



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas

**Klientų segmentavimas ir lojalumo įvertinimas: poligrafijos
įmonės atvejis**

Baigiamasis magistro projektas

Aistė Kazakevičiūtė
Projekto autorius

Doc. dr. Evaldas Vaičiukynas
Vadovas
Doc. dr. Aistė Dovalienė
Vadovas

Kaunas, 2018



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas

Klientų segmentavimas ir lojalumo įvertinimas: poligrafijos įmonės atvejis

Baigiamasis magistro projektas
Didžiųjų verslo duomenų analitika (621G12002)

Aistė Kazakevičiūtė
Projekto autorius

Doc. dr. Evaldas Vaičiukynas
Vadovas
Doc. dr. Aistė Dovalienė
Vadovas

Doc. dr. Vytautas Janilionis
Recenzentas
Doc. dr. Agnė Gadeikienė
Recenzentas

Kaunas, 2018



Kauno technologijos universitetas
Matematikos ir gamtos mokslų fakultetas
Aistė Kazakevičiūtė

Klientų segmentavimas ir lojalumo įvertinimas: poligrafijos įmonės atvejis

Akademinio sąžiningumo deklaracija

Patvirtinu, kad mano, Aistės Kazakevičiūtės, baigiamasis projektas tema „Klientų segmentavimas ir lojalumo įvertinimas: poligrafijos įmonės atvejis“ yra parašytas visiškai savarankiškai ir visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjusi.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

(vardą ir pavardę įrašyti ranka)

(parašas)

Turinys

Įvadas	9
1. Vartotojų lojalumas ir jų segmentavimas	11
1.1. Vartotojų segmentavimas, jo tikslai, kriterijai	11
1.2. Vartotojo lojalumas ir jo skatinimas	15
1.3. Metodai klientų segmentavimui ir lojalumo įvertinimui	20
2. Segmentavimui ir lojalumo įvertinimui naudojami metodai	27
2.1. ABC analizė	27
2.2. RFM analizė	28
2.3. Dviejų žingsnių klasterizavimas	33
2.4. Naudojami detektoriai	34
2.5. Detekcijos gerumo įvertinimas	36
3. Poligrafijos įmonės klientų segmentavimas ir lojalumo patikra	42
3.1. Įmonės ir jos klientų analizė	42
3.2. Klientų segmentavimas	46
3.3. Klientų lojalumo detekcija	56
Išvados	62
Literatūros sąrašas	63
Priedai	65

Paveikslų sąrašas

1 pav. Segmentavimo proceso koncepcinis modelis (Hwang et al., 2004)	16
2 pav. Segmentavimo pavyzdys pagal LTV (Hwang et al., 2004)	17
3 pav. Lojalumą įtakojantys veiksniai (Janita & Miranda, 2013)	19
4 pav. ABC analizė (sudaryta autorės remiantis Chu et al., 2008; Sutevski, 2018).....	27
5 pav. RFM metodo segmentavimas (sudaryta autorės, remiantis Gajdhar, 2017).....	30
6 pav. Trijų blokų kryžminės validacijos pavyzdys (Mudry & Tjellström, 2011).....	37
7 pav. DET (kairėje) ir ROC (dešinėje) kreivės (Sáenz-Lechón et al., 2006).....	38
8 pav. Preciziškumo – prisiminimo kreivės pavyzdys (Tolay, 2018)	39
9 pav. Paslaugų pasiskirstymas įmonėje pagal populiarumą.....	42
10 pav. Paslaugų pasiskirstymas pagal gaunamą pelną.....	43
11 pav. Pagrindinės klientų grupės.....	44
12 pav. Klientų pasiskirstymas Europoje (spalvos ryškumas parodo klientų tankį šalyje).....	45
13 pav. Klientų pasiskirstymas tarp užsienio šalių.....	45
14 pav. Klientų segmentavimas ABC metodu.....	46
15 pav. RFMT visų duomenų klasterizavimo kokybė naudojant dviejų žingsnių metodą (silueto plotis = 0,6).....	48
16 pav. Visų klientų segmentavimas naudojant dviejų žingsnių klasterizavimo metodą (silueto plotis = 0,6).....	49
17 pav. Auksinių klientų segmentavimas (silueto plotis = 0,5)	51
18 pav. Sidabrinių klientų segmentavimas (silueto plotis = 0,5)	52
19 pav. Bronzinių klientų segmentavimas (silueto plotis = 0,5)	53
20 pav. Visų įmonės klientų segmentų vizualizacija pagal klientų kiekį segmentuose	54
21 pav. Klientų segmentavimo pagal paslaugas modelio kokybė (silueto plotis = 0,7)	55
22 pav. Klientų segmentai pagal paslaugų užsakymo dažnumą (silueto plotis = 0,7)	55
23 pav. Lojalumo atskyrimo riba.....	57
24 pav. ROC tikslumo kreivės.....	58
25 pav. DET klaidų dažnio kreivės	59
26 pav. Apibendrinti rezultatai	60
27 pav. Suprognuoto lojalumo iš atsitiktinio miško detektoriaus pasiskirstymas	60
28 pav. Sumaišymų matrica	61

Lentelių sąrašas

1 lentelė. LTV įvertinimo rodikliai (sudaryta autorės, remiantis Hwang et al., 2004).....	16
2 lentelė. Dažniausiai naudojami duomenų tyrybos modeliavimo metodai klientų analizei (Tsiptsis & Chorianopoulos, 2010)	25
3 lentelė. Klasterizavimo metodų palyginimas (Tsiptsis & Chorianopoulos, 2010).....	25
4 lentelė. RFM modelių patobulinimas (sudaryta autorės)	29
5 lentelė. RFM kriterijų reikšmės (Gajdhar, 2017)	30
6 lentelė. RFM analizės deriniai (Gajdhar, 2017)	31
7 lentelė. Klientų segmentavimo pavyzdys (Birant, 2016)	32
8 lentelė. Sumaišymo matricos klasės (sudaryta autorės, remiantis Fawcett, 2004).....	40
9 lentelė. Visų užsakymų vertės aprašomoji statistika klientų segmentams pagal ABC analizę	47
10 lentelė. Visų įmonės klientų RFMT kintamųjų aprašomoji statistika	47
11 lentelė. Klientų segmentai atsižvelgiant į RFMT reikšmes	48
12 lentelė. Auksinės klientų grupės RFMT kintamųjų aprašomoji statistika.....	50
13 lentelė. Sidabrinės klientų grupės RFMT kintamųjų aprašomoji statistika.....	51
14 lentelė. Bronzinės klientų grupės RFMT kintamųjų aprašomoji statistika	53
15 lentelė. Visų ABC įmonės klientų segmentų išskaidymas	54

Kazakevičiūtė, Aistė. Klientų segmentavimas ir lojalumo įvertinimas: poligrafijos įmonės atvejis. Magistro baigiamasis projektas / vadovai: doc. dr. Evaldas Vaičiukynas; doc. dr. Aistė Dovalienė; Kauno technologijos universitetas, Matematikos ir gamtos mokslų fakultetas.

Studijų kryptis ir sritis (studijų krypčių grupė): Taikomoji matematika (A02), Matematikos mokslai (A).

Reikšminiai žodžiai: klientų lojalumas, klientų segmentavimas, ABC, RFM, dviejų žingsnių klasterizavimas, lojalumo detekcija.

Kaunas, 2018. 71 p.

Santrauka

Pagrindinis baigiamojo darbo tikslas – atlikti spaudos įmonės klientų segmentavimą ir lojalumo detekciją pelningiausia spaudos įmonės klientų segmentui.

Darbas susideda iš trijų pagrindinių dalių: literatūros apžvalgos, darbe naudojamų tyrimo metodų apžvalgos ir tiriamosios dalies, kurioje pateikti tyrimo rezultatai.

Literatūros apžvalgoje aptariama klientų segmentavimo ir lojalumo nauda teikiama įmonei. Aptariami populiariausi klientų segmentų tipai ir jų išskirtiniai bruožai. Taip pat išskiriama lojalių klientų grupė, jos charakteristika, privalumai ir lojalumo skatinimo galimybės.

Antroje dalyje aptariami šiame darbe naudojami klientų segmentavimo ir lojalumo detekcijos metodai: ABC analizė, RFM analizė, dviejų žingsnių klasterizavimo metodas ir detekcijai skirti mašininio mokymo algoritmai. Apibrėžiami jų naudojimo tikslai, veikimo principai ir skaičiavimo pavyzdžiai rasti literatūroje.

Trečioje dalyje pateikiama įmonės klientų apžvalga ir gauti rezultatai po metodų, aptartų antroje dalyje, panaudojimo. Atliekami keli klientų segmentavimo į įvairius klasterius variantai, panaudojant RFMT kintamuosius. Toliau atliekama klientų lojalumo patikra vertingiausiems klientams iš A ir B segmento, kurie buvo nustatyti atlikus ABC analizę. Naudojant mašininio mokymo algoritmus kartu su kryžmine patikra, įvertinamas klientų lojalumas sprendžiant detekcijos uždavinį. Po atliktos klientų analizės, įmonei pateikiamos rekomendacijos.

Kazakevičiūtė, Aistė. Customer Segmentation and Churn Detection: A Case of Printing House. Master's Final Degree Project / supervisors: assoc. prof. Evaldas Vaičiukynas; assoc. prof. Aistė Dovalienė; Faculty of School of Mathematics and Natural Sciences, Kaunas University of Technology.

Study field and area (study field group): Applied mathematics (A02), Mathematics (A).

Keywords: customer loyalty, customer segmentation, ABC, RFM, Two Step cluster analysis, churn detection.

Kaunas, 2018. 71 pages.

Summary

The main aim of this thesis is to conduct a segmentation of clients and churn detection for a most profitable segment of a printing company.

The thesis consists of three main parts: a literature review, investigation of methods used at work and a research part in which the results of a research are presented.

A discussion about customer segmentation and loyalty is presented in a literature review part as well as a short review of the most popular customer types and their most distinctive features. In this part there is also an analysis of most loyal customers separating them by distinction of their characteristics, benefits and the opportunities of encouragement loyalty.

In the second part of this thesis customer segmentation and loyalty detection methods, such as ABC analysis, RMF analysis, Two Step cluster analysis and machine learning algorithms for detection were carried out using the algorithms of classification. Also, the aims of using above mentioned tools, their operating principles and the examples of calculations which were found in literature are defined in the second part.

In the third part, the overview of clients and the outcome of implementing the methods mentioned in second part is presented. Several customer segmentation cases are carried out dividing them in to various clusters using variables of RFMT. During next step customer loyalty study is conducted to the most valuable clients from segments A and B, which were determined after carrying on ABC analysis. Using machine learning algorithms together with cross validation the loyalty of clients is evaluated by solving detection tasks. After conducting analysis of clients, the recommendations are suggested to the company.

Įvadas

Darbo aktualumas. Verslo pasauliui augant, kartu didėja ir konkurencija tarp įmonių. Klientai įgauna vis didesnę pasirinkimo laisvę, todėl tampa vis paprasčiau pakeisti nepatikusią prekę ar paslaugų tiekėją, jei pirkėjo poreikiai nėra patenkinami. Prarasti klientai įmonėms sudaro papildomus nuostolius, ypač jei prarandamas lojalus klientas, nes sumažėja įplaukos ir tenka skirti papildomai lėšų rinkodarai. Organizacijai nepalankų klientų apsisprendimą gali paskatinti įvairios priežastys, tokios kaip: prasta produktų ar paslaugų kokybė, aukšta kaina, netinkamas aptarnavimas ar geresnės alternatyvos. Dėl anksčiau minėtų priežasčių, įmonėms labai svarbu laiku gerai pažinti rinką ir išvelgti, kada klientai ruošiasi atsisakyti teikiamų paslaugų. Siekiant įmonėms sėkmingai vykdyti savo veiklą, vis daugiau dėmesio skiriama rinkos pažinimui, jos vertės gerinimui ir išlaikymui.

