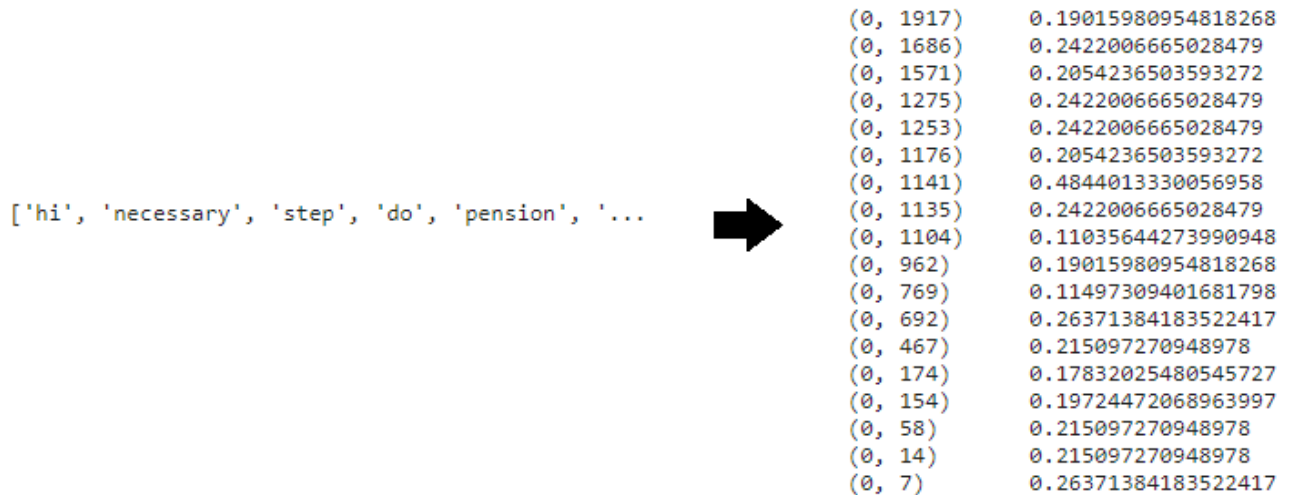
Jei nėra pasirenkami kitokie funkcijos nustatymai, žodynas sukuriamas remiantis funkcijai pateikiamu duomenų masyvu. Termino dažnio – atvirkštinio dažnio kodavimo kontekste žodynas susieja terminus su funkcijų indeksais. Įrašų integravimui sukurto žodyno dalis pavaizduota 5 lentelėje.

5 lentelė. Įrašų žodynas

	Žodis	Indeksas
1.	<i>Hello</i>	761
2.	<i>Name</i>	1122
3.	<i>Pierre</i>	1281
4.	<i>Long</i>	1000
5.	<i>Speak</i>	1649
6.	<i>French</i>	676
7.	<i>English</i>	541
8.	<i>Happy</i>	751
9.	<i>Coffee</i>	287
10.	<i>Laugh</i>	954
11.	<i>Beer</i>	167
12.	<i>Canadian</i>	233
13.	<i>Retire</i>	1482
14.	<i>Nurse</i>	1178
15.	<i>Expat</i>	580
16.	<i>Portugal</i>	1314
17.	<i>Husband</i>	808
18.	<i>Enjoy</i>	542
19.	<i>Love</i>	1009

	Žodis	Indeksas
20.	Climate	284

Pritaikius *TfidfVectorizer* funkciją turimam sutvarkytam duomenų masyvui, terminai buvo paversti ypatybių vektoriais. Tai leido toliau naudoti skaitinę teksto išraišką duomenų klasifikavimui. Vieno iš įrašų pokyčio pavyzdys pateikiamas 16 pav.



16 pav. Teksto integravimo pavyzdys

Teksto integravimo pavyzdys parodo kaip kiekvienas įrašo žodis įgauna skaitinę reikšmę. Terminai ir jų vektoriai dešinėje pusėje nėra išrikiuoti ta pačia tvarka, kaip originaliame sakinyje, o išrikiuoti pagal žodyne jiems priskirtą indeksą, mažėjimo tvarka. Antrojo įrašo žodžio būtina (angl. *necessary*) skaitinės išraiškos pavaizduotos aštuntoje eilutėje. Nulis skliausteliuose parodo, kad nagrinėjama pirmoji duomenų rinkinio eilutė, o 1135 leidžia susieti įrašą su konkrečiu žodžiu panaudojant 5 lentelėje pavaizduotą žodyną. Antrasis stulpelis nurodo termino dažnio – atvirkštinio dokumento dažnio balą. Žodžio pasikartojimas dokumente (šiuo atveju įrašė) kelia balo įvertį, tačiau žodžio pasikartojimas dokumentuose (šiuo atveju kituose įrašuose), jį mažina. Tai užtikrina, kad daugelyje įrašų pasikartojantys žodžiai, nepadedantys atskirti skirtingų įrašų klasių, neturės labai aukšto termino dažnio – atvirkštinio dažnio balo.

3.5. Duomenų klasifikavimas

Duomenų klasifikavimui naudojami 2.4 poskyryje aprašyti algoritmai (K-artimiausių kaimynų, Naive Bayes bei atraminių vektorių klasifikatoriai). Atlikus teksto rekonstravimą, tolimesni veiksmai buvo tęsiami naudojant *Python* programą.

„*MultinomialNB*“ funkcija *Python* programavimo kalboje yra skirta klasifikacijai su diskrečiais kintamaisiais, tačiau praktikoje funkcija pritaikoma ir trupmeniniams skaičiams (pavyzdžiui, termino dažnio – atvirkštinio dažnio balams). Taikant daugianarį (angl. *multinomial*) Naive Bayes algoritmą tikimybė apskaičiuojama remiantis 3 lygtyje pateikiama formule.

$$p(x | C_k) = \frac{(\sum_{i=1}^n x_i)!}{\prod_{i=1}^n x_i!} \prod_{i=1}^n p_{ki}^{x_i}$$

3 lygtis. Tikimybės apskaičiavimas su daugianariu Naive Bayes

3 lygtyje pavaizduota formulė remiasi prielaida, jog naudojami diskretieji skaičiai. Tačiau, dėl pavaizduotos gamos funkcijos tapatybės $n! = \Gamma(n+1)$, daugianario Naive Bayes tikimybės funkcija gali būti perrašoma į leidžiančią naudoti neneigiamus realiuosius skaičius. Iš 3 lygtyje pavaizduotos formulės išvesta neneigiamų realiųjų skaičių naudojimą leidžianti formulė yra pavaizduota toliau pateiktoje 4 lygtyje.

$$p(x | C_k) = \frac{(\sum_{i=1}^n x_i)!}{\prod_{i=1}^n x_i!} \prod_{i=1}^n p_{ki}^{x_i} = \frac{\Gamma(1 + \sum_i x_i)}{\prod_i \Gamma(1 + x_i)} \prod_{i=1}^n p_{ki}^{x_i}$$

4 lygtis. Tikimybės apskaičiavimas su daugianariu Naive Bayes (leidžiantis naudoti neneigiamus realiuosius skaičius)

Apmokant ir vėliau pritaikant Naive Bayes algoritmą prognozėms rengti nebuvo keičiami numatytieji parametrai.

Naudojant atraminių vektorių mašinos algoritmą buvo pasirinkti konkretūs parametrai. Išbandžius kelis skirtingus branduolio parametro variantus buvo nustatyta, jog geriausi rezultatai pasiekiami naudojant linijinį branduolį. Išbandžius skirtingas C parametro, nurodančio sureguliuojimo stiprumą, reikšmes, apmokymo imties prognozių tikslumas vis tiek siekdavo 100 proc. Taigi, siekiant surasti tinkamą balansą tarp algoritmo paskatinimo kuo tiksliau atskirti klases ir modelio lankstumo išlaikymo, buvo paliktas standartinis $C=1,0$ parametras. Pasirinkus linijinį branduolį, nebuvo svarbūs laipsnio ir gama parametrai.

K-artimiausių kaimynų algoritmo panaudojimo metu taip pat buvo pasirenkami konkretūs parametrai. Siekiant, kad algoritmas veiktų remiantis Euklido atstumu, buvo pasirinkta Minkovskio metrika bei dviem lygus galios parametras. Buvo išbandyti keli skirtingi kaimynų skaičiaus parametro pasirinkimai. Apmokymo imties tikslumo rezultatai naudojant skirtingas kaimynų skaičiaus parametro reikšmes pateikiami 6 lentelėje.

6 lentelė. K-artimiausių kaimynų algoritmo apmokymo imties prognozių tikslumas

n_neighbors parametro reikšmė	Apmokymo imties prognozių tikslumas
3	83,56 proc.
5	86,30 proc.
7	82,19 proc.

Atsižvelgus į gautus rezultatus, tolimesni veiksmai buvo atliekami prilyginus parametro reikšmę penkiems.

Algoritmų palyginimas atliekamas remiantis algoritmų tikslumu, jautrumu bei specifiškumu. 7 lentelėje pateikiamas sumaišymo matricos pavyzdys su apibrėžimais, skliausteliuose pateikiami anglų kalboje naudojami sutrumpinimai.

7 lentelė. Sumaišymo matricos pavyzdys (sudaryta remiantis Zhu, Zengu ir Wangu [75])

Prognozės rezultatas	Reali reikšmė	
	Pozityvi	Negatyvi
Pozityvus	Teisingas teigiamas (TP)	Klaidingas teigiamas (FP)
Negatyvus	Klaidingas neigiamas (FN)	Teisingas neigiamas (TN)

Tikslumas parodo kokia dalis iš visų įvertinimų buvo teisingi. Taigi, tikslumas yra apskaičiuojamas sudedant teisingų teigiamų bei teisingų neigiamų prognozių skaičių ir padalijant šią sumą iš bendro prognozių skaičiaus $((TP + TN)/(TP+TN+FP+FN))$.

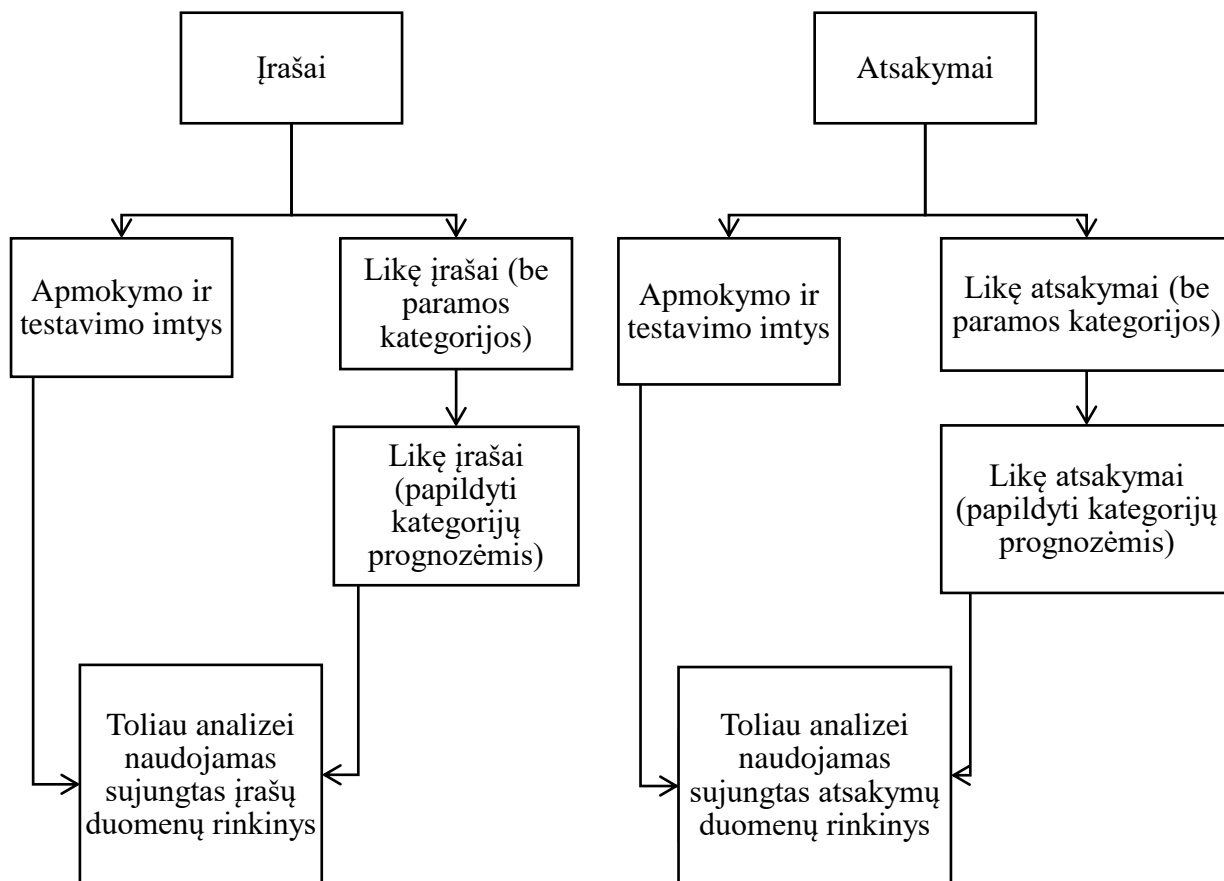
Jautrumas atskleidžia kokią dalį pozityvios klasės reikšmių algoritmas prognozuoja teisingai. Taigi, jautrumas yra apskaičiuojamas padalijant teisingų teigiamų prognozių skaičių iš realių teigiamų reikšmių skaičiaus, kuris gaunamas sudėjus teisingų teigiamų bei klaidingų neigiamų prognozių skaičių $(TP/(TP+FN))$.

Specifiškumas reprezentuoja negatyvios klasės reikšmių dalį, kuri buvo prognozuota teisingai. Siekiant apskaičiuoti specifiškumą reikia teisingų neigiamų reikšmių skaičių pagalinti iš teisingų neigiamų bei klaidingų teigiamų prognozių sumos $(TN/(TN+FP))$.

Rezultatai, gauti pritaikius skyriuje apžvelgtus veiksmus ir poskyryje aprašytus algoritmus, pateikiami 4 skyriuje.

4. Tyrimo rezultatai

Šiame skyriuje apžvelgiami ir palyginami skirtingų algoritmų klasifikavimo rezultatai, analizuojami sudaryti sintetiniai duomenų rinkiniai, aprašomos analizės įžvalgos. Siekiant aiškiai pateikti duomenų rinkinių skaidymo ir sudarymo procesą, 17 pav. pateikiama duomenų rinkinių pokyčių schema.



17 pav. Įrašų bei atsakymų duomenų rinkinių pokyčių schema

Ir įrašų, ir atsakymų atvejais, buvo atrinkta dalis žinučių ir sukurtos apmokymo bei testavimo imtys. Atrinkus tinkamiausią paramos tipo klasifikavimo metodą, vėliau ir likusiai duomenų imčiai buvo priskirtos klasės, bei visi duomenys apjungti siekiant juos panaudoti tolimesnei analizei. Detaliai duomenų klasifikavimo, sintetinių duomenų rinkinių kūrimo bei jų analizavimo procesas aprašomas tolimesniuose poskyriuose.

4.1. Paramos tipų identifikavimas pasitelkiant matematinius algoritmus

Siekiant apmokyti algoritmus, daliai duomenų buvo priskirtos paramos kategorijos. Vertinant įrašus, identifikuoti tik keli atvejai, kai būtų galima prašomą pagalbą įvardinti kaip informacinę-instrumentinę. Taigi, buvo priimtas sprendimas vertinti pasiskirstymą tik tarp emocinės ir informacinės paramos prašymo. Informacinės paramos kategorijai dažniausiai priskirti prašymai supažindinti su vizų ar kitų dokumentų išdavimo tvarka, patarti dėl nekilnojamojo turto pirkimo ar nuomos ir pan. Tuo tarpu emocinės paramos kategorijai priskirti įrašai, kuriuose forumo nariai mini, jog norėtų susitikti su kitais bendraminčiais ar pabendrauti nuotoliniu būdu, pasidalinti savo patirtimis ar susirasti draugų.

4.1.1. Įrašų klasifikavimas pagal prašomos paramos tipą

Modelio apmokymui ir testavimui buvo panaudoti 105 įrašai. 30 proc. įrašų buvo priskirta testavimo imčiai. Iš pradžių 2.4 ir 3.5 poskyriuose aprašomi algoritmai buvo apmokyti naudojant mokomąją duomenų imties dalį. Vėliau jų rezultatai buvo įvertinti naudojant testavimo imtį. Testavimo imties tikslumo rezultatai pateikiami 8 lentelėje.

8 lentelė. Algoritmų tikslumo prognozuojant įrašuose prašomos paramos tipą palyginimas

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas	78,13 proc.	81,25 proc.	84,38 proc.

Remiantis gautais rezultatais, sėkmingiausiai įrašų paramos tipus prognozavo K-artimiausių kaimynų algoritmas. Atraminių vektorių mašinos algoritmo tikslumo rezultatas buvo 3,13 proc. punkto mažesnis nei K-artimiausių kaimynų algoritmo, o Naive Bayes nuo geriausio rezultato atsiliko 6,25 proc. punktais.

Siekiant įvertinti ne tik algoritmų prognozių tikslumą, bet ir jautrumą bei specifiškumą, buvo sudarytos sumaišymo matricos, kurių apibendrinti rezultatai yra pateikiami 9 lentelėje. Neretai jautrumas ir specifiškumas naudojamas vertinant ligų prognozes, tačiau šiuo atveju, galima pasirinkti, kuri klasė yra interpretuojama kaip teigiama. Toliau pateikiami rezultatai atspindi jautrumą ir specifiškumą, kai informacinės paramos kategorija traktuojama kaip teigiama klasė. Taigi, jautrumas atspindi, kokia dalis informacinės paramos tipo įrašų buvo priskirta informacinės paramos tipui, o specifiškumas parodo, kokia dalis emocinės paramos tipo įrašų buvo teisingai priskirti emocinės paramos tipo klasei.

9 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant įrašuose prašomos paramos tipą palyginimas

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Jautrumas	100 proc.	100 proc.	86,96 proc.
Specifiškumas	22,22 proc.	33,33 proc.	77,78 proc.

Remiantis lentelėje pateiktais rezultatais, Naive Bayes ir atraminių vektorių mašinos algoritmai turėjo 100 proc. siekiantį jautrumą, tačiau labai žemą specifiškumą. Tuo tarpu K-artimiausių kaimynų algoritmas pasižymėjo labiau subalansuotais rezultatais – jautrumas siekė apytiksliai 87 proc., o specifiškumas – 78 proc. Tai parodo, jog Naive Bayes ir atraminių vektorių mašinos algoritmai retai teisingai identifikuoja emocinės paramos tipo įrašus.

Nepaisant skirtingų rezultatų, visi trys algoritmai buvo toliau naudojami sintetinio duomenų rinkinio kūrimui. Pritaikius jau apmokytus algoritmus likusiems įrašams, buvo prognozuotos įrašuose prašytos paramos klasės. Siekiant įvertinti, kuris algoritmas pasižymėjo geriausiomis prognozėmis vertinant likusius duomenis, nedidelė dalis įrašų buvo atrinkta tikslumo apskaičiavimui. Įrašų klasės buvo įvertintos rankiniu būdu ir palygintos su algoritmų prognozėmis. Rezultatai pateikiami 10 lentelėje.

10 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas įrašams

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas	92 proc.	80 proc.	68 proc.

Nors Naive Bayes algoritmas pasižymėjo mažiausiu tikslumu vertinant testavimo imties prognozes, vertint dalį prognozių likusiame įrašų duomenų rinkinyje, algoritmo tikslumas buvo pats aukščiausias. Tam įtakos galėjo turėti stiprus klasių disbalansas originaliame duomenų rinkinyje. Tuo tarpu K-artimiausių kaimynų tikslumas buvo net 16,4 proc. punktais mažesnis nei nagrinėjant testavimo imtį. Atraminių vektorių mašinos algoritmas išlaikė panašų prognozių tikslumo lygį ir vertinant testavimo ir sintetinę duomenų imtis (skirtumas siekė 1,25 proc. punktus).

4.1.2. Atsakymų klasifikavimas pagal suteikiamos paramos tipą

Sukūrus įrašuose prašomos paramos tipo prognozes, minėti algoritmai buvo panaudoti sintetinio atsakymų ir juose teikiamos paramos tipų duomenų rinkinio kūrimui. Kaip ir įrašų atveju, nedidelė dalis duomenų buvo atrinkta apmokymo ir testavimo imtims. Atrinktų atsakymų paramos tipai buvo prognozuojami rankiniu būdu, objektyviai įvertinant atsakymų turinį. Apmokius algoritmus, modelių tikslumas buvo įvertintas panaudojant atsakymų duomenų rinkinio testavimo imtį. Rezultatai pateikiami 11 lentelėje.