Tinkamoms išvalgoms padaryti labai populiarus yra klientų segmentavimas, kuris svarbus šiuolaikinėms įmonėms dėl geresnio klientų poreikių supratimo, kurio pagalba įmonės taip pat gali sustiprinti klientų lojalumą. Daugelis neskirsto klientų į grupes ir laikosi nuomonės, kad visi klientai yra svarbūs, tačiau vieni jų yra svarbesni už kitus ir tą įmanoma identifikuoti, o tai padėtų organizacijai orientuotis tik į vertę įmonei atnešančius klientus ir išvengti nereikalingų išlaidų. Išskirstant klientus į grupes pagal jų pirkimo ar asmeninius požymius, galima išvelgti, kodėl jie perka ir ką reikia padaryti, kad nenutruktų jų santykiai su kompanija. Remiantis segmentų charakteristika galima sukurti keletą skirtingų strategijų, kaip išlaikyti ar pritraukti įmonės klientus, tačiau tokie procesai yra sudėtingi ir reikalauja daug laiko bei dėmesio. Netinkamai parinktos segmentavimo strategijos gali būti nuostolingos įmonei, nes gali skatinti vartotojų nepasitenkinimą, o tuo pačiu sumažinti pardavimus, todėl svarbu pasirinktas strategijas nuolatos analizuoti ir tobulinti.

Spaudos įmonės atveju klientų segmentavimo pavyzdžių literatūroje nėra pateikiama. Klientai gali būti ne tik įmonės, bet ir asmenys, kurie yra iš skirtingų šalių ir perka dažniausiai vienos rūšies gaminius dideliais tiražais. Šioje srityje klientai yra labai reiklūs, jų užsakymai retai būna standartiniai ar pasikartojantys ir jų poreikių įgyvendinimas reikalauja daug pastangų. Nestandartinės produkcijos gamybos įmonės klientų segmentavimas padėtų įmonėms atkreipti dėmesį į jau pamirštus, tačiau vertingus klientus, taip pat apžvelgti lojalius klientus, kuriuos reikėtų puoselėti. Kiekvienam segmentui pritaikius išskirtinę strategiją, būtų galima sustiprinti ryšius tarp įmonės ir kliento ir lengviau pasiekti pardavimo tikslus. Taip pat atlikus segmentavimą, pastebimi vertingų klientų išskirtiniai bruožai, kuriuos žinant būtų galima pastebėti kliento vertę jau ankstesnėse, neseniai pradėjusio pirkti kliento stadijose ir jam laiku pritaikyti tinkamą rinkodaros strategiją.

Darbo objektas. Poligrafijos įmonės klientai, kurie užsako standartines bei nestandartines spaudos paslaugas.

Darbo tikslas. Atlikti spaudos įmonės klientų segmentavimą ir sumodeliuoti pelningiausių segmentų klientų lojalumą.

Darbo uždaviniai:

1. apžvelgti literatūroje apibrėžiamą klientų segmentų ir lojalumo sampratą bei teikiamą naudą;
2. aptarti šiame darbe taikytinus klientų segmentavimo ir lojalumo detekcijos metodus;
3. atlikti spaustuvės klientų žvalgomąją analizę ir išskirti klientų segmentus, remiantis aptartais metodais;
4. vertingiausiems klientams atlikti lojalumo detekciją ir įvertinti detekcijos gerumą;
5. remiantis gautais rezultatais, pateikti rekomendacijas įmonei.

1. Vartotojų lojalumas ir jų segmentavimas

Ryšių su klientais valdymo tikslas – gauti, išlaikyti ir padidinti klientų vertę. Svarbiausia yra veiksmingai suskirstyti klientus į skirtingus segmentus, o tada pagal svarbumą įmonė atitinkamai paskirsto išteklius skirtingoms klientų grupėms, taupydama savo lėšas, tobulindama įmonės paslaugų lygį, didindama klientų lojalumą ir tuo pačiu padidindama įmonės pelną bei klientų susidomėjimą (Li & Li, 2011).

1.1. Vartotojų segmentavimas, jo tikslai, kriterijai

Klientų segmentavimas (taip pat dar vadinamas rinkos segmentavimu) – tai klientų paskirstymas į grupes, kurios išsiskiria panašiomis charakteristikomis. Klientų segmentavimas padeda pasirinkti atitinkamą komunikaciją su klientais, taip pat padeda identifikuoti pelningiausias įmonės klientus ir pasiūlyti jiems atitinkamus produktus bei paslaugas, kad būtų patenkinti jų poreikiai. Segmentavimas yra susijęs su pirkimų patirtimi, kur pastebimas ir klientų lojalumas prekės ženklui („Magenta“, 2013). Įmonės klientus reikia segmentuoti periodiškai, nes grupės gali kisti. Stebint vartotojų elgesio skirtumus, renkantis paslaugą ar prekę, galima atrasti naujų grupių. Segmentavimas prasmingas, kada segmentai yra išmatuojami (dydžiu, perkamąja galia, laiku ir pan.), tinkamo dydžio, užtikrinantys atitinkamas pajamas, pasiekiami ir išsiskiriantys reakcija į organizacijų naudojamas rinkodaros priemones (Urbonavičius & Ivanauskas, 2006). Viena rinkodaros strategija negali tikti visiems klientams vienodai, o kai kuriais atvejais, pasirinkus netinkamą bendravimo būdą su atitinkamu klientu, gali pareikalauti tik didelių išlaidų.

Segmentavimas – tai klientų bazės padalijimo į skirtingas ir viduje homogeniškas grupes, siekiant parengti diferencijuotas rinkodaros strategijas, atitinkančias jų charakteristikas. Yra daug segmentavimo tipų, grindžiamų tam tikrais kriterijais ar atributais (Tsipstis & Chorianopoulos, 2010).

Duomenų tyryba padeda identifikuoti naujus klientus, diferencijuoti esamus klientus ir rasti geriausią būdą užmegzti pokalbį su klientu. Duomenų tyrybos būdai, tokie kaip segmentavimas ir klasifikavimas, padeda siekti šių tikslų. Įmonės gali labiau pažinti savo klientus dėl segmentavimo ir aptarnauti juos geriau, pateikiant jiems tinkamus paslaugų ir produktų tipus. Segmentavimas padeda suskirstyti klientus į įvairias grupes pagal jų poreikius. Tokie klientų segmentai gali būti geras pagrindas tikslinei rinkodaros strategijai, kur vienas segmentas skirtas vienam rinkodaros produktų tipui (Kadiyala & Srivastava, 2011).

Taip pat literatūroje galima rasti, kad segmentavimo procesas naudojamas norint atskirti klientus nuo ne klientų, kur žodis „klientas“ apima pirkėjus, mokėtojus, lojalius klientus ir daug kitų, taip pat suprasti jų sudėtį ir ypatybes, atsakant į klausimus: kas jie tokie? kaip jie atrodo? kokios jų

charakteristikos? kur jie gyvena? Ši analizė padeda priimti daug sprendimų, kaip sprendimų taikymą, nustatant veiksmingas ir ekonomiškai efektyvias rinkodaros strategijas, įvertinant ir rinkos konkurenciją (Levin & Zahavi, 2001).

Kituose šaltiniuose segmentavimas apibrėžiamas kaip procesas, padalijantis klientus į skirtingas, prasmingas ir vienaarūšes grupes, kurios pagrįstos įvairiais požymiais bei savybėmis. Ši technika naudojama kaip rinkodaros įrankis ir tai leidžia organizacijoms labiau suprasti savo klientus ir kurti diferencijuotas strategijas (Ziafat & Shakeri, 2014).

Įprastai organizacijos, nepriklausomai nuo pramonės šakos, kurioje jie vykdo savo veiklą, yra linkusios naudoti rinkos segmentavimo sistemas, kurios remiasi demografiniais duomenimis, tokiais kaip lytis, amžius, pajamos bei kita vertinga informacija, nes juos paprasta išmatuoti ir palyginti su pastebėtais vartotojų elgesio pokyčiais. Per kelis pastaruosius dešimtmečius organizacijos pradėjo remtis šia informacija pateikdamos rinkai naujus produktus ir paslaugas (Ziafat & Shakeri, 2014).

Klientų segmentavimas yra populiarus duomenų tyrybos taikymas. Segmentavimo procesas prasideda nuo verslo tikslų apibrėžimo ir baigiamas pristatant diferencijuotas rinkodaros strategijas išskirtiems segmentams. Yra didelis kiekis skirtingų segmentų tipų, pagrįstų konkrečiais kriterijais ar atributais, taigi naudojamas segmentavimo tipas priklauso nuo įmonės išsikeltų tikslų. Pavyzdžiui, klientai gali būti segmentuojami pagal jų vertę, socialinę ir demografinę padėtį, jų elgesį, poreikius, požiūrį ir lojalumą. Skirtingi kriterijai ir segmentavimo metodai yra tinkami skirtingoms situacijoms ir verslo tikslams, todėl prieš atliekant segmentavimą reikia ne tik nusistatyti siekiamą tikslą, bet ir apžvelgti turimus duomenis modeliui sudaryti. Taigi segmentavimo tikslai gali būti (Ziafat & Shakeri, 2014):

- geriau suprasti klientus, siekiant identifikuoti naujas galimybes rinkoje;
- sukurti naujus produktus / paslaugas, siekiant pritaikyti juos atskiriems segmentams, o ne masinei rinkai;
- sukurti naują produktą skirtingiems klientų segmentams pagal nustatytus norus ir poreikius;
- pasiūlyti tinkamas dovanas ar paskatinimus;
- pasirinkti tinkamą reklamos ir komunikacijos pranešimą bei kanalą;
- pasirinkti tinkamą pardavimų / paslaugų kanalą;
- nustatyti prekinio ženklo įvaizdį ir pagrindinio produkto pranašumus, kurie turi būti pateikti atsižvelgiant į konkrečias kiekvieno segmento ypatybes;
- diferencijuoti klientų aptarnavimą pagal kiekvieno segmento svarbą;
- veiksmingiau paskirstyti išteklius pagal kiekvieno segmento potencialią grąžą;

- paskirstyti rinkodaros iniciatyvas pagal prioritetus, kurie orientuoti į klientų išlaikymą ir plėtrą, pagal kiekvieno segmento svarbą ir vertę.

Kadangi pagal išsikeltus tikslus reikia rinktis segmentavimo tipą, reikia žinoti, pagal ką galima diferencijuoti vartotojus. Išskiriami du apibendrinti segmentavimo variantai – pagal išorinius vartotojų bruožus, kaip geografines, demografines ir psichologines savybes, bei pagal vartotojų elgseną, kur atsižvelgiama į kliento žinias apie produktą, požiūrį ir nuostatas, o tai galima sužinoti pagal įvairius kriterijus (Urbonavičius & Ivanauskas, 2006). Literatūroje galima rasti ir daugiau išskirstytų kriterijų (Ziafat & Shakeri, 2014):

- verte pagrįstas segmentavimas – klientai sugrupuojami pagal vertę, tai labai aktualu, kada norima nustatyti vertingiausias organizacijos klientus ir stebėti vertės pokyčius laikui bėgant. Šis segmentavimas taip pat naudojamas, siekiant atskirti paslaugų teikimo strategijas ir optimizuoti rinkodaros išteklių paskirstymą;
- segmentavimas pagal elgseną – šioje grupėje klientai išskirstomi pagal jų elgesį ir vartojimo įpročius. Šis tipas paprastai naudojamas kuriant produktų pateikimo strategijas, naujų produktų vystymui ir lojalumo programų kūrimui;
- orientacija pagrįstas segmentavimas – šis segmentavimas skirtas suskirstyti pagal polinkio rodiklius į grupes, tokius kaip pasipriešinimo, kryžminio pardavimo balai ir kiti, kurie vertinami pagal atitinkamus klasifikavimo modelius. Pavyzdžiui, pagal šį segmentavimą klientai gali būti suskirstyti į „blogas“, „geras“ ir „labai geras“ grupes;
- lojalumo segmentavimas – šis grupavimas skirtas identifikuoti klientų lojalumą, ar jis yra lojalus ar migruojantis. Tokiu atveju sulaikymo veiksmai gali būti nukreipti į aukštos vertės klientus su neištikimumo rodikliais, taip pat į perspektyviai lojalius klientus;
- socialinis-demografinis ir gyvenimo būdo segmentavimas – pagal šį tipą atskleidžiamos skirtingos klientų grupės, pagrįstos socialine-demografinė ir gyvenimo būdo informacija, tokia kaip lytis, rasė, amžius, socialinė padėtis, išsilavinimas, profesija, pajamos ar šeimyninė padėtis. Šis segmentavimas tinkamas skatinti tam tikram gyvenimo būdui skirtą produktų rinkodarą;
- reikalavimų / ir požiūrio segmentavimas – segmentai nustatomi pagal klientų poreikius, nuostatas ir suvokimą, susijusiais su įmonės produktais ir paslaugomis. Tai gali būti naudojama naujo produkto vystymui, prekinio ženklo įvaizdžio ir pagrindinių produkto savybių sudarymui.