11 lentelė. Algoritmų tikslumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas	56,25 proc.	78,13 proc.	78,13 proc.

Visų algoritmų tikslumas prognozuojant atsakymuose teikiamos paramos tipus buvo prastesnis nei prognozuojant įrašuose prašomos paramos tipus. Ypač žemu tikslumu pasižymėjo Naive Bayes algoritmas, kuris retai pateikdavo emocinės informacijos tipo prognozę. Siekiant išsiaiškinti ar abi kategorijos buvo prastai prognozuojamos, buvo sudarytos sumaišymo matricos kiekvieno algoritmo rezultatams. Pagal sumaišymo matricas apskaičiuoti jautrumo ir specifiškumo rodikliai pateikiami 12 lentelėje.

12 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Jautrumas	100 proc.	94,12 proc.	82,35 proc.
Specifiškumas	6,67 proc.	60 proc.	73,33 proc.

Prastus Naive Bayes ir atraminių vektorių mašinos modelių rezultatus lėmė skirtingas klasių pasiskirstymas apmokymo bei testavimo imtyse, nulemtas atsitiktinio duomenų suskirstymo į imtis. Taigi, modeliai buvo iš naujo apmokyti naudojant parametą „*stratify*“, kuris nurodė, kad apmokymo ir testavimo imtyse klasių pasiskirstymas turi atitikti bendrą klasių pasiskirstymą duomenų rinkinyje. Atliekant įrašų analizę šis parametras nebuvo įtrauktas, nes klasės atsitiktinai buvo subalansuotai

paskirstytos tarp imčių. Duomenų stratifikavimo pritaikymas leido pasiekti geresnius rezultatus. Atnaujinti algoritmų tikslumo rezultatai pateikiami 13 lentelėje.

13 lentelė. Algoritmų tikslumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas (sulyginus klasių balansą apmokymo ir testavimo imtyse)

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas	71,88 proc.	84,34 proc.	87,5 proc.

Visų algoritmų prognozių tikslumas pagerėjo, ypač Naive Bayes algoritmo. Lyginant su algoritmo prognozių tikslumo įverčiu 11 lentelėje, tikslumas pakilo 15,63 proc. punktais. K-artimiausių kaimynų algoritmo tikslumas lyginant su rezultatu gautu nestratifikuotą duomenų imtį, pakilo 9,37 proc. punktais ir buvo pats aukščiausias tarp trijų vertintų algoritmų rezultatų. Nuo K-artimiausių kaimynų algoritmo rezultato 3,16 proc. punkto atsiliko atraminių vektorių mašinos algoritmas, o Naive Bayes algoritmas pateikė prasčiausius rezultatus. 14 lentelėje pateikiami atnaujinti jautrumo bei specifiškumo rezultatai.

14 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant atsakymuose suteikiamos paramos tipą palyginimas (sulyginus klasių balansą apmokymo ir testavimo imtyse)

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Jautrumas	100 proc.	100 proc.	90,48 proc.
Specifiškumas	18,18 proc.	54,54 proc.	81,81 proc.

Lyginant duomenis 14 ir 12 lentelėse pastebėta, jog sumažėjo tik atraminių vektorių mašinos algoritmo prognozių specifiškumas, visi kiti rodikliai padidėjo arba išliko tokie patys. Tačiau Naive Bayes specifiškumo rezultatai indikuoja, jog didžioji dalis emocinės paramos atvejų naudojant šį algoritmą nebūtų identifikuojami.

Naujai apmokyti algoritmai buvo panaudoti likusių duomenų paramos kategorijų prognozėms parengti. Kaip ir įrašų atveju, dalis įrašų buvo atrinkti, jų klasės priskirtos rankiniu būdu, siekiant įvertinti sukurtų sintetinių duomenų rinkinių tikslumą. Rezultatai pateikiami 15 lentelėje.

15 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas atsakymams

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas	85 proc.	85 proc.	80 proc.

Nors Naive Bayes algoritmas pasižymėjo mažiausiu tikslumu vertinant testavimo imties, algoritmo rezultatų tikslumas vertinant dalį sukurtos sintetinės duomenų imties buvo didesnis nei K-artimiausių kaimynų algoritmo ir sutapo su atraminių vektorių mašinos algoritmo rezultatais. Kaip ir įrašų duomenų vertinimo atveju, tokiems galutiniams rezultatams įtaką turėjo mažesnė emocinės paramos atvejų koncentracija duomenų imtyje.

Prieš įvertinant galutinę paramos prašymų ir suteikimų pasiskirstymą tarp emocinės ir informacinės paramos, buvo išbandytas duomenų apmokymas remiantis bendra (įrašų ir atsakymų) imtimi. Sujungus visas ankstesniuose žingsniuose naudotas apmokymo bei testavimo imtis bei jas kartu

konvertavus į vektorius vėl buvo apmokyti Naive Bayes, atraminių vektorių mašinos bei K-artimiausių kaimynų algoritmai. Dėl platesnės apmokymo imties, papildžiusio duomenų žodyno bei atsinaujinusių termino dažnio – atvirktinio dokumentų dažnio balų, nauji rezultatai gali reikšmingai skirtis nuo rezultatų, gautų remiantis tik atsakymų arba įrašų imtimis. Testavimo imties tikslumo rezultatai pateikiami 16 lentelėje.

16 lentelė. Algoritmų tikslumo prognozuojant paramos tipą bendroje imtyje palyginimas

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas	69,84 proc.	79,37 proc.	92,06 proc.

Remiantis gautais rezultatais, K-artimiausių kaimynų algoritmo tikslumas buvo pats aukščiausias. Antroje vietoje buvo atraminių vektorių mašinos algoritmas, kurio tikslumas buvo 12,69 proc. punkto žemesnis nei K-artimiausių kaimynų algoritmo, o trečioje vietoje liko Naive Bayes algoritmas, kurio rezultatas nuo geriausio skyrėsi 22,22 proc. punktais. Kadangi tikslumas pateiktas visai imčiai, aprėpiančiai ir atsakymus ir įrašus, algoritmų tikslumo rezultatai nėra lyginami su anksčiau gautais atskirais įrašų ir atsakymų tikslumo rodikliais. 17 lentelėje pateikiami jautrumo bei specifiškumo rezultatai gauti remiantis skirtingų algoritmų parengtų prognozių sumaišymo matricomis, gautomis vertinant bendrą duomenų imtį.

17 lentelė. Algoritmų jautrumo ir specifiškumo prognozuojant paramos tipą bendroje imtyje palyginimas

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Jautrumas	100 proc.	100 proc.	92,86 proc.
Specifiškumas	9,52 proc.	38,10 proc.	90,48 proc.

Nors Naive Bayes ir atraminių vektorių mašinos algoritmų sukurtų prognozių jautrumas siekė 100 proc., specifiškumas buvo labai žemas. Tuo tarpu K-artimiausių kaimynų algoritmas pasižymėjo gana aukštu jautrumu ir specifiškumu (>90 proc.). Jautrumo arba specifiškumo svarba vertinant algoritmą priklauso ir nuo vertinamų duomenų. Atvejais, kai teigiama klasė identifikuojama ligos, pavyzdžiui, užkrečiamos infekcijos, atvejį, 100 proc. siekiantis jautrumas gali būti kur kas svarbesnis, net jei tai reiškia žemą specifiškumą, nes pakankamai aukštas jautrumas leistų išvengti ligos plitimo. Tačiau šiuo atveju, kai abi klasės yra svarbios tyrimo apimtyje, ir jautrumas, ir specifiškumas gali būti laikomi vienodai svarbiais.

Apmokyti Naive Bayes, atraminių vektorių ir K-artimiausių kaimynų algoritmai buvo panaudoti prognozuojant likusių įrašų ir atsakymų paramos tipus. Algoritmai pritaikyti atskirai likusių įrašų ir atsakymų duomenų imtims. Sintetinių duomenų rinkinių dalies tikslumo vertinimo rezultatai pateikiami 18 lentelėje.

18 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas įrašams ir atsakymams

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas (įrašai)	84 proc.	84 proc.	76 proc.
Tikslumas (atsakymai)	85 proc.	85 proc.	70 proc.

Rezultatai parodo, kad nepaisant gerų K-artimiausių kaimynų algoritmo rezultatų vertinant testavimo imties prognozes, sintetinę imtį tiksliau sudarė Naive Bayes bei atraminių vektorių mašinos algoritmai. Nors K-artimiausių kaimynų algoritmo jautrumo ir specifiškumo rezultatai buvo gana aukšti vertinant testavimo imtį, vertinant sintetinių duomenų rinkinį pastebėta, jog algoritmas neretai priskiria emocinės paramos tipą žinutėms, kuriose iš tikrųjų ieškoma informacinės pagalbos arba ji suteikiama.

Siekiant testuoti algoritmų suklasifikuotų duomenų analizę, buvo pasirinktos geriausių tikslumą turinčios duomenų imtys. Visi sintetinių duomenų imčių tikslumo rodikliai lyginami toliau pateikiamoje 19 lentelėje. Aukščiausi tikslumo rezultatai įrašų ir atsakymų imtyse išskirti paryškintu ir palinkusiu šriftu.

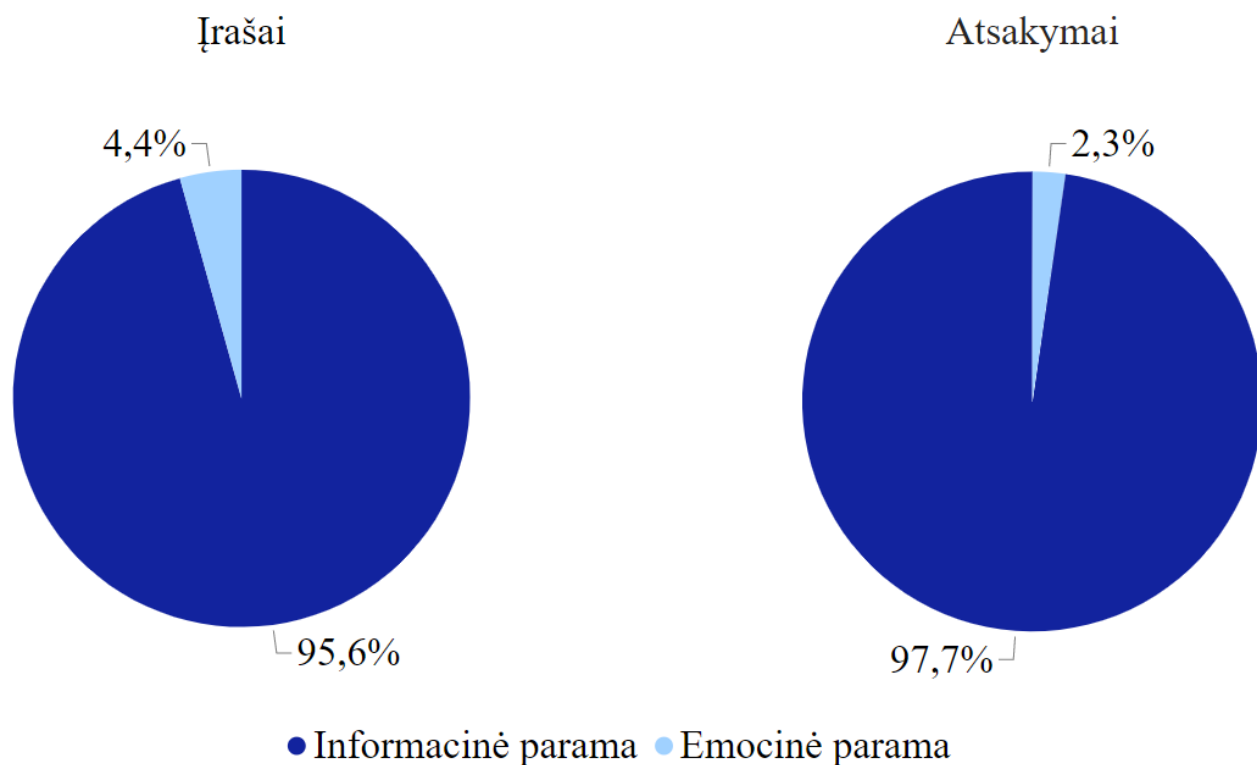
19 lentelė. Algoritmų tikslumo palyginimas vertinant paramos klases priskirtas įrašams ir atsakymams

Algoritmas	Naive Bayes	Atraminių vektorių mašinos algoritmas	K-artimiausių kaimynų algoritmas
Tikslumas (įrašai, remiantis įrašų imtimi)	<i>92 proc.</i>	80 proc.	68 proc.
Tikslumas (atsakymai, remiantis atsakymų imtimi)	<i>85 proc.</i>	<i>85 proc.</i>	80 proc.
Tikslumas (įrašai, remiantis bendra imtimi)	84 proc.	84 proc.	76 proc.
Tikslumas (atsakymai, remiantis bendra imtimi)	<i>85 proc.</i>	<i>85 proc.</i>	70 proc.

Geriausi rezultatai prognozuojant likusiuose įrašuose išreikštas paramos klases buvo gauti naudojant Naive Bayes algoritmą (92 proc.), o prognozuojant atsakymuose suteikiamos paramos tipus, Naive Bayes ir atraminių vektorių mašinos metodų tikslumo rezultatai buvo vienodi (85 proc.). Taigi, tolimesnei analizei naudojamas įrašų rinkinys su Naive Bayes algoritmo parengtomis paramos tipo prognozėmis bei atsakymų rinkinys su atraminių vektorių mašinos algoritmo parengtomis prognozėmis.

4.2. Sintetinių duomenų rinkinių analizė

Pasirinkus tiksliausias prognozes bei apjungus visus turimus duomenis, buvo atlikta tolimesnė duomenų analizė. Sukurtų sintetinių duomenų rinkinių analizė atskleidė, jog nagrinėtuose įrašuose ir atsakymuose didžiąją dalį žinučių sudarė informacinės pagalbos prašymai ir jos suteikimai. Įrašų ir atsakymų pasiskirstymas pagal paramos tipą pateikiamas 18 pav.

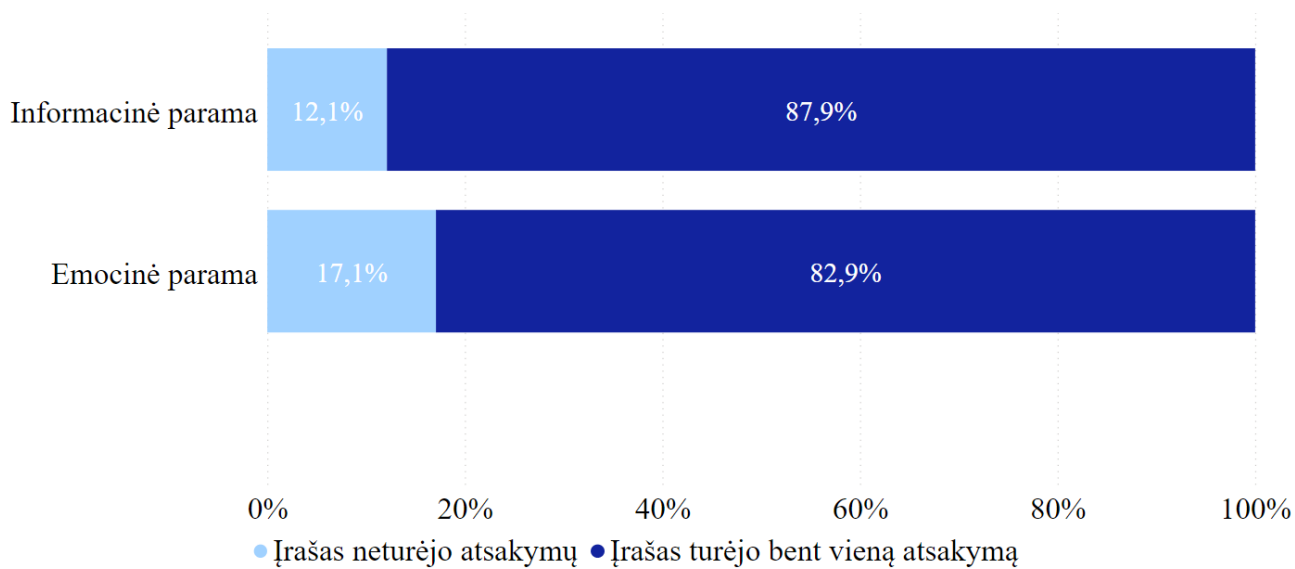


18 pav. Įrašų pasiskirstymas pagal prašomos paramos tipą, proc.

Remiantis vien informacinės ir emocinės paramos proporcijomis įrašų ir atsakymų tarpe, būtų galima daryti prielaidą, jog didesnė dalis emocinės paramos prašymų nesulaukė atsakymų. Atsižvelgus į forumo, iš kurio buvo renkami duomenys, pobūdį, pasiskirstymas tarp informacinės bei emocinės paramos atrodo logiškas. Galima daryti prielaidą, jog instrumentinės bei emocinės paramos asmenys dažniau yra linkę ieškoti artimoje aplinkoje, be to, gyvai gaunama instrumentinė ar emocinė parama yra vertinama kaip naudingesnė [68].

4.2.1. Sintetinio įrašų duomenų rinkinio analizė

Nagrinėjant sudarytą įrašų duomenų rinkinį, visų pirma buvo įvertinta, kokia dalis įrašų nesulaukė atsakymo skirtingose paramos grupėse. Rezultatai pateikiami 19 pav.



19 pav. Įrašų pasiskirstymas remiantis kitų asmenų paliktais komentarais, proc.

12,1 proc. įrašų, kuriuose buvo prašoma informacinės pagalbos, buvo neatsakyti. Atsakymo nesulaukusių emocinės paramos prašymų dalis siekė 17,1 proc. Taigi, neatsakytų įrašų, kuriuose buvo prašoma emocinės paramos, dalis buvo 5 proc. punktais didesnė nei neatsakytų įrašų, kuriuose buvo prašoma informacinės pagalbos, dalis. Suteikus įrašams, kurie neturėjo atsakymų, nulinę vertę, o įrašus, turinčius bent vieną atsakymą, įvertinus vienetu, buvo įvertintas skirtumas tarp įverčių vidurkių emocinės ir informacinės paramos prašymų tarpe. Gauti statistinio palyginimo rezultatai pateikiami 20 lentelėje.

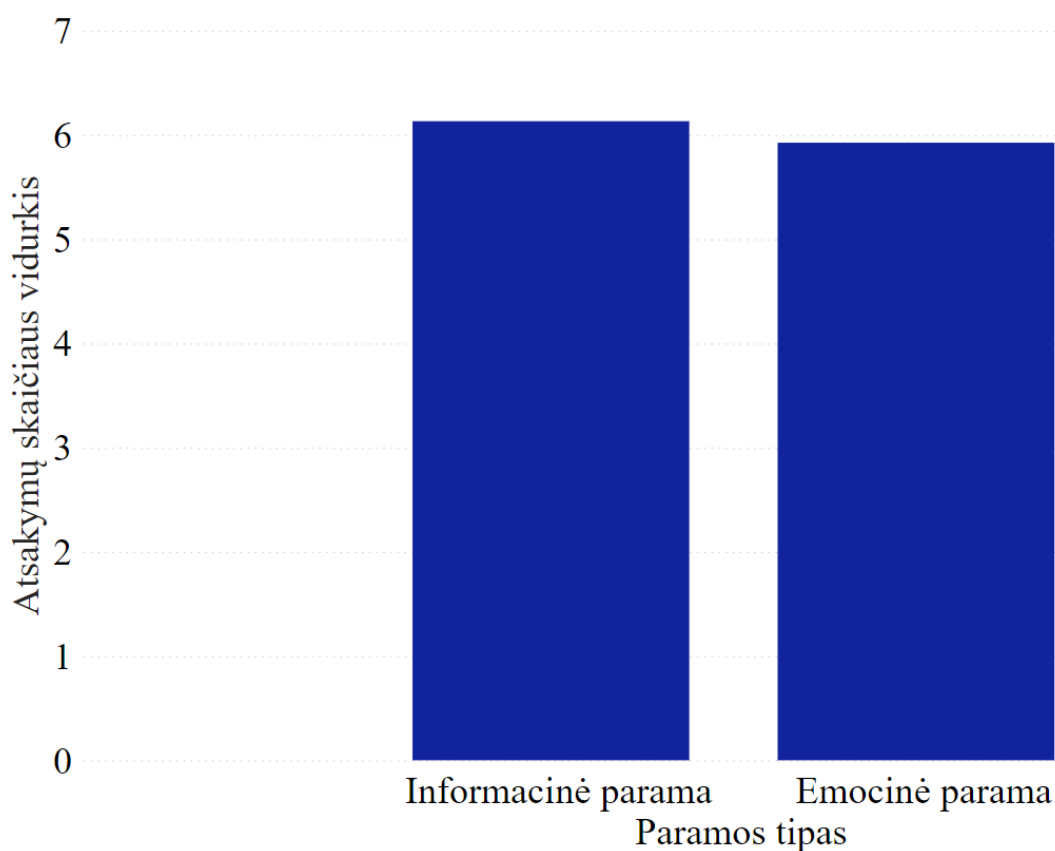
20 lentelė. Įrašų be atsakymo įverčio palyginimas tarp emocinės ir informacinės paramos prašymų