Weinsteinas trumpai išskiria tik penkis segmentavimo kriterijus (Levin & Zahavi, 2001):

1. geografinis – tai rinkos suskirstymas pagal geografines aplinkybes;

2. socialinis-ekonominis – tai segmentavimas grindžiamas veiksniais, atspindinčiais klientų socialinį ir ekonominį statusą;
3. psichologinis – tai rinkos diferencijavimas, atsižvelgiant į klientų poreikius, asmenybės bruožus ir gyvenimo būdą;
4. produkto vartojamumas – rinkos suskaidymas į įvairias vartotojų grupes pagal vartojimo lygį;
5. privalumai – rinkos suskaidymas grindžiamas nauda, gauta iš produkto ar paslaugos, tokia kaip kaina, aptarnavimas, ypatingi bruožai ir veiksniai.

Taigi dauguma autorių pabrėžia segmentavimą pagal elgseną, lojalumą ir socialinę-demografinę padėtį. Taip pat yra nepaminėtų tokių segmentavimo galimybių, pavyzdžiui, pagal paskirstymą, kainą, laiką ir t. t. (Thomas, 2016) Tačiau, kaip jau buvo minėta, skirtingiems tikslams pasiekti naudojama skirtingi diferencijavimo kriterijai. Mažmeninėje prekyboje arba verslas – klientui (angl. *business to customer*) segmente (toliau – B2C), pagrindinis tikslas yra – kuo geriau tenkinti pirkėjo poreikius, o tam pasiekti reikia nustatyti pirkėjų poreikius, taigi tam dažniausiai taikomas tradicinis demografinis, geografinis ir psichologinis segmentavimas. Tačiau pastebėta, kad toks segmentavimas tinka nedidelės ar siauriau specializuotos mažmeninės prekybos įmonės klientams. Kitais atvejais, kada mažmeninės prekybos įmonės yra didelės ar parduotuvių tinklai, teikiantys paslaugas daug skirtingo amžiaus, pajamų ir gyvenamosios vietos klientams, reikia detalesnio segmentavimo, pavyzdžiui pagal pirkėjų elgsenos bruožus (Urbonavičius & Ivanauskas, 2006).

Verslas – verslui (angl. *business to business*) rinkoje (toliau – B2B) dažniau susiduriama su tuo, kad kiekvienas klientas yra unikalus, todėl ne visi kriterijai nurodyti literatūroje yra tinkami didmeninės klientų rinkos segmentavimui. Pirmiausia reikia atskirti, kuo skiriasi B2C ir B2B segmentavimas (Willan, 2014):

- B2B įprastai yra mažiau segmentų, nes rinkoje yra daug mažiau pirkėjų ir pirkėjai paprastai turi panašių reikalavimų, o tai reiškia, kad yra mažiau įvairovės, todėl paprastai būna keturi ar penki segmentai;
- B2B rinkoms įprastai galima taikyti 80/20 taisyklę – net 80% pajamų sukuria tik 20% klientų. Vadinasi, mažas skaičius vertingų klientų kartais turėtų būti priskiriamas atskiram segmentui;
- nors B2B segmentavimas iš pirmo žvilgsnio atrodo nesudėtingas, tačiau jis turi savų komplikuočių dalių, pavyzdžiui, kas vis dėl to yra priskiriamas segmentui: individualus pirkėjas ar įmonė, kurioje jis dirba;
- segmentai dažniausiai grindžiami racionaliais aspektais, tokiais kaip pirkimo aspektai, o ne emociniais, kaip saviraiška ir kiti elgesio bruožai, kas įprasta B2C atveju;

- B2B rinka kinta lėčiau nei B2C, todėl segmentavimas ilgesnį laiką išlieka tikslus, priešingai nei mažmeninėje prekyboje.

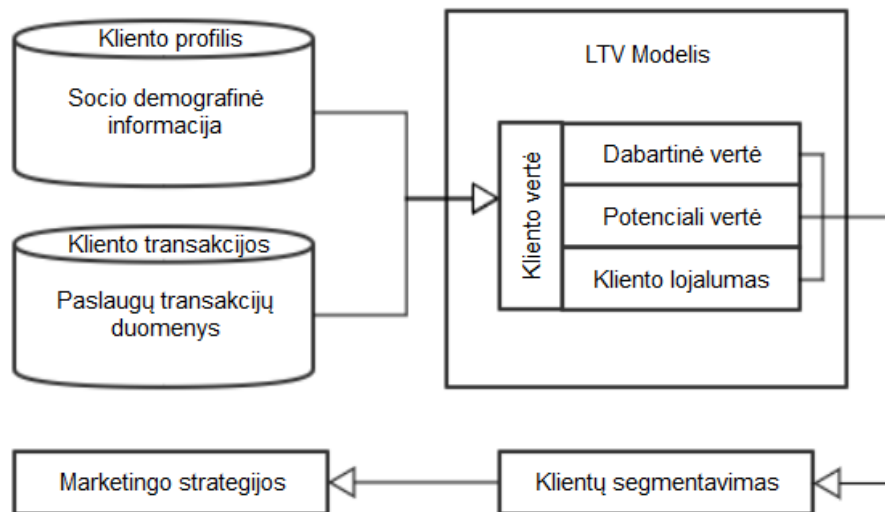
Taigi didmeninės rinkos klientai suskirstomi į grupes pagal firmografinius kintamuosius (tai yra specifiniai kompanijų kintamieji) (Willan, 2014):

1. įmonės dydis – apyvartos arba darbuotojų skaičiaus požiūriu;
2. įmonės vieta – šalis, regionas;
3. įmonės veikla arba pramonės sektorius.

Šiame poskyryje apžvelgti skirtingų atvejų pagrindiniai segmentavimo kriterijai, kuriais remiantis dauguma tyrėjų skirsto klientus į atitinkamas grupes, kurie yra skirtingos vertės ir bendrovė gali gauti nevienodą naudą iš jų. Tačiau vienas dažniausiai naudojamų segmentavimo kriterijų yra grupavimas pagal lojalumą, kuriuo metu išskiriami labai vertingi įmonės pirkėjai ne tik pagal išleidžiamą pinigų sumą.

1.2. Vartotojo lojalumas ir jo skatinimas

Praeitame poskyryje apie segmentavimą galima pastebėti, kad dauguma autorių siūlo išskirstyti vartotojus pagal jų lojalumą – tai svarbu įmonėms, nes apibrėžia kliento vertę ir iš to galima daryti išvadas, kokį pasirinkti tolimesnį bendravimą su klientu. Vartotojo vertę galima vertinti iš trijų taškų: dabartinė vertė, potenciali vertė ir klientų lojalumas (Hwang, Jung, & Suh, 2004). Potenciali vertė vaizduoja pardavimo galimybę, o lojalumas reiškia klientų išlaikymą. Didėjant įmonių konkurencingumui, vienu svarbiausiu įmonių tikslų tampa klientų lojalumas ir jo stiprinimas, dėl įmonei teikiamos naudos. Atlikus klientų segmentavimą, galima lengviau įvertinti klientų lojalumą įmonei ir parinkti apibrėžtai grupei klientų tinkamiausią strategiją. Hwangas su kitais autoriais sudarė koncepcinį modelį, kuriame atvaizduojama seka, kaip iš turimos informacijos apie klientus, įvertinama jų gyvavimo trukmės vertė (*angl. LTV – Lifetime value*) ir sudaroma rinkodaros strategija (žiūrėti 1 pav.) (Hwang et al., 2004).



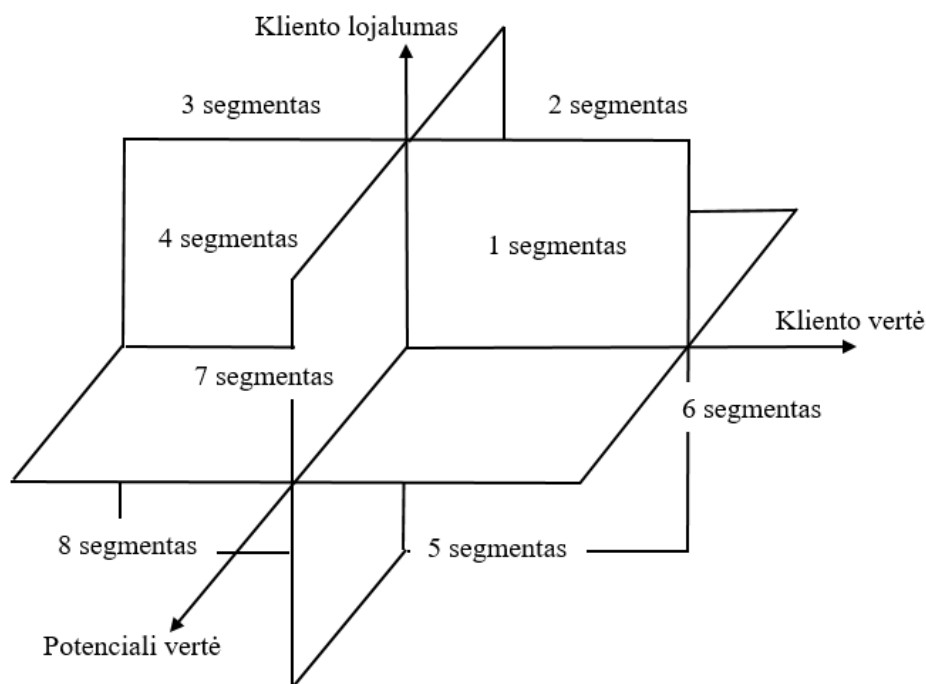
1 pav. Segmentavimo proceso koncepcinis modelis (Hwang et al., 2004)

Šiame modelyje pavaizduota, kaip informacija apie klientus ir jų užsakymus leidžia nustatyti kliento vertę, kuri susideda iš dabartinės, potencialios vertės ir kliento lojalumo (1 lent.). Visi išvardinti vertės rodikliai dar vadinami kliento gyvavimo trukmės (toliau – LTV) modeliu, kurio dėka galima atlikti segmentavimą, o remiantis gautais segmentais sudaryti rinkodaros strategijas.

1 lentelė. LTV įvertinimo rodikliai (sudaryta autorės, remiantis Hwang et al., 2004)

Kliento vertės rodikliai	Apibrėžimas
Dabartinė vertė	Kliento įneštas pelnas per tam tikrą laikotarpį
Potenciali vertė	Tikėtinas pelnas, kurį galima gauti iš tam tikro vartotojo, naudojant papildomas rinkodaros priemones
Lojalumas	Indeksas nustatantis, kad tai yra įmonės klientas: lojalumas = 1 – nubyrijimo koeficientas (angl. <i>Churn rate</i>). Nubyrijimas aprašo nuolatinių klientų, kurie atsisako santykių su įmone skaičių arba procentą

Kaip jau buvo minėta anksčiau, naudojant kliento vertės rodiklius, galima diferencijuoti klientus į grupes, kurioms galima pritaikyti tik jiems tinkamas strategijas. Iš šių trijų LTV modelio rodiklių galima sudaryti aštuonis segmentus (žiūrėti 2 pav.), tai reiškia, kad galima sukurti ir aštuonis ar mažiau rinkodaros planus.