Nulinė hipotezė: vidurkių skirtumas = 0		
1 imtis:		
n = 900	vidurkis = 0,878889	standartinis nuokrypis = 0,326438
standartinė vidurkio paklaida = 0,0108813		
95 proc. Pasikliautinumo intervalas vidurkiui: nuo 0,857533 iki 0,900245		
2 imtis:		
n = 41	vidurkis = 0,829268	standartinis nuokrypis = 0,380949
standartinė vidurkio paklaida = 0,0594942		
95 proc. Pasikliautinumo intervalas vidurkiui: nuo 0,709026 iki 0,949511		
Testo statistiniai rezultatai:		
$t(939) = (0,878889 - 0,829268) / (0,0525296) = 0,944622$		
Dvipusė p-reiškė = 0,3451		Vienpusė p-reiškė = 0,1725

Remiantis gauta p-reikšme, kuri viršija 0,05 statistinio reikšmingumo lygį, nulinė lygių vidurkių hipotezė yra neatmetama. Taigi, skirtumas nėra pakankamai reikšmingas, kad būtų galima teigti, jog emocinės paramos prašymai rečiau sulaukia atsakymų nei informacinės paramos prašymai.

Bendrai vertinant neatsakytų įrašų kiekį, galima daryti prielaidą, jog visose viešose platformose yra rizika nesulaukti atsakymo. Tačiau tokie atvejai gali mažinti migrantų motyvaciją ieškoti paramos internete, taip sumažinant jų gaunamos paramos kiekį. Informacinės paramos atveju ekspatriantai bei kiti migrantai gali lengviau surasti kitus paramos resursus ir tęsti informacijos paiešką, tačiau emocinės paramos ieškojimo atveju neigiamos patirtys gali paskatinti emocinių problemų vystymąsi vietoj to, kad prisidėtų prie jų sprendimo.

Analizuojant įrašų ir atsakymų duomenų rinkinius taip pat buvo įvertinta, kiek atsakymų vidutiniškai sulaukia vienas įrašas. Bendroje imtyje šis skaičius siekė 6,12 įrašo. Rezultatai informacinės ir emocinės paramos prašančių įrašų imtyse pavaizduoti 20 pav.



20 pav. Vidutinis vieno įrašo sulauktų atsakymų skaičius

Įrašai, kuriuose buvo prašoma informacinio tipo paramos, vidutiniškai sulaukdavo 6,13 atsakymų, o įrašai, kuriuose buvo prašoma emocinio paramos tipo, vidutiniškai sulaukdavo 5,93 atsakymų. Iš duomenų imties pašalinus įrašus, kurie neturėjo nė vieno atsakymo, bendras vidutinis sulauktų atsakymų skaičius siekė 7,29. Informacinės paramos prašymai vidutiniškai sulaukdavo 7,33 atsakymų, o emocinės paramos prašymai vidutiniškai sulaukdavo 6,93 atsakymų.

Skirtumas tarp vidurkių vertinant visus įrašus siekė 0,2, tačiau, siekiant įvertinti, ar tokio dydžio skirtumas tikrai nėra statistiškai reikšmingas, buvo atliktas t-testas, kurio rezultatai pateikiami 22 lentelėje.

21 lentelė. Atsakymų vidurkio palyginimas tarp emocinės ir informacinės paramos prašymų

Nulinė hipotezė: vidurkių skirtumas = 0

1 imtis:		
n = 900	vidurkis = 6,13333	standartinis nuokrypis = 9,25357
standartinė vidurkio paklaida = 0,308452		
95 proc. Pasikliautinumo intervalas vidurkiui: nuo 5,52796 iki 6,7387		

2 imtis:		
n = 41	vidurkis = 5,92683	standartinis nuokrypis = 9,19073
standartinė vidurkio paklaida = 1,43535		
95 proc. Pasikliautinumo intervalas vidurkiui: nuo 3,02588 iki 8,82778		

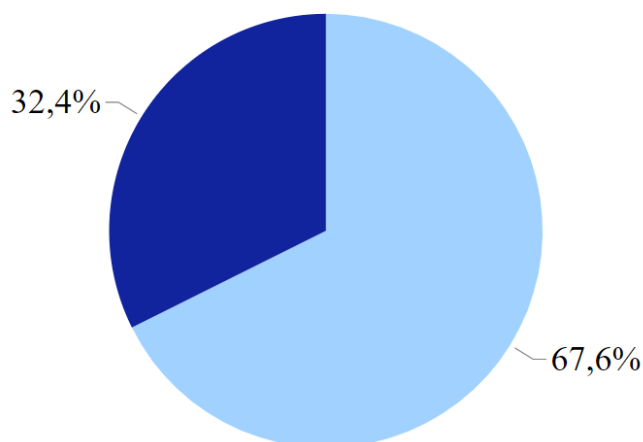
Testo statistiniai rezultatai:	
$t(939) = (6,13333 - 5,92683) / (1,47729) = 0,139786$	
Dvipusė p-reikšmė = 0,8889	Vienpusė p-reikšmė = 0,4444

Remiantis aukšta p-reikšme, nulinė hipotezė negali būti atmesta, taigi, skirtumas tarp vidutinio atsakymų skaičiaus emocinės ir informacinės paramos įrašų tarpe laikomas statistiškai nereikšmingu. Taigi, negalima teigti, jog vieno ar kito socialinės paramos tipo prašymai sulaukia daugiau atsakymų, nei kiti.

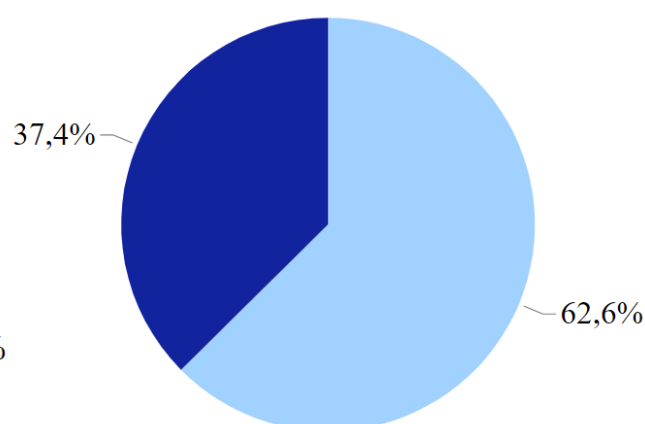
Kaip ir vertinant originalius duomenis, atliekant sintetinio įrašų duomenų rinkinio analizę buvo įvertinti įrašai, kurių autoriai vėliau pakomentavo juos naudodami padėką išreiškiančius žodžius, taip indikuodami, kad jie sulaukė naudingos pagalbos. Įvertinus įrašų dalį, kurie neturėjo jokio atsakymo bei atskyrus įrašus, kuriuose prašoma emocinės bei instrumentinės paramos, parengti grafikai vaizduojami 21 pav.

Gauti rezultatai turėtų būti interpretuojami kaip galimai turintys paklaidą, nes dalis įrašų autorių, gavusių pagalbą, galėjo neišreikšti dėkingumo komentuose, arba, išreikšta padėka galėjo atspindėti asmens mandagumą net jei gauta pagalba neatitiko jo poreikių. Tačiau, neatliekant forumo narių, ieškojusių socialinės paramos, apklausos, neįmanoma išsiaiškinti kaip asmenys vertino suteiktą paramą, taigi, vertinama dėkingumo ir naudingos pagalbos sąryšio prielaida.

Emocinė parama



Informacinė parama



- Komentaruose nebuvo autoriaus padėkos
- Autorius padėjo įrašo komentaruose

21 pav. Įrašų pasiskirstymas pagal įrašų autorių išreikštą padėką, proc.

Remiantis skritulinėse diagramose pateikta informacija, beveik trečdaliu atvejų, kai buvo prašoma emocinės paramos, įrašo autorius padėjo už atsakymus komentaruose. Dėkojančių autorių dalis buvo šiek tiek didesnė tarp informacinės paramos prašymų ir siekė 37,4 proc. Siekiant įvertinti ar skirtumas tarp emocinės ir informacinės paramos yra reikšmingas, buvo palyginti padėkų vidurkiai abiejose imtyse (siekiant atlikti palyginimą, įrašams, kurie turėjo autoriaus padėką atsakymuose, buvo suteikiama vieneto reikšmė, o likusiems įrašams, kurie autoriaus padėkos neturėjo – nulio reikšmė). Rezultatai pateikiami 22 lentelėje.

22 lentelė. Padėkų vidurkio palyginimas tarp emocinės ir informacinės paramos prašymų

Nulinė hipotezė: vidurkių skirtumas = 0

1 imtis:

n = 791	vidurkis = 0,37421	standartinis nuokrypis = 0,484224
---------	--------------------	-----------------------------------

standartinė vidurkio paklaida = 0,017217

95 proc. Pasikliautinimo intervalas vidurkiui: nuo 0,340413 iki 0,408006

2 imtis:

n = 34	vidurkis = 0,323529	standartinis nuokrypis = 0,474858
--------	---------------------	-----------------------------------

standartinė vidurkio paklaida = 0,0814375

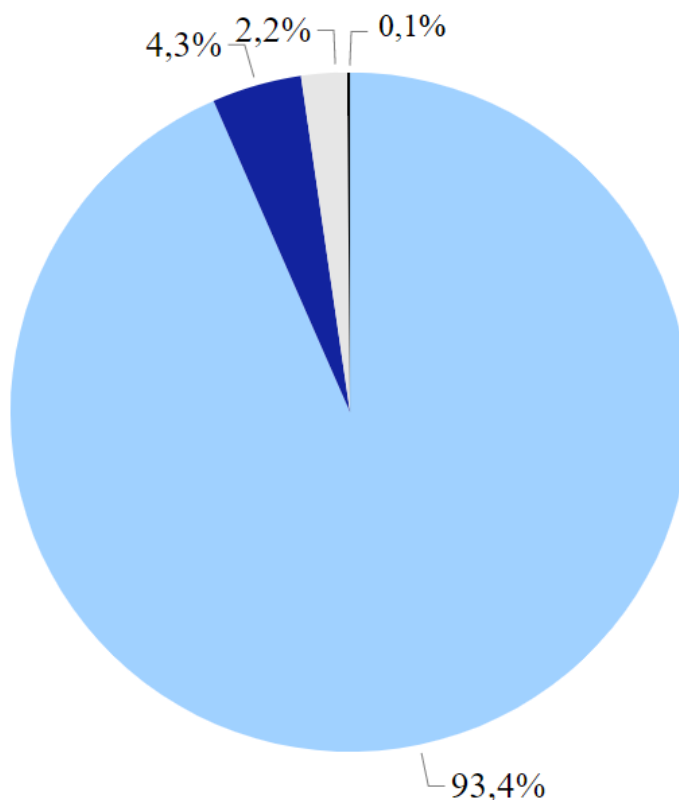
95 proc. Pasikliautinimo intervalas vidurkiui: nuo 0,157844 iki 0,489215

Testo statistiniai rezultatai:	
$t(823) = (0,37421 - 0,323529) / (0,0847446) = 0,598037$	
Dvipusė p-reiškė = 0,55	Vienpusė p-reiškė = 0,275

Remiantis apskaičiuota p-reiškė, kuri viršija 0,05 reikšmingumo lygmenį, nulinės hipotezės negalima pagrįstai atmesti, taigi, skirtumas tarp vidurkių vertinamas kaip nereikšmingas. Prie šio rezultato prisideda ir stiprus disbalansas tarp emocinės paramos ir instrumentinės paramos prašymų skaičiaus.

4.2.2. Sintetinio atsakymų duomenų rinkinio analizė

Analizuojant atsakymų duomenų rinkinį nebuvo galima įvertinti tų pačių pjūvių, kurie buvo atlikti įrašų duomenų rinkiniui, nes atsakymai buvo talpinami viename lygyje (nebuvo galima atsakyti į atsakymus). Tačiau buvo įvertinta ar atsakyme suteiktos paramos tipas atitiko įrašė prašomos paramos tipą. Rezultatai pateikiami 22 pav.



- Prašoma ir suteikiama informacinė parama
- Prašoma emocinės paramos, bet suteikiama informacinė parama
- Prašoma informacinės paramos, bet suteikiama emocinė parama
- Prašoma ir suteikiama emocinė parama

22 pav. Prašomos ir suteikiamos paramos tipų sutapimas

Didžiausią dalį atsakymų, kaip ir įrašų, sudarė informacinės paramos tipo žinutės, taigi, didžioji dalis atsakymų (93,4 proc.) pateko į kategoriją, apibrėžiančią atvejį, kai buvo prašoma informacinės paramos ir ji buvo suteikta. Nors emocinės paramos tipo žinučių įrašų bei atsakymų tarpe dalys buvo panašios (4,4 proc. ir 2,3 proc.), įvertinus kurie atsakymai buvo skirti konkrečioms žinutėms, atvejai, kai buvo prašoma ir suteikiama emocinė pagalba, sudarė vos 0,1 proc. visų atsakymų. Daugiau nei 6 proc. atsakymų neatitiko prašomos paramos tipo – 4,3 proc. atvejų, forumo nariams paprašius emocinės paramos, jiems suteikta informacinė, o 2,2 proc. atvejų, forumo nariams paprašius informacinės paramos, jiems buvo suteikta emocinė.

Nepaisant paramos tipo neatitinkančių atsakymų, net 80 proc. įrašų sulaukė bent vieno jų prašomos paramos tipą atitinkančio atsakymo. Vertinant vien įrašus, kurie sulaukė atsakymų, tinkamo paramos tipo atsakymų turinčių įrašų dalis siekia 91 proc.

4.3. Tyrimo apibendrinimas ir diskusija

Apibendrinant įgyvendintą tyrimą galima teigti, jog viešuose forumuose dažniausiai prašoma (95,6 proc. atvejų) ir suteikiama (97,7 proc. atvejų) informacinė parama. Likusiais atvejais buvo prašoma emocinės paramos (4,4 proc. atvejų) arba ji buvo suteikiama (2,3 proc. atvejų). Instrumentinės paramos suteikimo arba gavimo atvejai nebuvo identifikuoti duomenų rinkinyje. Ir emocinės, ir informacinės paramos prašymai sulaukė panašaus atsakymų dažnio – forumo nariai atsakė į 87,9 proc. informacinės pagalbos prašymų ir į 82,9 proc. emocinės pagalbos prašymų. Vienas įrašas vidutiniškai sulaukdavo 6 atsakymų. Statistiškai reikšmingi skirtumai tarp skirtingų paramos tipų imčių nebuvo nustatyti ir vertinant autorių padėkomis pakomentuotų įrašų dalį. 93,5 proc. visų nagrinėtų atsakymų suteikiamas paramos tipas atitiko įrašuose prašomą paramos tipą. Prašomos ir suteikiamos paramos neatitikimo atvejais dažniau buvo suteikiama informacinė parama vietoj prašomos emocinės (4,3 proc.), beveik dvigubai rečiau – emocinė vietoj prašomos instrumentinės (2,2 proc.).

Tyrimo rezultatai atitinka iki šiol akademinėje literatūroje matytas tendencijas. Dažniausiai tyrimo atveju prašoma ir suteikiama informacinė parama laikoma vienintele socialinės pagalbos rūšimi, kurią pakankamai lengva teikti naudojant socialinius tinklus, netaikant hibridinio bendravimo [12].

Nors tyrime nagrinėti duomenys parodė, kad emocinės paramos forumo nariai ieškojo retai ir ją retai gavo, būtų naudinga įvertinti ir socialinių tinklų, kuriuose parama teikiama ne tik viešinamomis žinutėmis, bet ir reakcijomis (pavyzdžiui, Facebook mygtukai patinka, labai patinka, rūpi ir pan.), duomenis. Tokiose platformose būtų lengviau sulaukti emocinės paramos, jeigu ir reakcijos yra interpretuojamos kaip socialinės paramos forma [25].

Tyrimas leidžia patvirtinti ir Hayeso bei kolegų pateiktą išvalgą, jog instrumentinės paramos socialiniuose tinkluose asmenys ieškotų nebent bendraudami tiesiogiai su konkrečiu asmeniu (pavyzdžiui, rašydami asmenines žinutes asmenims savo tinkle), dažniau draugu, o ne nepažįstamu platformos nariu [25]. Tyrime nagrinėjami duomenys iš atviro forumo, kuriame įrašus ir atsakymus talpinti gali asmenys iš bet kurios pasaulio valstybės. Nors yra galimybė, kad kai kurie forumo nariai bendravo hibridiniu būdu ir buvo gyvai susitikę (ypač tais atvejais, kai buvo prašoma emocinės paramos ir siūloma susitikti bei kartu pasidalinti savo gyvenimo patirtimi), tyrime nagrinėjama tik nuotoliniu būdu teikiama socialinė pagalba. Taigi, paramos teikimo būdas apriboja galimos suteikti paramos tipą – informacinė parama ieškoma dažniau, nes yra nedidelis šansas, kad instrumentinė

parama (pavyzdžiui, trumpalaikė neapmokama vaikų ar naminių gyvūnų priežiūra, pinigų paskolinimas, maisto pagaminimas ar kt.) bus suteikta. Galima daryti prielaidą, jog retą arba neegzistuojantį instrumentinės paramos ieškojimą viešuose forumuose lemia silpni forumo narių tarpusavio ryšiai ir skirtingos asmenų gyvenamosios vietos.

Remiantis nagrinėta literatūra ir tyrimo rezultatais, galima daryti prielaidą, jog daugiau emocinės ir instrumentinės paramos būtų prašoma ir suteikiama uždaroje bendruomenėje (pavyzdžiui, privačiose Facebook grupėse), ypač jei dalijamasi siauresnėje asmenų grupėje (pavyzdžiui, vien tik tarp kitų migrantų iš tos pačios valstybės, atvykusių į tą patį miestą).

Rekomenduojama šiame poskyryje pristatytas prielaidas dėl asmenų ryšio stiprumo bei dalijimosi platformos tipo (uždara / atvira, mažas / didelis narių skaičius, plati / siaura demografinė auditorija grupėje) nagrinėti tolimesniuose akademinuose tyrimuose.

Išvados

1. Remiantis atlikta literatūros analize, ekspatriantai bei kiti migrantai gali sulaukti visų trijų socialinės paramos rūšių (informacinės, emocinės, instrumentinės).
2. Socialiniuose tinkluose suteikiama ir gaunama parama gali padėti migrantams lengviau prisitaikyti, užmegzti ryšius, pagerinti jų emocinę savijautą.
3. Atlikus literatūros šaltinių analizę, parengtas konceptualus tyrimo modelis, kurį sudaro įrašų surinkimo, teksto valymo, rekonstravimo, analizės bei ieškomos ir teikiamos paramos identifikavimo žingsniai.
4. Ekspatriantams socialiniuose tinkluose teikiamos ir jų sulaukiamos paramos tipo identifikavimui gali būti naudojami įvairūs klasifikavimo algoritmai.
5. Tyrimo metu nustatyta, jog naudojant mokomąją duomenų dalį, geriausiai klasifikaciją atlieka K-artimiausių kaimynų algoritmas, tačiau bendroje duomenų imtyje tikslesnes prognozes pateikia atraminių vektorių mašinos bei Naive Bayes algoritmai.
6. Ekspatriantų bendravimui skirtame viešame forume didžioji dalis narių ieško informacinės paramos ir ją gauna. Kur kas rečiau ieškoma ir gaunama emocinė parama, o instrumentinės paramos ieškojimo arba suteikimo atvejai nebuvo identifikuoti. Dažniausiai nagrinėjamos temos apie darbą, persikėlimui reikalingus dokumentus bei gyvenamosios vietos pasirinkimą.