2 pav. Segmentavimo pavyzdys pagal LTV (Hwang et al., 2004)

Taigi kiekviena įmonė siekia turėti lojalių klientų. Segmentavimas padeda atskirti juos į jau lojalius ir klientus, kuriems reikia pritaikyti kitokią strategiją, siekiant užsitarnauti vartotojų lojalumą. Organizacijoms lojalus klientas yra vertingas dėl kelių priežasčių (Glinskienė, Kvedaraitė, & Kvedaras, 2010):

1. įgyti, išlaikyti bei paversti lojaliais naujus klientus kainuoja daugiau nei išlaikyti jau esančius ir formuoti jų lojalumą;
2. lojalūs klientai mažiau jautrūs kainos pokyčiams;
3. lojalūs klientai linkę skleisti teigiamas žodines reklamas bei rekomendacijas;
4. lojalūs vartotojai perka daugiau su mažesnėmis rinkodaros pastangomis per tam tikrą laiką;
5. lojalūs klientai linkę pateikti dažnus panašius užsakymus, todėl mažiau išlaidų pareikalauja jų aptarnavimas;
6. klientų ir įmonės darbuotojų lojalumas yra tiesiogiai susiję;
7. žinios apie lojalius vartotojus naudojamos įmonėje, siekiant padidinti rinkodaros veiksmų efektyvumą ir sustiprinti santykius su klientais;
8. lojalūs klientai į įmonę žiūri teigiamai, todėl gali pateikti neigiamus pastebėjimus ir duoti laiko išspręsti iškilusias problemas.

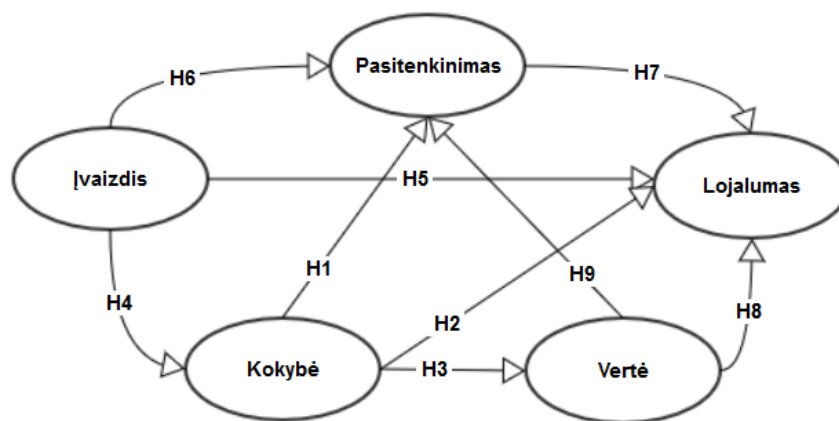
Remiantis šiais faktais, vartotojų lojalumo formavimas yra vienas svarbiausių rinkodaros tikslų, kuris padeda užtikrinti verslo sėkmę bei ilgalaikiuose vartotojo ir įmonės santykiuose, stiprinant klientų lojalumą, gauti naudos abiem šalims.

Lojalumas – tai vartotojo įsipareigojimas ateityje pakartotinai įsigyti paslaugą / prekę arba pakartotinai apsilankyti organizacijoje, nepaisant situacijos ar rinkodaros pastangų pakeisti vartotojo elgseną (Auruškevičienė & Šalčiuvienė, 2003). Taip pat lojalumas gali būti apibrėžiamas kaip kliento noras būti nuolatinio pirkėju, kuris per ilgą laiko tarpą perka ir naudoja tam tikros bendrovės prekes ar paslaugas, tuo pačiu teikia rekomendacijas pažįstamiems. Lojalumas gali būti išreiškiamas per ištikimybę prekei ar jos ženklui, o kartais pardavėjui ar pirkimo vietai (Glinskienė et al., 2010). Galima rasti ir daug kitų lojalumo apibrėžimų, tačiau jie visi yra panašūs, dažniausiai juose išreiškiamas pakartotinis vartotojų pirkimas, nors tai neturėtų indikuoti lojalumo, norint išsiaiškinti, ar vis dėl to klientas yra lojalus, reikia sužinoti jo nuostatą parduotuvės atžvilgiu. Taigi gali būti išskiriami 7 veiksniai, kurie gali formuoti vartotojų lojalumą (Bakanauskas & Pilelienė, 2009):

1. parduotuvės patogumas, kuris susijęs su parduotuvės pasiekiamumu, darbo laiku, prekių išdėstymu ir t. t.;
2. parduotuvės asortimentas – tai pasirinkimo galimybių įvairovė pagal produktų kategorijas, kainas ir kitas išskirtines produktų savybes;
3. prekių kokybė – tai vienas svarbiausių veiksnių, formuojančių parduotuvės įvaizdį;
4. personalo aptarnavimas – šis veiksnys siejamas su darbuotojų elgsena, jų patikimumu;
5. palaikančios paslaugos – tai įmonės teikiamos papildomos paslaugos, siekiant suteikti pridėtinę vertę vartotojui;
6. vertės suvokimas – tai suvokiamos pridėtinės vertės skirtumas lyginant su sumokėta kaina;
7. bendras įspūdis apie parduotuvę.

Šiuose 7 punktuose atsispindi pagrindiniai veiksniai traukiantys vartotojus ir priklausomai nuo to, kas yra jo lojalumo objektas, skirsis lojalumo valdymo priemonės. Literatūroje aptiktas koncepcinis modelis, kuris išskiria tik keturis lojalumui įtaką darančius veiksnius: įvaizdis, kokybė, pasitenkinimas ir vertė (žr. 3 pav.). Kiekvienas ryšys tarp veiksnių turi savo reikšmę (Janita & Miranda, 2013):

- H1, H2 ir H3 – kuo didesnė kokybė, tuo didesnis ir klientų pasitenkinimas, vartotojo lojalumas ir jo suvokiama vertė;
- H4, H5 ir H6 – kuo geresnis įmonės įvaizdis, tuo didesnė suvokiama kokybė, kliento lojalumas ir jo pasitenkinimas;
- H7 – kuo didesnis vartotojų pasitenkinimas, tuo didesnis lojalumas;
- H8 ir H9 – kuo didesnė paslaugų ir produktų teikiama vertė, tuo didesnis klientų lojalumas ir pasitenkinimas.



3 pav. Lojalumą įtakojantys veiksniai (Janita & Miranda, 2013)

Kaip matoma, lojalumas susiformuoja iš kelių kriterijų, kurie gali būti susiję su įmone, preke ar aptarnavimu, dėl kurių bendrovės turi stengtis, kad galėtų sėkmingai išlaikyti esančius lojalus klientus ir įgytus naujus klientus paversti lojaliais. Įmonių atžvilgiu naujas klientas nėra lojalus, kol jis nepereina septynių vartotojo evoliucijos stadijų (Bakanauskas & Pilelienė, 2009):

1. įtariamasis – šiai stadijai priskiriami visi, kurie galėtų tapti įmonės klientais, išsigyjant paslaugą ar prekę. Šie vartotojai vadinami „įtariamaisiais“, nes galima tik įtarti, kad jie pirks, nors tiksliai nėra žinoma;
2. potencialus vartotojas – tai galimas klientas, kuriam reikia įmonės produkto ar paslaugos ir jie gali pirkti, nors dar nėra pirkę;
3. diskvalifikuotas potencialus vartotojas – vartotojas, kurio organizacija nepažįsta tiek, kad suprastų, jog jam nereikia ar negali išsigyti produkto / paslaugos;
4. vartotojas, perkantis pirmą kartą – tai vartotojas nusipirkęs produktą / paslaugą vieną kartą, todėl jis gali būti tiek organizacijos, tiek konkurentų klientas;
5. vartotojas, perkantis pakartotinai – pirkėjas, išsigijęs tas pačias ar skirtingas prekes organizacijoje du ar daugiau kartų;
6. klientas – tai vartotojai perkantys viską, kas yra aktualu. Pirkėjai perka reguliariai, o ryšys tarp kliento ir organizacijos yra stiprus ir tęstinis, kuris apsaugo nuo konkurencijos spaudimo;
7. advokatas – jie perka kaip klientai, tik papildomai skatina ir kitus pirkti įmonės produktus ar paslaugas, skatindami kitus, jie atlieka rinkodaros funkciją.

Be šių apibrėžtų lojalumo stadijų išskiriama dar viena – tai neveiklus vartotojas, kuris buvo įmonės klientas, tačiau nėra nieko pirkęs ilgesnį laiko tarpą nei įprastas pirkimo ciklas. Aukščiau buvo apibrėžtos kliento evoliucijos stadijos, o toliau įvardinamos kliento lojalumo keturios stadijos (Bakanauskas & Pilelienė, 2009):

1. pažintinis lojalumas – tai silpnas lojalumas, kada vartotojo lojalumą lemia pasiūlymo teikiama informacija, kaip kaina ar kokybė. Tokie vartotojai gali pakeisti prekės ženklą, kai tik ras alternatyvų pasiūlymą, labiau tenkinantį naudos ir kaštų santykį;
2. emocinis lojalumas – ši stadija siejama su geru požiūriu į prekės ženklą, o pasitenkinimą vartotojui suteikia lūkesčių patenkinimas;
3. paprastasis lojalumas – tai nuostatų lojalumas, lydimas noro pirkti patikusią prekę ar paslaugą pakartotinai. Šis lojalumas yra stipresnis už emocinį, tačiau vartotojas išlieka jautrus ir susidūręs su trūkumais, gali pradėti ieškoti alternatyvų;
4. veiksmų lojalumas – tai vartotojai, kurie ieškos mėgstamo pasiūlymo, neatsižvelgiant į pastangas, kurių gali prireikti.

Taigi pažintinio lojalumo stadijai priklauso klientai, kuriems svarbios produkto charakteristikos, emocinio lojalumo stadijoje – simpatija prekės ženklui, paprastojo lojalumo stadijoje pasireiškia noras pirkti produktą, o veiksmų lojalumo stadiją pasiekęs vartotojas yra atsidavęs įmonei ir stengiasi pirkti pakartotinai (Bakanauskas & Pilelienė, 2009).

Norint išvengti klientų praradimo išskiriami septyni žingsniai, skirti atsikratyti nepageidaujamo klientų nusigręžimo (Barladean, 2017):

1. išanalizuoti atkrentančių klientų charakteristikas;
2. susipažinti su rinka ir konkurencija;
3. įtraukti ir mokyti savo klientus;
4. pasiūlyti puikų klientų aptarnavimą;
5. ankstyvojoje stadijoje aptikti klientų nusivylimą;
6. pasirūpinti rekomendacinio pobūdžio klientais;
7. pritraukti pasyvius klientus.

Įmonė norėdama optimizuoti ryšių su klientais valdymą, pirmiausia turėtų išanalizuoti savo turimą klientų duomenų bazę ir atlikti klientų segmentavimą, o po jo įvertinti turimų klientų lojalumą. Tačiau norint atlikti sėkmingą analizę svarbu, remiantis turimais duomenimis ir žiniomis, pasirinkti tinkamus segmentavimo ir lojalumo vertinimo metodus, kurie įmonėse turėtų būti naudojami periodiškai.

1.3. Metodai klientų segmentavimui ir lojalumo įvertinimui

Dėmesys klientų nubyrejimui yra svarbus siekiant nustatyti riziką, kurie klientai planuoja palikti įmonę ir kuriuos klientus verta išlaikyti. Klientų atsitraukimo rodiklis yra glaudžiai susijęs su klientų

išlaikymo ir lojalumo rodikliu. Taip pat nubyrejimo rodiklis turi stiprią įtaką klientų vertės apibrėžimui (LTV) (Mutanen, 2006).

Nubyrejimo / neloyalumo (angl. *churn*) valdymas yra neatsiejama dalis nuo ryšių su klientais valdymo (toliau – CRM), tai ypač svarbu įmonėms, siekiančioms užmegzti ilgalaikius santykius su klientais ir didinant jų teikiančią vertę. Klientų praradimai neigiamai veikia įmonės veiklą keliais būdais. Pirmiausia, dėl to iš karto sumažėja pardavimo pajamos ir bendrovė turės pritraukti daugiau klientų, dėl ko patirs papildomas išlaidas. Antra, sutarties nutraukimo atveju, tai ne tik tam tikrų produktų ar paslaugų pardavimų sumažėjimas, bet ir aiškiai apibrėžtas santykių nutraukimas su klientu. Potencialūs būsimieji pinigų srautai per kryžminius ir papildomus pardavimus prarandami. Taigi, tikslios nubyrejimo tikimybės prognozės yra pagrindinis elementas apskaičiuojant kliento gyvavimo trukmės vertę taip pat ir CRM. Tikslumas tampa dar svarbesnis, jei LTV naudojamas išteklių paskirstymui (Risselada, Verhoef, & Bijmolt, 2010).

Nubyrejimas – tai su rinkodara susijęs terminas apibūdinantis vartotoją, kuris migruoja iš vienos įmonės į kitą. Kaip klientas, šis vis dar turi ryšių su pagrindine įmone, tačiau artimiausiu metu taps konkurentų klientu. Jei įmonė nori užkirsti klientui kelią palikti bendrovę, tam reikalingi išlaikymo veiksmai. Nubyrejimo modeliavimas yra svarbus tik išlaikymo perspektyvoje (Glady, Baesens, & Croux, 2009).

Siekiant išsiaiškinti, koks yra santykis tarp klientų atsitraukimo ir klientų lojalumo rodiklių yra atlikta daug tyrimų. Didelė tikimybė, kad klientai su aukštu pasitenkinimu išliks esamais klientais. Lojalūs klientai sunkiai pereina pas kitą paslaugų tiekėją, jie nėra savaime atkrentantys klientai (Dehghan & Trafalis, 2012). Taigi klientų lojalumą galima apskaičiuoti pagal lygtį:

$$\text{Lojalumas} = 1 - \text{Nubyrejimas}$$

Kadangi kliento lojalumas yra susijęs su noru tęsti santykius su kompanija, sprendimas nepabėgti nuo dabartinio paslaugų tiekėjo gali būti prilygintas lojalumui šiai bendrovei, kuris taip pat priklauso nuo laiko trukmės. Klientų elgsenos prognozavimas yra viena pagrindinių kompanijos nerimą keliančių priežasčių, o tiksliai prognozuojant gali būti nustatytos strateginės ir taktinės išsaugojimo priemonės, skirtos klientų judėjimui sumažinti (Dehghan & Trafalis, 2012).

Tradiciniai nubyrejimo valdymo metodai susideda iš dviejų pakopų: modelis skirtas prognozuoti kiekvieno kliento polinkį į pabėgimą, po to keliems pasirinktiems klientams siūlomos kelios išlaikymo paskatos (Lemmens & Gupta, 2013). Dažniausiai nubyrejimo aptikimui naudojami metodai yra klasifikavimo medžių ir logistinės regresijos modeliai, taip pat neseniai buvo pritaikytas mašininis mokymasis (Risselada et al., 2010). Pirmame žingsnyje gali būti naudojama ir daugiau

prognozavimui skirtų metodų: logistinė regresija, diskriminantinė analizė, sprendimų medžiai, neuroniniai tinklai, atsitiktiniai miškai ir kiti. Visi šie metodai tarpusavyje labai skiriasi, tačiau visi jie siekia sumažinti netinkamai suklasifikuotų klientų procentą. Šie metodai ignoruoja, kad klientai nėra vienodai svarbūs įmonei. Todėl mažinant prognozavimo modelio neteisingos klasifikacijos balą nėra lygiavertis procesas įmonės pelno didinimui (Lemmens & Gupta, 2013).

Egzistuojantys algoritmai vis dar turi tam tikrų apribojimų dėl specifinių problemų, išskylančių prognozuojant nubyrežimą. Apibrėžiamos trys pagrindinės neigiamos savybės (Xie, Li, Ngai, & Ying, 2009):

1. duomenys įprastai būna nesubalansuoti, tai reiškia, kad ketinančių palikti įmonę klientų skaičius sudaro nedidelę duomenų dalį;
2. didelis apmokymo taikymas neišvengiamai turės tam tikrą triukšmo lygį;
3. nubyrežimo prognozavimo užduotis reikalauja nubyrežimo tikimybes įvertinimo.

Gerai išanalizuotų duomenų rezultatai duoda pagrindą įstaigų vadovų priimamų sprendimų kokybės didėjimui. Kasdienis klientų praradimas įmonėse parodo žinių trūkumą apie klientų poreikius, kad jie būtų išlaikyti. Tokiu atveju duomenų tyryba ir jos algoritmai yra reikalingi išgauti informacijai slypinčiai duomenyse ir remiantis ja priimti reikiamus sprendimus, todėl klientų nubyrežimo prognozavimas yra dar viena įdomi sritis, kuri traukia valdymo tikslais. Trumpai apibendrinant, prognozavimo modeliai surenka duomenis, suskirsto juos į kategorijas, kad galėtų priimti sprendimą (Nabareseh, 2017).

Dalis ankstesniame poskyryje išvardintų klientų praradimo išvengimo žingsnių gali būti pasiekti panaudojant prognozavimo ir segmentavimo metodus. Rinkodaros specialistai dažnai naudojami kitų įmonių teikiamomis paslaugomis, kad atliktų segmentavimo analizę, ypač jei jiems trūksta statistinių gebėjimų analizuoti. Norint identifikuoti segmentus duomenys turi atitikti kelis kriterijus („Decision Support Systems, LP.“, 2018):

- užtektinai dideli – rinka turi būti pakankamai didelė, kad būtų galima ją segmentuoti;
- įvairūs – tarp rinkos dalyvių turi egzistuoti skirtumai, o jie turi būti išmatuojami, naudojant tradicinius duomenų rinkimo būdus;
- reaguojantys – kada rinka sugrupuojama, turi būti galimybė nukreipti rinkodaros komunikaciją į segmentus, kitu atveju neverta segmentuoti;
- pasiekiami – kiekvienas segmentas turi būti pasiekiamas per vieną ar daugiau žiniasklaidos priemonių;

- nukreipti į skirtingus tikslus – segmentai turi skirtis ne tik demografinėmis ir psichologinėmis savybėmis, bet ir pagal naudą gaunamą iš produkto;
- naudingi – segmentavimas turi sugeneruoti pelną, praplečiant rinkas ir orientuojantis į atitinkamus segmentus.

Vertinant klientus, dažnai prisimenama 80/20 taisyklė arba Pareto principas, kuris teigia, kad 80% pelno gaunama iš 20% pelningų klientų, o likę 80% vartotojų atneša tik 20% pelno. Ši taisyklė leidžia atskirti klientus, kurie generuoja didžiausią pelną įmonei (Raulerson, Malraison, & Leboyer, 2009). Pagal šį Pareto principą, manoma, kad reikia orientuotis į pelningiausius klientus, tačiau yra kitų klientų analizės metodų, kuriais klientai apibūdinami ne tik pelnu, bet ir kitomis pirkimo charakteristikomis, pagal kurias vėliau galima atlikti segmentavimą.

Yra du būdai susegmentuoti rinką – tai apriorinis (angl. *a priori*) ir post hoc (angl. *post hoc*). Apriorinis segmentavimas apima rinkos padalijimą į segmentus be pirminio rinkos tyrimo. Vadybininko intuicija, antrinių duomenų šaltinių analizė, vidinių klientų duomenų bazių analizė ir kiti įvairūs metodai yra naudojami klientų diferencijavimui. Ankstesni post hoc segmentavimo tyrimai laikomi aprioriniais, kada jie yra taikomi ateityje toms pačioms rinkoms analizuoti. Post hoc – pirminės rinkos tyrimai naudojami tikslinės rinkos narių klasifikacijos ir deskriptorių kintamųjų rinkimui. Segmentai nėra apibrėžti iki tol, kol nėra surinkta ir išanalizuota visa susijusi informacija. Daugiamačiai analitiniai metodai naudojami kiekvienam segmentui apibrėžti ir parengti vertinimo algoritmą, skirtą visiems tikslinės rinkos nariams („Decision Support Systems, LP.“, 2018).

Tiesiogiai pagrįsti (angl. *Judgmentally based*) arba rankiniai (*manual*) segmentavimo metodai – dažniausiai naudojami, norint suskaidyti turimus klientus į homogeniškus segmentus. Tipiniu atveju, segmentavimo kriterijai apima ankstesnę pirkimo elgseną, demografinius, geografinius ir psichologinius rodiklius. Ankstesnis pirkimo elgesys laikomas stipriausiu kriterijumi, prognozuojant būsimos reakcijos tikimybę. Literatūroje apibrėžiami klientų LTV nustatymui naudojami modeliai pagal jų elgesį praeityje: RFM, SOW, PCV modeliai. Piniginės dalis (angl. *Share of Wallet – SOW*) – kliento išlaidų dalis tam tikros įmonės produktui ar paslaugai. Kliento vertė praeityje (angl. *Past Customer Value – PCV*) – vertė paremta kliento pelno įnašu įmonei praeityje (Kumar & Shah, 2004). Vienas plačiausiai naudojamų metodų yra RFM, kurį sudaro naujumas (angl. *recency*), dažnumas (angl. *frequency*) ir perkamumas (angl. *monetary*). Naujumas atitinka dienų ar mėnesių skaičių nuo paskutinio pirkimo, dažnumas – tai ankstesnių transakcijų kiekis, perkamumas – tai piniginė suma išleista visiems apsipirkimams arba vidutinė suma išleista apsipirkimo metu. Dažnai naudojamas papildomas išskaidymas pagal produkto kategoriją (angl. *Category of product*), tada metodas vadinamas FRAC. Rankiniu būdu naudojami šie du metodai yra kritiški ir subjektyvūs, taip pat ne

visoms rinkoms šie metodai yra tinkami, ilgaamžiams gaminiams šių rodiklių reikšmės gali būti interpretuojamos kitaip nei trumpalaikiams produktams. Taip pat siekiant patenkinti segmento dydžio apribojimus, derinant smulkius segmentus ir skaidant didelius, kol bus rastas tinkamas sprendimas, gali labai padidinti skaičiavimo laiką (Levin & Zahavi, 2001).

Rečiau segmentavimui taip pat gali būti naudojama latentinės klasės analizė (angl. *latent class analysis*), dažnai apibūdinama kaip kategorijų kintamųjų faktorinė analizė. Ji naudojama siekiant rasti pagrindines konstrukcijas kintamųjų rinkiniuose, tačiau latentinių klasių analizė taip pat gali būti naudojama, grupuojant kategorinius kintamuosius į segmentus, remiantis atsakymais gautais iš daug įvairių kategorinių kintamųjų. Latentinės klasės bando rasti pagrindinius konstruktus, kurie motyvuoja žmones nusipirkti konkretų produktą ar pageidauti tam tikrų produkto savybių („Decision Support Systems, LP.“, 2018).

Literatūroje taip pat minimi automatiniai metodai, kuriais siekiama atmesti priimamus kritiškus ir subjektyvius sprendimus rankinio segmentavimo proceso metu. Šie metodai sujungia duomenų elementus į vieną iš preliminarių klasių. Paprasčiausiu atveju tikslas yra klientus paskirstyti į vieną iš dviejų klasių, pagrįstu tam tikro tipo binariniu atsakymu, pavyzdžiui, lojalus / nelojalus, pirkto / nepirkto ir panašiai. Taigi, medžių klasifikacija (angl. *Tree classification*) yra grįsta pasirinkimu. Keletas literatūroje minimų automatinių medžių klasifikacijų metodų: automatinis sąveikos aptikimas (angl. *Automatic Interaction Detection – AID*), Chi kvadrato AID (angl. *Chi square AID – CHAID*), klasifikacijos ir regresijos medžiai (angl. *Classification and Regression Trees – CART*) ir kt.