Literatūros sąrašas

1. ABDEL-HALIM, A. Social Support and Managerial Affective Responses to Job Stress. *Journal of Occupational Behaviour*. 1982, **3**(4), 281-295. Prieiga per: <http://www.jstor.org/stable/3000115>
2. ADELMAN, M. B. Cross-cultural adjustment: A theoretical perspective on social support. *International Journal of Intercultural Relations*. 1988, **12**(3), 183–204. Prieiga per: [https://doi.org/10.1016/0147-1767\(88\)90015-6](https://doi.org/10.1016/0147-1767(88)90015-6)
3. AKSOY, Z. Exploring Intercultural Communication in Expats' Social Media Groups. *International Conference on Studies in Humanities and Social Sciences (SHSS-2016)*. DOI:10.17758/ERPUB.EA0516088
4. AL-NALDAWI, R. Migration-related stress and psychosomatic consequences. *International Congress Series*. 2002, **1241**, 271-278. Prieiga per: [https://doi.org/10.1016/S0531-5131\(02\)00649-0](https://doi.org/10.1016/S0531-5131(02)00649-0)
5. ANDRESEN, M., F. BERGDOLT, J. MARGENFELD ir M. DICKMANN. Addressing international mobility confusion – developing definitions and differentiations for self-initiated and assigned expatriates as well as migrants. *The International Journal of Human Resource Management*. 2014, **25**(16), 2295-2318. Prieiga per: <https://doi.org/10.1080/09585192.2013.877058>
6. ASHTON, D. N. ir M. J. MAGUIRE. Dual Labour Market Theory and the Organisation of Local Labour Markets. *International Journal of Social Economics*, 1984, **11**(7), 106-120. Prieiga per: <https://doi.org/10.1108/eb013980>
7. BAYES, T. An Essay towards solving a Problem in the Doctrine of Chances. *Philosophical Transactions*. 1763, **53**, 370–418. DOI: doi:10.1098/rstl.1763.0053
8. BERGER, C. R. ir R. J. CALABRESE. Some Explorations in Initial Interaction and Beyond: Toward a Developmental Theory of Interpersonal Communication. *Human Communication Research*. 1975, **1**(2), 99–112. Prieiga per: <https://academic.oup.com/hcr/article-abstract/1/2/99/4637500>
9. BROWNLEE, J. What are word embeddings for text? *Machine Learning Mastery*. 2019. Prieiga per: <https://machinelearningmastery.com/what-are-word-embeddings/>
10. BZDOK, D., M. KRZYWINSKI ir N. ALTMAN. Machine learning: Supervised methods. *Nature Methods*. 2018, **15**(1), 5–6. Prieiga per: <https://doi.org/10.1038/nmeth.4551>
11. CALIGIURI, P. ir M. LAZAROVA. A model for the influence of social interaction and social support on female expatriates' cross-cultural adjustment. *International Journal of Human Resource Management*. 2002, **13**(5), 761-772. Prieiga per: <https://doi.org/10.1080/09585190210125903>
12. CANHILAL, K., B. CANBOY ir T. BAKICI. Social support for expatriates through virtual platforms: exploring the role of online and offline participation. *The International Journal of Human Resource Management*. 2020, **1**(32). Prieiga per: <https://doi.org/10.1080/09585192.2020.1752283>
13. CHEN, H., S. HU, R. HUA ir X. ZHAO. Improved naive Bayes classification algorithm for Traffic Risk Management. *EURASIP Journal on Advances in Signal Processing*, 2021, **2021**(1). Prieiga per: <https://doi.org/10.1186/s13634-021-00742-6>
14. CHEN, W. Ir A. S. K. CHOI. Internet and social support among Chinese migrants in Singapore. *New Media & Society*. 2011, **13**(7), 1067–1084. Prieiga per: <https://doi.org/10.1177/1461444810396311>
15. CHOUDHURY, M. ir E. KICIMAN. The Language of Social Support in Social Media and its Effect on Suicidal Ideation Risk. *International AAAI Conference on Weblogs and Social Media*. 2017, 32-41.

16. COHEN, S., B. H. GOTTLIEB ir L. G. UNDERWOOD. Social relationships and health. Iš COHEN, S., B. H. GOTTLIEB ir L. G. UNDERWOOD, *Social support measurement and intervention: A guide for health and social scientists*. Oxford University Press, 2020, 3–25. Prieiga per: <https://doi.org/10.1093/med:psych/9780195126709.003.0001>
17. DEY, L., S. CHAKRABORTY, A/ BISWAS, B. BOSE ir S. TIWARI. Sentiment analysis of review datasets using naïve Bayes' and k-nn classifier. *International Journal of Information Engineering and Electronic Business*. 2016 **8**(4), 54–62. Prieiga per: <https://doi.org/10.5815/ijieeb.2016.04.07>
18. DEKKER, R. ir G. ENGBERSEN. How social media transform migrant networks and facilitate migration. *Global networks*. 2013, **14**(4), 401-418. Prieiga per: <https://doi.org/10.1111/glob.12040>
19. DENISI, A. S. ir S. SONESEH. Success and failure in international assignments: A review and a proposed multi-dimensional model. *Journal of Global Mobility*. 2016, **4**(4), 386-407. Prieiga per: <https://doi.org/10.1108/JGM-12-2015-0060>
20. DOKI, S., S. SASAHARA ir I. MATSUZAKI. Stress of working abroad: a systematic review. *International archives of occupational and environmental health*. 2018, **91**(7), 767–784. Prieiga per: <https://doi.org/10.1007/s00420-018-1333-4>
21. DU PLOOY, D. R., A. LYONS ir E. S. KASHIMA. The Effect of Social Support on Psychological Flourishing and Distress Among Migrants in Australia. *Journal of immigrant and minority health*. 2019, **21**(2), 278–289. Prieiga per: <https://doi.org/10.1007/s10903-018-0745-2>
22. FARH, C. I. C., K. M. BARTOL, D. L. SHAPIRO ir J. SHIN. Networking Abroad: A Process Model of How Expatriates Form Support Ties to Facilitate Adjustment. *Academy of Management Review*. 2017, **35**(3), 434–454. Prieiga per: <https://doi.org/10.5465/amr.35.3.zok434>
23. FIX, E. ir J. L. HODGES. Discriminatory Analysis - Nonparametric Discrimination: Consistency Properties. *International Statistical Review*. 1989, **57**, 238. Prieiga per: <https://doi.org/10.2307/1403797>
24. GU, F., S. SARKAR, Y. SUN, H. WU, ir K. WU. NLP: Word embedding techniques for text analysis. *Medium*. 2020. Prieiga per: <https://medium.com/sfu-cspmp/nlp-word-embedding-techniques-for-text-analysis-ec4e91bb886f>
25. HAYES, R. A., C. T. CARR ir Y. WOHN. It's the Audience: Differences in Social Support Across Social Media. *Social Media + Society*. 2016, **2**(4). Prieiga per: <https://doi.org/10.1177/2056305116678894>
26. HÄMMIG, O. Health and well-being at work: The key role of supervisor support. *SSM - Population Health*. 2017, **3**, 393-402. ISSN 2352-8273. Prieiga per: <https://doi.org/10.1016/j.ssmph.2017.04.002>
27. HEDEMALM, A., M. SCHAUFELBERGER ir I. EKMAN. Reported Adherence and Social Support Among Immigrants Compared to Native Swedes. *Journal Immigrant Minority Health*. 2010, **12**, 124. Prieiga per: <https://doi.org/10.1007/s10903-009-9235-x>
28. HERNÁNDEZ-PLAZA, S., E. ALONSO-MORILLEJO ir C. POZO-MUÑOZ. Social Support Interventions in Migrant Populations. *The British Journal of Social Work*. 2006, **36**(7), 1151-1169. Prieiga per: <http://www.jstor.org/stable/23721354>
29. HOFHUIS, J., K. HANKE ir T. RUTTEN. Social network sites and acculturation of international sojourners in the Netherlands: The mediating role of psychological alienation and online social support. *International Journal of Intercultural Relations*. 2019, **69**, 120-130. Prieiga per: <https://doi.org/10.1016/j.ijintrel.2019.02.002>
30. HOMBRADOS-MENDIETA, I., M. MILLÁN-FRANCO, L. GÓMEZ-JACINTO, F. GONZALEZ-CASTRO, M. J. MARTOS-MÉNDEZ, ir A. GARCÍA-CID. Positive Influences of Social Support on Sense of Community, Life Satisfaction and the Health of Immigrants in Spain. *Frontiers in psychology*. 2019, **10**, 2555. Prieiga per: <https://doi.org/10.3389/fpsyg.2019.02555>