Nėra idealaus metodo, kuris būtų naudingas kiekvienam segmentavimo tyrimui, nes kiekviena metodika turi savų privalumų ir trūkumų. Segmentavimo tyrimams paprastai reikia naudoti dvi ar daugiau metodikų, kad būtų pasiekti geriausi rezultatai. Beveik kiekvienu atveju, prieš pasirinkdami geriausią sprendimą, reikia patikrinti keletą metodų. Rinkos segmentavimui taikomi trijų kategorijų analizės metodai: duomenų rengimas, duomenų analizė ir klasifikavimas („Decision Support Systems, LP.“, 2018).

Toliau pateikiami keli plačiau naudojami duomenų tyrybos metodai tarp organizacijų skirti lojalumo detekcijai, segmentavimui bei rinkos analizei (žiūrėti 2 lent.) (Tsiptsis & Chorianopoulos, 2010).

2 lentelė. Dažniausiai naudojami duomenų tyrybos modeliavimo metodai klientų analizei (Tsipstis & Chorianopoulos, 2010)

Modelio kategorija	Modelis	Taikymas
Klasifikavimo modeliai	Neuroniniai tinklai, sprendimų medžiai, logistinė regresija ir kiti	Lojalumo įvertinimui, palankumo produktui detekcijai (cross ¹ /up ² /deep ³ pardavimai)
Klasterizavimo modeliai	K-vidurkio, dviejų žingsnių, Kohonen tinklų/savarankiškai organizuojamas žemėlapis ir kiti	Segmentavimui
Asociacijos ir sekos modeliai	Apriorinis, apibendrinta indukcijos taisyklė, seka	Rinkos krepšelio analizei, prieinamų kanalų analizei

Prieš tai lentelėje išskirti trys populiariausi klientų klasterizavimo metodai, toliau pateikiamas šių metodų palyginimas (žr. 3 lent.).

3 lentelė. Klasterizavimo metodų palyginimas (Tsipstis & Chorianopoulos, 2010)

	K-vidurkio	Dviejų žingsnių	Kohonen tinklų/SOM
Metodikos aprašymas	Pasikartojanti procedūra pagrįsta pasirinkta atstumo priemone (dažniausiai Euklido)	Pirmoje fazėje įrašai per vieną duomenų perdavimą yra sugrupuojami į išankstinius klasterius. Antroje fazėje turimi klasteriai suskirstomi į galutinius klasterius per hierarchinį klasterizavimą	Paremti neuroniniais tinklais. Klasteriai yra išdėstomi tinklelio žemėlapyje su atstumais, parodančiais jų panašumus
Kategorinių klasterizavimo laukų tvarkymas	Taip, perrašant į indikatorius laukus	Taip	Taip, perrašant į indikatorius laukus
Klasterių skaičius	Iš anksto nurodomas klasterių skaičius, todėl tam prireikia keleto bandymų	Klasterių skaičius yra automatiškai nustatomas, remiantis specifiniais kriterijais	Nustatomas maksimalus skaičius išeinančių neuronų skaičius, šie neuronai nustato galimus klasterius
Greitis	Greičiausias	Greitas	Negreitas
Integruotas išvesties tvarkymas	Ne	Taip	Ne
Rekomenduojamas duomenų paruošimas	Standartizavimas ir išskirčių pašalinimas	Standartizavimas ir išskirčių pašalinimas	Standartizavimas ir išskirčių pašalinimas

¹ Cross-selling (liet. *kryžminiai pardavimai*) – tai rinkodaros praktika, kai parduodant pagrindinius produktus, siūloma įsigyti ir susijusias ar papildančias prekes (MarketingSolutions.lt, 2015a).

² Up-selling (liet. *papildomi pardavimai*) – rinkodaros praktika, kada pirkėjui pasiūloma brangiau kainuojanti prekė, siekiant kompanijai sudaryti finansiškai vertingesnę sandorį (MarketingSolutions.lt, 2015c).

³ Deep/down-selling (liet. *pigesnių prekių pardavimai*) – rinkodaros praktika, kada pirkėjui pasiūloma pigesnė prekė, siekiant parduoti daugiau prekių arba parduoti bent vieną, užuot netekus kliento (MarketingSolutions.lt, 2015b).

Palyginus tris klientų segmentavimui skirtus metodus išsiskiria k-vidurkio ir dviejų žingsnių klasterizavimo metodai savo privalumais, o Kahonen tinklų metodas jokių patogumų klasterizavimo procesui nepasiūlo. K-vidurkio metodas yra greičiausias iš visų trijų naudojamų metodų, o dviejų žingsnių metodas yra patrauklus, nes automatiškai parenka klasterių skaičių, leidžia naudoti kategorinius kintamuosius ir suteikia galimybę tvarkyti išvesties informaciją.

Aptarti metodai plačiai naudojami tyrimuose, kurie padeda įvardinti bendrovių klientų segmentus ir išskirti lojalių klientų dalį. Tačiau vieni metodai veikia naudojant visų rūšių duomenis, kiti nesupranta kategorinių kintamųjų, kas yra didelis trūkumas. Taip pat vieni metodai yra labiau automatizuoti nei kiti, kas palengvina klasterizavimo eigą, pagal tai atsirinkti metodai, kurie bus naudojami šio darbo tyrime.

2. Segmentavimui ir lojalumo įvertinimui naudojami metodai

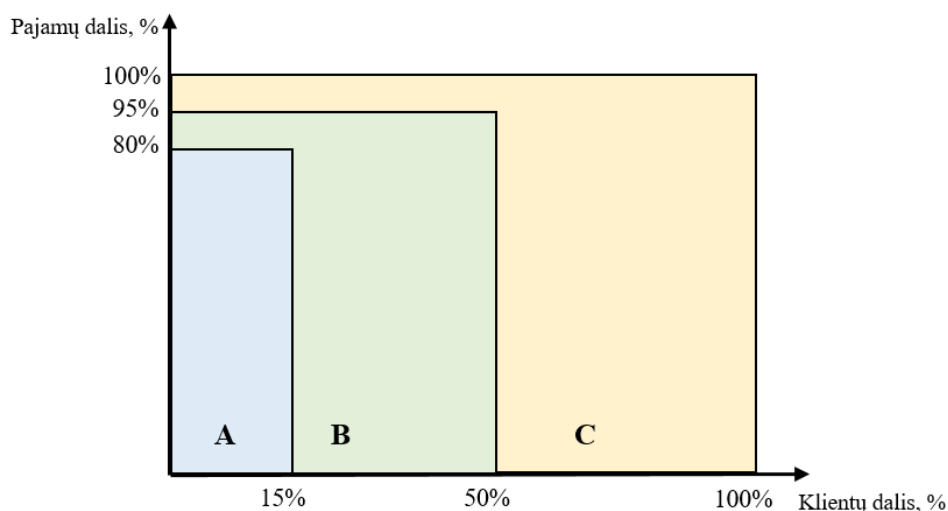
Duomenų tyryba padeda išgauti žinias ir įžvalgas analizuojant didelius duomenų kiekius bei naudojant sudėtingus metodus. Šiuo atveju tolimesniam darbui pasirinkti keli patogūs ir plačiai naudojami metodai, rasti literatūros apžvalgos metu, kurių pagalba bus atliekamas segmentavimas ir klientų lojalumo įvertinimas.

2.1. ABC analizė

Vienas dažniausiai vertės apžvalgai naudojamų metodų ir Pareto taisyklė, kuri turi kelias alternatyvas. ABC analizė yra panaši į 80/20 taisyklę, kitaip dar vadinamu Pareto principu. Ši taisyklė nurodo, kad 80% rezultatų bus gaunama iš 20% pastangų įdėtų šiam rezultatui pasiekti. Naudojant ABC analizę, galima klasifikuoti produktus ar klientus į tris skirtingus segmentus, remiantis iš jų gautu pelnu (4 pav.) (Chu, Liang, & Liao, 2008; Sutevski, 2018):

- A segmentas – šioje grupėje yra didžiausia apyvarta, apie 15 – 20% klientų / produktų sugeneruoja apie 75 – 80% pajamų;
- B segmentas – šioje grupėje yra didesnis klientų kiekis apie 30 – 40% visų klientų / produktų ir jie sudaro tik 15 – 20% įmonės pajamų (nuo 80% iki 95%);
- C segmentas – trečioje grupėje gali būti apie 40 – 50% klientų / produktų, bet jie duos tik 10 – 15% pajamų.

Procentai segmentams pateikiami kaip pavyzdys kiekvienos įmonės atveju, jie gali būti skirtingi, tačiau arti šių skaičių. Taigi 80/20 taisyklė yra panaši į ABC, tik ji apibrėžia dvi grupes, tai 20% klientų, kurie atneša įmonei 80% pajamų ir 80% klientų, kurie sudaro 20% pelno.



4 pav. ABC analizė (sudaryta autorės remiantis Chu et al., 2008; Sutevski, 2018)

Taip pat literatūroje panašiai aprašomas verte pagrįstas segmentavimas – tai klientų bazės padalijimas pagal vertę. Reikėtų pabrėžti, kad tai nėra vienkartinė užduotis. Labai svarbu, kad organizacija galėtų stebėti vertės pokyčius laike. Organizacija turėtų stebėti ir, jei įmanoma įsikišti, kad būtų užkirsta klientų migracija.

Verte pagrįstas segmentavimas yra sukurtas naudojant paprastus skaičiavimus ir nereikalauja sudėtingo duomenų modelio. Vertės segmentų identifikavimas apima klientų rūšiavimą pagal jų vertę. Tipiškų segmentų sąrašas pagal vertę (Tsiptsis & Chorianopoulos, 2010):

- auksas – 20% klientų, kurių vertė yra didžiausia;
- sidabras – 30% klientų, antra pagal vertingumą grupė;
- bronzos – 50% klientų, kurių vertė yra mažiausia.

Taigi visus tris modelius apibendrinus (ABC, 80/20 taisyklę ir segmentavimą pagal vertę), galima daryti išvadą, kad idealus klientų segmentavimas būtų į tris grupes: A segmentas sudarytas iš 20% klientų, B – 30% klientų ir C – 50% klientų, tokiu atveju šiose grupėse atsispindėtų visi paminėti modeliai.

2.2. RFM analizė

Literatūros apžvalgoje pastebėtas populiarus klientų analizei skirtas RFM metodas, kuriuo remiantis galima atlikti klientų segmentavimą. RFM analizė – tai rinkodaros metodas, naudojamas analizuoti klientų elgseną, taip pat padeda atsakyti į klausimus, kada paskutinį kartą klientas pirko, kaip dažnai klientas perka ir kiek klientas išleidžia. Tai naudingas metodas, siekiant geriau diferencijuoti klientus į grupes, kurioms ateityje bus pateikti personalizuoti pasiūlymai. Taip pat ši analizė padeda išskirti klientus, kurie labiau yra linkę reaguoti į reklamas. Taigi RFM analizė priklauso nuo trijų svarbių su pirkimais susijusių kintamųjų: naujumo (R), dažnumo (F) ir perkamumo (M) (Birant, 2016).

RFM analizė turi ir savų privalumų ir trūkumų. Pagrindiniai analizės privalumai yra (Dursun & Caber, 2016):

1. tai galinga priemonė klientų LTV įvertinimui, kuri taip pat gali būti derinama su kitais įprastais duomenų tyrybos metodais;
2. šis metodas laikomas pagrindu kitiems metodams, skirtiems sėkmingam klientų segmentavimui;
3. veiksmingas, prognozuojant klientų reakciją ir siekiant padidinti įmonės pajamas per trumpą laiką.