31. HOUSE, J. S. Social Support and Social Structure. *Sociological Forum*. 1987, **2**(1), 135–146. Prieiga per: <https://doi.org/10.1007/BF01107897>
32. HU, C., S. WANG, C. YANG ir T. WU. When mentors feel supported: Relationships with mentoring functions and protégés' perceived organizational support. *Journal of Organizational Behavior*. 2013, **35**(1), 22-37. Prieiga per: <https://doi.org/10.1002/job.1851>
33. YAKUSHKO, O. Xenophobia. *Counseling Psychologist - COUNS PSYCHOL*. 2008, **37**, 36-66. Prieiga per: <https://doi.org/10.1177/0011000008316034>
34. IRFAN, M., W. URIAWAN, O. T. KURAHMAN, M. A. RAMDHANI ir I. A. DAHLIA. Comparison of naive bayes and K-nearest neighbor methods to predict divorce issues. *IOP Conference Series: Materials Science and Engineering*. 2018, 434, 012047. Prieiga per: <https://doi.org/10.1088/1757-899x/434/1/012047>
35. JIANG, L., D. WANG, Z. CAI ir X. YAN. Survey of improving naive Bayes for classification. *Advanced Data Mining and Applications*. 2007, 134–145. Prieiga per: https://doi.org/10.1007/978-3-540-73871-8_14
36. JIANG, L., Z. CAI, D. WANG ir S. JIANG. Survey of improving K-nearest-neighbor for classification. *Fourth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD 2007)*. 2007. Prieiga per: <https://doi.org/10.1109/fskd.2007.552>
37. JUNG TINIŪ TAUTŪ ORGANIZACIJA. Toolkit on International Migration. 2012. Prieiga per: https://www.un.org/en/development/desa/population/migration/publications/others/docs/toolkit_DESA_June%202012.pdf
38. KIM, H. Enacted Social Support on Social Media and Subjective Well-Being. *International Journal of Communication*. 2014, **8**, 2340–2342.
39. KIM, Y. Y. Cross-Cultural Adaptation. Oxford Research Encyclopedia of Communication. 2017. Prieiga per: <https://doi.org/10.1093/acrefore/9780190228613.013.21>
40. KOELET, S. ir H. A. G. DE VALK. Social networks and feelings of social loneliness after migration: The case of European migrants with a native partner in Belgium. *Ethnicities*. 2016, **16**(4), 610-630. Prieiga per: <https://doi.org/10.1177/1468796816638398>
41. KUSUMAWATI, R., A. D'AROFAN ir P. A. PRAMANA. Comparison performance of naive Bayes classifier and support vector machine algorithm for Twitter's classification of Tokopedia Services. *Journal of Physics: Conference Series*. 2019, **1320**(1), 012016. Prieiga per: <https://doi.org/10.1088/1742-6596/1320/1/012016>
42. LANDTHALER, J., B. WALTL, D. HUTH, D. BRAUN, F. MATTHES, C. STOCKER ir T. GEIGER. Extending Thesauri Using Word Embeddings and the Intersection Method. *Conference: Second Workshop on Automated Semantic Analysis of Information in Legal Text (ASAIL 2017)*. 2017. Prieiga per: https://www.researchgate.net/publication/326588219_Extending_Thesauri_Using_Word_Embeddings_and_the_Intersection_Method
43. LOAYZA-RIVAS, J. ir J. FERNÁNDEZ-CASTRO. Perceived stress and well-being: The role of social support as a protective factor among Peruvian immigrants in Spain. *Ansiedad y Estrés*. 2020, **26**(2-3), 67-72. Prieiga per: <https://www.elsevier.es/es-revista-ansiedad-estres-242-articulo-perceived-stress-well-being-the-role-S1134793720300191>
44. MANCOSU, M. ir F. VEGETTI. What you can scrape and what is right to scrape: A proposal for a tool to collect public Facebook data. *Social Media + Society*. 2020, **6**(3), 205630512094070. Prieiga per: <https://doi.org/10.1177/2056305120940703>
45. MASSEY, D., J. ARANGO, G. HUGO, A. KOUAOUCCI, A. PELLEGRINO ir J. TAYLOR. Theories of International Migration: A Review and Appraisal. *Population and Development Review*. 1993, **19**(3), 431-466. Prieiga per: <https://doi.org/10.2307/2938462>
46. MIKOLOV, T., K. CHEN, G. CORRADO, G. ir J. DEAN. Efficient estimation of word representations in vector space. *Proceedings of Workshop at ICLR*. 2013. Prieiga per: <https://doi.org/10.48550/arXiv.1301.3781>

47. MITRA, A. ir Q. EVANSLUONG. Narratives of integration: Liminality in migrant acculturation through social media. *Technological Forecasting and Social Change*. 2019, **145**, 474-480. Prieiga per: <https://doi.org/10.1016/j.techfore.2019.01.011>
48. MORELLI, S. A., I. A. LEE, M. E. ARNN ir J. ZAKI. Emotional and instrumental support provision interact to predict well-being. *Emotion (Washington, D.C.)*. 2015, **15**(4), 484-493. Prieiga per: <https://doi.org/10.1037/emo0000084>
49. MUNOZ-LABOY, M., N. SEVERSON, A. PERRY, A. ir V. GUILAMO-RAMOS. Differential Impact of Types of Social Support in the Mental Health of Formerly Incarcerated Latino Men. *American Journal of Men's Health*. 2013, 226-239. Prieiga per: <https://doi.org/10.1177/1557988313508303>
50. NARDON, L., K. ATEN ir D. GULANOWSKI. Expatriate adjustment in the digital age: The co-creation of online social support resources through blogging. *International Journal of Intercultural Relations*. 2015, **47**, 41-55. Prieiga per: <https://doi.org/10.1016/j.ijintrel.2015.04.001>
51. NG, A. Y. ir M. I. JORDAN. On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *Neural Information Processing Systems*. 2022, **14**, 841-848.
52. OH, H. J., E. OZKAYA ir R. LAROSE. How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction. *Computers in Human Behavior*. 2014, **30**, 69-78. Prieiga per: <https://doi.org/10.1016/j.chb.2013.07.053>
53. POHL, S. ir M. GALLETTA. The role of supervisor emotional support on individual job satisfaction: A multilevel analysis. *Applied Nursing Research*. 2016, **33**, 61-66. Prieiga per: <https://doi.org/10.1016/j.apnr.2016.10.004>
54. PRATAMA, B. Y. ir R. SARNO. Personality classification based on Twitter text using naive Bayes, KNN and SVM. *2015 International Conference on Data and Software Engineering (ICoDSE)*. 2015. Prieiga per: <https://doi.org/10.1109/icodse.2015.7436992>
55. RISH, I. An Empirical Study of the Naive Bayes Classifier. *IJCAI 2001 Work Empir Methods Artif Intell*. 2001, **3**, 41-46. Prieiga per: <https://www.cc.gatech.edu/home/isbell/classes/reading/papers/Rish.pdf>
56. ROBERT, A.M. ir T. GILKINSON. Mental health and well-being of recent immigrants in Canada: Evidence from the Longitudinal Survey of Immigrants to Canada (LSIC). *Ottawa, ON: Citizenship and Immigration Canada*. 2012, 27. Prieiga per: <https://www.canada.ca/en/immigration-refugees-citizenship/corporate/reports-statistics/research/mental-health-well-being-recent-immigrants-canada-evidence-longitudinal-survey-immigrants-canada-lsic.html>
57. SANDOZ, L. ir F. SANTI. Who Receives More Help? The Role of Employer Support in Migration Processes. Iš STEINER I. ir P. WANNER, *Migrants and Expats: The Swiss Migration and Mobility Nexus*. IMISCOE Research Series, Springer, 2019, 57-81. Prieiga per: https://doi.org/10.1007/978-3-030-05671-1_3
58. SARICA, S. ir J. LUO. Stopwords in technical language processing. *PLOS ONE*. 2021, **16**(8). <https://doi.org/10.1371/journal.pone.0254937>
59. SEMMER, N. K., ELFERING, A., JACOBSHAGEN, N., PERROT, T., BEEHR, T. A. IR BOOS, N.. The emotional meaning of instrumental social support. *International Journal of Stress Management*. 2008, **15**(3), 235–251. Prieiga per: <https://doi.org/10.1037/1072-5245.15.3.235>
60. SHAW, L. H. ir L. M. GANT. In Defense of the Internet: The Relationship between Internet Communication and Depression, Loneliness, Self-Esteem, and Perceived Social Support. *CyberPsychology & Behavior*. 2004, **5**(2), 157-171. Prieiga per: <https://doi.org/10.1089/109493102753770552>

61. SIMOSI, M. The moderating role of self-efficacy in the organizational culture–training transfer relationship. *International Journal of Training and Development*. 2012, **16**(2), 92–106. Prieiga per: <https://doi.org/10.1111/j.1468-2419.2011.00396.x>
62. SLOAN, M. M., R. J. EVENSON NEWHOUSE ir A. B. THOMPSON. Counting on Coworkers: Race, Social Support, and Emotional Experiences on the Job. *Social Psychology Quarterly*. 2013, **76**(4), 343–372. Prieiga per: <https://doi.org/10.1177/0190272513504937>
63. SUNDARI, A. ir R. PEARSON. Types of migrants. *Striking women*. 2013. Prieiga per: <https://www.striking-women.org/main-module-page/types-migrants>
64. TARPTAUTINĖ MIGRACIJOS ORGANIZACIJA. Key migration terms. Prieiga per: <https://www.iom.int/key-migration-terms>
65. TARPTAUTINĖ MIGRACIJOS ORGANIZACIJA. World migration report 2020. 2019, ISSN 1561-5502. Prieiga per: https://publications.iom.int/system/files/pdf/wmr_2020.pdf
66. TEN KATE, R. L. F., B. BILECEN ir N. STEVERINK. Closer Look at Loneliness: Why Do First-Generation Migrants Feel More Lonely Than Their Native Dutch Counterparts? *The Gerontologist*. 2020, **60**(2), 291–301. Prieiga per: <https://doi.org/10.1093/geront/gnz192>
67. THEEMLING, R. The Use of Online Social Media and Expatriate Adjustment and Integration in the Netherlands. 2013. Prieiga per: <https://thesis.eur.nl/pub/15589/Theemling-R.pdf>
68. TREPTE, S., T. DIENLIN ir L. REINECKE. Influence of Social Support Received in Online and Offline Contexts on Satisfaction With Social Support and Satisfaction With Life: A Longitudinal Study. *Media Psychology*. 2015, **18**(1), 74–105. Prieiga per: <https://doi.org/10.1080/15213269.2013.838904>
69. TRSTENJAK, B., S. MIKAC ir D. DONKO. KNN with TF-IDF based framework for text categorization. *Procedia Engineering*. 2014, **69**, 1356–1364. Prieiga per: <https://doi.org/10.1016/j.proeng.2014.03.129>
70. VOGLER, R. The TF-IDF-statistic for keyword extraction: R-bloggers. 2014. Prieiga per: <https://www.r-bloggers.com/2014/02/the-tf-idf-statistic-for-keyword-extraction/>
71. WEE CHAN, A. ir K. A. PERVAIZ. Exploring the effects of workplace support on work-life experience: a study of Malaysia. *Human Resource Development International*. 2015, **18**(4), 346–365. Prieiga per: <https://doi.org/10.1080/13678868.2015.1019816>
72. WETTSCHERECK, D., D. W. AHA, ir T. MOHRI. A Review and Empirical Evaluation of Feature Weighting Methods for a Class of Lazy Learning Algorithms. *Artificial Intelligence Review*. 1997, **11**(1/5), 273–314. Prieiga per: <https://doi.org/10.1023/a:1006593614256>
73. ZHANG, W. ir F. GAO. An improvement to naive Bayes for text classification. *Procedia Engineering*. 2011, **15**, 2160–2164. Prieiga per: <https://doi.org/10.1016/j.proeng.2011.08.404>
74. ZHANG, Z. Introduction to machine learning: k-nearest neighbors. *Annals of translational medicine*. 2016, **4**(11), 218. Prieiga per: <https://doi.org/10.21037/atm.2016.03.37>
75. ZHU, W., N. ZENG ir N. WANG. Sensitivity, Specificity, Accuracy, Associated Confidence Interval and ROC Analysis with Practical SAS Implementations. *NorthEast SAS users group, health care and life sciences*. 2010. Prieiga per: <http://www.nesug.org/Proceedings/nesug10/hl/hl07.pdf>