RFM analizės trūkumai (Dursun & Caber, 2016):

1. šis metodas nėra pakankamas sėkmingam rinkodaros programų generavimui naudojant tik tris kriterijus, nes kiti požymiai, kaip klientų pajamos, gyvenimo būdas ir susidomėjimas skirtingais produktais yra ignoruojami ir neįtraukiami į analizę;
2. didelė koreliacija tarp dažnumo (F) ir perkamumo (M) verčių;
3. nežinojimas potencialių ir nepelningų klientų;
4. skirtingose verslo srityse RFM rodiklių svarba nėra lygiavertė.

Taigi norint panaikinti metodo trūkumus ir atsižvelgti į daugiau kriterijų yra sukurti keli modelio patobulinimai, kuriuos naudojant atliekama platesnė klientų apžvalga (žr. 4 lent.).

4 lentelė. RFM modelių patobulinimas (sudaryta autorės)

Trumpinys	Pavadinimas	Reikšmė
RFM (Birant, 2016)		
R	Naujumas (angl. <i>Recency</i>)	Laiko tarpas nuo paskutinio pirkimo (dienų arba mėnesių skaičius)
F	Dažnumas (angl. <i>Frequency</i>)	Bendras pirkimų skaičius per nurodytą laikotarpį arba vidutinė trukmė tarp pirkimų
M	Perkamumas (angl. <i>Monetary</i>)	Bendra visų pirkimų suma arba vidutinė užsakymo vertė
FRAC (Levin & Zahavi, 2001)		
A	Piniginė vertė (angl. <i>Amount</i>)	Vidutinė kliento vieno pirkimo vertė
C	Produkto kategorija (angl. <i>Category of product</i>)	Perkamų produktų grupė
RFMD (Chiang, 2011)		
D	Nuolaida (angl. <i>Discount</i>)	Vidutinė taikoma nuolaida
LRFM (Wei, Lin, Weng, & Wu, 2012)		
L	Laikas (angl. <i>Length</i>)	Laiko tarpas nuo pirmo pirkimo iki paskutinio pirkimo (dienų arba mėnesių skaičius)
RFMTC (Dursun & Caber, 2016)		
T	Laikas (angl. <i>Time since first purchase</i>)	Laiko tarpas nuo pirmo pirkimo (dienų arba mėnesių skaičius)
C	Praradimo tikimybė (angl. <i>Churn probability</i>)	Naudojama Bernoulli tikimybė atsakyti į klausimą ar klientas pirks

Taip pat siūlomas paprastesnis modelio apskaičiavimas naudojant RFM balą (Wollen, 2017):

$$RFM = \frac{1}{R} * F * \sqrt{M}$$

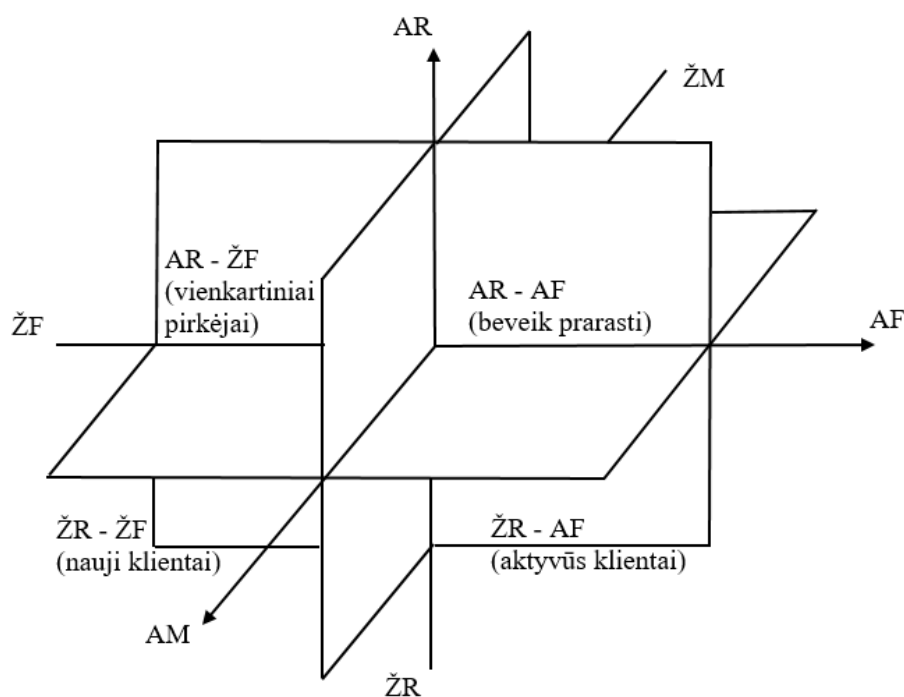
Šioje formulėje naujumas (R) yra neigiamai greitėjanti kreivė. Matematiškai naujumas turi didžiausią poveikį RFM balui (geriausias laikas būtų 1 diena – 1/1, tokiu atveju poveikio nebūtų, jei klientas neaktyvus, pavyzdžiui 2 dienos – 1/2, toks dydis RFM balą sumažina 50%). Kvadratinė šaknis naudojama siekiant sumažinti perkamumo (M) efektą balui.

Priklausomai nuo duomenų kiekio, duomenų dydžio svyravimų, galima jų apimtį suskirstyti į kvartilius ar smulkiau. Galima RFM kriterijus padalinti į tris reikšmes: aukštą, vidutinę ir žemą (žr. 5 lent.). Naujumas šiuo atveju turi priešingas reikšmes, t. y. ilgas laiko tarpas nuo paskutinio pirkimo bus aukštas, nors vertinamas neigiamai, o trumpas laiko tarpas bus vertinamas kaip žemas.

5 lentelė. RFM kriterijų reikšmės (Gajdhar, 2017)

	R	F	M
Aukštas	AR	AF	AM
Vidutinis	VR	VF	VM
Žemas	ŽR	ŽF	ŽM

Šią informaciją galima aiškiau perteikti per modelį (žr. 5 pav.), kuriame išskiriamos keturios klientų kategorijos: vienkartiniai, beveik prarasti, nauji ir aktyvūs klientai, remiantis, kaip dažnai jie vykdo pirkimus ir kada paskutinį kartą pirkė.



5 pav. RFM metodo segmentavimas (sudaryta autorės, remiantis Gajdhar, 2017)

Taigi jei klientas seniai pirko, tačiau iki tol jis įvykdęs daug transakcijų (AH – AF), jis „beveik prarastas“, nes jo apsipirkimai retėja. Tačiau jei klientas neseniai pirko ir dažnai perka (ŽR – AF), tada toks klientas yra aktyvus ar net lojalus. Jei klientas seniai pirko ir mažai, tada galima daryti išvadą, kad tai vienkartinis klientas ir maža tikimybė, kad grįš, o jei klientas mažai pirko, bet neseniai, tai jis yra naujas klientas ir vėliau jis persikels į bet kurią iš išvardintų trijų grupių.

Naudojant tris reikšmes kiekvienam kriterijui, iš viso susidaro 27 RFM analizės deriniai, kuriuos galima suskirstyti į šešis segmentus pagal RFM kriterijų vertę (žr. 6 lent.) (Gajdhar, 2017). Iš lentelės matoma, kad keliems segmentams svarbiausi yra vienas ar du kriterijai iš trijų, pavyzdžiui, pelningiausiems klientams svarbiausia, kad būtų didelis perkamumas, o naujiems klientams pagal prieš tai pateiktą modelį, turi būti neseniai įvykęs pirkimas ir jų turi būti nedidelis kiekis. Tačiau yra segmentų, kuriems svarbūs visi trys kriterijai, pavyzdžiui, geriausi klientai turi būti neseniai pirkę, atlikę daug ir vertingų užsakymų.

6 lentelė. RFM analizės deriniai (Gajdhar, 2017)

Segmentas	RFM	Aprašymas	Rinkodaros veiksmai
Geriausi klientai	ŽAA	Klientai, kurie sudaro daug sandėrių, išleidžia daug ir yra neseniai pirkę	Panaudoti lojalumo programą, kad jie būtų išlaikyti
Pelningiausi klientai	XXA	Klientai, kurie išleidžia dideles sumas	Galimybė up ir cross pardavimams
Lojalūs klientai	ŽAX	Klientai, kurie yra neseniai ir daug atlikę pirkimų	Grįžtamojo ryšio programa
Blėstantys klientai	AAA	Aukštos vertės klientai, tačiau didelė tikimybė, kad jie bus prarasti	Skatinti pakartotinius klientų pirkimus
Mažos vertės prarasti klientai	AŽŽ	Klientai, kurie yra nedaug pirkę ir pirkimai yra mažos vertės, o paskutinis pirkimas atliktas seniai	Pakartotinis klientų pritraukimas
Nauji klientai	ŽŽX	Neseniai tapę klientais	Personalizuota komunikacija orientuotą į cross pardavimus

Taigi remiantis RFM analize galima susegmentuoti klientus ir pagal tai pateikti tinkamą rinkodaros strategiją, tačiau ši analizė gali būti tik kaip pagrindas apžvelgiant klientų lojalumo ir segmentavimo temą, o tolimesniems grupavimo veiksams gali būti naudojami ir kiti duomenų tyryboje populiarūs metodai.

Duomenų tyryba, naudojant RFM kintamuosius, taip pat gali apimti skirtingus duomenų tyrybos metodus, tokius kaip neuroniniai tinklai, sprendimų medis, CHAID, genetinis algoritmas ir kiti (Birant, 2016).

Toliau pateikiamas pavyzdys (žr. 7 lent.), kaip galima klientus suskirstyti į segmentus, naudojant RFM ir k-vidurkio klasterizavimą (Birant, 2016).

7 lentelė. Klientų segmentavimo pavyzdys (Birant, 2016)

Klasteris	Dydis	Naujumas (vidurkis)		Dažnumas (vidurkis)		Perkamumas (vidurkis)		RFM modelis	Kliento grupė
		Day	R	#	F	TL	M		
C1	309	65,4	4,57	6,28	4,89	485,1	4,79	R↑F↑M↑	Best
C2	392	83,5	4,32	1,52	3,44	146,8	3,42	R↑F↑M↑	Valuable
C3	415	75,1	4,44	1,18	3,05	70,1	1,49	R↑F↑M↓	Shopper
C4	300	202,4	2,86	1,01	2,02	69,5	1,47	R↑F↓M↓	FirstTime
C5	425	247,8	2,22	4,27	4,51	378,4	4,67	R↓F↑M↑	Churn
C6	135	325,8	1,38	2,26	3,76	137,5	2,94	R↓F↑M↓	Frequent
C7	381	290,1	1,86	1,00	1,41	138,1	3,33	R↓F↓M↑	Spenders
C8	309	339,1	1,35	1,00	1,00	69,5	1,53	R↓F↓M↓	Uncertain
Iš viso	2666		2,85		3,01		2,95		

K-vidurkio klasterizavimas naudojamas, siekiant sugrupuoti klientus su panašiomis RFM vertėmis. Pateiktame pavyzdyje klientai sugrupuoti į aštuonias tikslines rinkas pagal laikotarpį nuo paskutinio sandorio – naujumą, dažnumą ir vieno apsipirkimo metu išleidžiamą pinigų sumą. Klasterių skaičius modeliavimo metu nustatytas 8, nes aštuoni (2x2x2) galimi įvesties deriniai (RFM), kurie gali būti gaunami priskiriant rodykles aukštyn arba žemyn pagal klasterio vidurkį, tai yra R, F, M vertė yra mažesnė arba didesnė už bendrą vidurkį. Jei klasterio vidutinė R (F, M) vertė viršijo bendrą R (F, M) vidurkį, tada parenkama rodyklė rodanti aukštyn, kitu atveju rodo žemyn. Pavyzdžiui, R ↑ F ↓ M ↓ reiškia, kad klientų segmento naujumo vidutinė pradinė vertė yra didesnė už bendrą vidurkį, o dažnumas ir piniginių vidutinės vertės yra mažesnės už bendrą vidurkį.

Įvardinamos aštuonios klientų grupės: geriausi klientai (angl. *Best*), kurie įmonei yra reikšmingiausi, jų visi rodikliai yra aukšti. Vertingi klientai (angl. *Valuable*), kurių taip pat visi rodikliai aukšti, bet žemesni už geriausių grupės. Pirkėjai (angl. *Shopper*) – klientai, kurie pastoviai perka, bet už nedideles sumas. Pirmą kartą apsilankę klientai (angl. *FirstTime*) – klientai, kurių naujumas yra aukštas, bet kiti rodikliai prasti. Prarasti klientai (angl. *Churn*) – priešingi rodikliai nei naujo kliento, tai toks klientas, kuris daug pirko, bet jau seniai. Dažni klientai (angl. *Frequent*) – tai klientai atliekantys daug užsakymų, bet už mažą vertę, o paskutinis užsakymas atliktas seniai. Investuotojai (angl. *Spenders*) – klientai, kurie seniai ir mažai pirko, bet didelės vertės užsakymus

darydavo. Neaiškūs klientai (angl. *Uncertain*), kurie atneša įmonei mažiausią vertę, nes visi RFM rodikliai yra prasti. Lentelėje taip pat pateikiami klasteriai su jiems priklausančiu klientų skaičiumi, jų vidutinės realias ir įvertintas R, F ir M vertes. Paskutinėje eilutėje taip pat rodomas bendras visų klientų vidurkis. RFM stulpelyje pateikiamas kiekvieno klasterio RFM modelis. Taigi šiame pavyzdyje pateiktas RFM modelio panaudojimas taikant sudėtingesnį metodą, taigi tokiu būdu gali būti RFM modelis panaudojamas ir su kitais klasterizavimo metodais.

2.3. Dviejų žingsnių klasterizavimas

Remiantis atlikta literatūros analize, apžvelgus klasterizavimui naudojamus metodus pastebėta, kad patogiausias klasterizavimo metodas su RFM klientų analize būtų dviejų žingsnių metodas, kuris yra greitas, leidžia naudoti kategorinius kintamuosius, taip pat automatinio būdu parenka tinkamiausią klasterių skaičių ir leidžia patogiai tvarkyti išvesties rezultatus – profiliuoti rastus klientų klasterius, suteikiant jiems aprašymą pagal būdingas charakteristikas.

Dviejų žingsnių (angl. *two steps*) klasterių analizės SPSS programos (tai specializuota statistinė programinė įranga) metodas, skirtas atpažinti natūralias grupes ar klasterius duomenyse, kurie kitaip nebūtų matomi, todėl dažnai naujose situacijose rekomenduojama klasterinę analizę pradėti šiuo metodu. Šioje procedūroje naudojamas algoritmas turi keletą svarbių savybių, kurios skiriasi nuo tradicinių klasterizavimo metodų (Bafadal, 2011):

- analizei tinka tiek kategoriniai, tiek tolydiniai kintamieji;
- automatinis optimalaus klasterių skaičiaus parinkimas;
- gali efektyviai analizuoti didelius duomenų masyvus.

Siekiant valdyti kategorinius ir kiekybinius kintamuosius šis metodas naudoja log-tikėtimumo (angl. *log-likelihood distance*) arba Euklido atstumo matą. Euklido atstumas naudojamas tik tais atvejais, kada visi kintamieji yra tolydiniai. Taip pat laikoma, kad kiekvienas tolydus kintamasis turi normalų (Gausinį) pasiskirstymą ir kiekvienas kategorinis kintamasis turi daugiareikšmį pasiskirstymą. Empirinis vidinis testavimas rodo, kad procedūra yra stipri nepriklausomybės prielaidų ir pasiskirstymo prielaidų pažeidimams (Bafadal, 2011).

Pirmajame žingsnyje duomenys sugrupuojami į daug mažų klasterių. Antrajame žingsnyje maži klasteriai sujungiami į norimą skaičių klasterių, o jei klasterių skaičius nėra žinomas, šis metodas automatiškai parinks geriausią klasterių kiekį. Dviejų žingsnių klasterinė analizė išsprendžia klasterizavimo greičio problemą, kuri iškyla analizuojant didelius duomenų rinkinius. Pirmajame žingsnyje naudojami greiti metodai suformuojantis mažus klasterius, o antrajame žingsnyje, kada maži klasteriai sujungiami į stambesnius, naudojami klasikiniai metodai, kurie yra tikslūs, tačiau lėti

dideliems duomenų rinkiniams (Vaitkevičius, 2012). Toliau pateikiami dviejų žingsnių algoritmo detalesni etapų veikimo principai (SPSS, 2001):

1. 1 žingsnis – pasiruošimas klasterizavimui. Pirminiame klasterizavimo žingsnyje naudojamas nuoseklus grupavimo būdas, jis po vieną nuskaito įrašus ir nusprendžia ar dabartinis įrašas turėtų būti prijungtas prie jau suformuotų klasterių, ar pradėti naują klasterį, atsižvelgiant į atstumo kriterijų. Toliau vykdoma procedūra, kurios metu sukuriamas modifikuotas klasterio bruožų (CF – angl. *Cluster Features*) medis. CF medis apima kelių lygių mazgus (angl. *nodes*), o kiekvienas mazgas susidaro iš įrašų. Lapo (angl. *leaf*) įrašas (įrašas lapo mazge) atvaizduoja norimus sub-klasterius. Ne lapų mazgai ir jų įrašai nukreipia naujus įrašus į teisingą lapo mazgą (Bafadal, 2011).
2. 2 žingsnis – duomenų sugrupavimas į galutinius klasterius. Šis žingsnis paima dalines grupes sudarytas pirmajame žingsnyje ir sugrupuoja į pageidaujamą skaičių grupių. Kadangi antrame etape dalinių grupių skaičius yra mažesnis už originalių įrašų skaičių, todėl galima naudoti ir tradicinius klasterizavimo metodus. Naudojant SPSS, pirmiausia naudojamas aglomeracinis hierarchinis klasterizavimo metodas, nes jis efektyviai veikia kartu su automatine klasterizavimo procedūra. Norint nustatyti, kuris klasterių skaičius yra geriausias, kiekvienas klasterių sprendimas lyginamas naudojant Bajeso (BIC) arba Akaike (AIC) klasterizavimo kriterijus (Bafadal, 2011).

Taigi šis metodas puikiai tinka dideliems duomenų masyvams segmentuoti dėl algoritmo veikimo logikos, kada grupavimas vyksta per du etapus ir jų dėka klasterizavimas atliekamas greitai.

2.4. Naudojami detektoriai

Lojalumo nustatymui gali būti naudojami įvairūs mašininio mokymo algoritmai: logistinė regresija, sprendimų medžiai, atsitiktiniai miškai, k-artimiausių kaimynų klasifikatorius, tiesinė diskriminantinė analizė ir daug kitų. Kiekvienas algoritmas veikia remdamasis skirtingais principais (Milošević, Živić, & Andjelković, 2017). Šiame darbe naudojami šeši detektoriai.

Logistinė regresija (angl. *logistic regression*) – tai tikimybinis dvejetainis klasifikavimo algoritmas, kuris naudoja svertinį funkcijų derinį išvedant teigiamą arba neigiamą klasę priklausančių atvejų tikimybę (Milošević et al., 2017). Logistinė regresija yra prognozavimo modeliavimo metodas, kada yra koreliacija tarp rezultato tikimybės r jo prognozuojamų kintamųjų, kaip pateikta lygtyje (Nabareseh, 2017):

$$\log \left[\frac{\pi_i}{(1 - \pi_i)} \right] = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik}$$

Lygtyje π_i – rezultatų tikimybė, β_1, \dots, β_k – koeficientai, X_1, \dots, X_{ik} – prognozuojami kintamieji. Šis metodas patrauklus tikimybių prognozavimui, nes yra matematiškai ribotas gauti tikimybes diapazone [0,1] ir parametrai vertinami pakankamai lengvai (Westreich, Lessler, & Funk, 2010).

Atsitiktinių miškų (angl. *Random Forest*) algoritmas generuoja daugybę skirtingų sprendimų medžių ir juos naudoja derinant daugumą rinkinių naujos klasės prognozavimui. Sprendimų medis – klasifikavimo algoritmas, kuris naudoja įvesties apmokymo duomenis, kad generuotų taisyklėmis pagrįstą medį. Prognozuodamas klasę algoritmas paprasčiausiai peržiūri medį formuojančius parametrus ir naujos klasės ypatybes (Milošević et al., 2017). Šis metodas yra patogus, nes turi tik du parametrus, kuriuos reikia apibrėžti (Maroco et al., 2011):

1. atsitiktinių medžių skaičių miške, t. y. koks bus klasifikatorių kolektyvo dydis;
2. dydį, kuris nusako, iš kiek atsitiktinai parinktų kintamųjų galima rinktis formuojant medžio mazgą.

Šie parametrai gali būti lengvai optimizuojami, nors atsitiktiniai miškai nėra labai jautrūs jų vertėms. Paprastai kuo daugiau medžių tuo geriau, tačiau klasifikavimo klaidai stabilizuotis dažniausiai pakanka 250 medžių (Probst & Boulesteix, 2017).

Tiesinė diskriminantinė analizė (angl. *Linear Discriminant Analysis*) – tai seniausiai naudojamas tiesinio modeliavimo metodas, analizuojantis skirtumus tarp dviejų ar daugiau stebėjimų ir padedantis rasti taisyklę, kaip atsižvelgiant į kiekybinių kintamųjų reikšmes, priskirti naujam stebėjimui grupę. Šis metodas naudojamas objektų kategorijų derinimui ar atskyrimui bei taikomas nustatant kintamąjį, kuris išskiria dvi ar daugiau kategorijų, kurios padeda prognozavime. Diskriminantinės analizės funkcija (Nabareseh, 2017):

$$\max \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2}$$

Diskriminantinė analizė stengiasi rasti tokią įėjimo kintamųjų projekciją, kad atstumas tarp klasių vidurkių joje būtų maksimalus, o sklaida klasės viduje būtų minimali.

K-artimiausių kaimynų (angl. *k-Nearest Neighbors*) metodas – naują objektą palygina su mokymo aibės objektais, kurie yra panašūs į naują. Norint naują objektą priskirti kuriai nors klasei yra skaičiuojami atstumai nuo naujo objekto iki visų apmokymo aibės objektų. Objektas priskiriamas klasei, kuriai priklauso dauguma k jo kaimynų. Artimiausio kaimyno vertė kiekvienam kintamajam y_i apskaičiuojamas pagal formulę (Nabareseh, 2017):

$$\hat{y}_i = \frac{1}{k} \sum_{x_j \in N(x_i)} x_j$$

Formulėje x atitinka x_i kaimyną, o $N(x_i)$ ir k yra fiksuotos konstantos. Klasifikavimo tikslumas priklauso nuo kaimynų skaičiaus k parinkimo, kuris parenkamas eksperimentiškai. Kuo didesnė mokymo aibė, tuo didesnis turėtų būti kaimynų skaičius. Šis algoritmas dažniausiai naudojamas klasifikavimui, tačiau gali būti naudojamas ir vertinimui bei prognozavimui (Nabareseh, 2017).

Neuroniniai tinklai (angl. *Neural Network*) plačiai naudojami prognozavimo ir klasifikavimo uždaviniams spręsti. Neuroninis tinklas – neuronų (apskaičiuotų vienetų) susiejimas atsižvelgiant į jų svorius. Šis metodas yra daugiasluoksnis, nes sudarytas iš įvesties, vieno ar daugiau paslėptų bei išvesties sluoksnių. Įvesties sluoksnyje randasi elementą aprašanti informacija, paslėptuose sluoksniuose realizuojama uždavinio logika, o išvesties – rezultatas (Maroco et al., 2011). Dirbtiniai neuroniniai tinklai skaičiuojami pagal įvesties signalus ir svarbos svorį. Įvesties signalai tvarko kombinuotą funkciją su svoriais ir slenkstine reikšme, tada aktyvuojama aktyvavimo funkcijos, kad pateikti išvedimo signalą (Nabareseh, 2017):

$$y_j = f(x, w_j) = f(P_j) = f\left(\sum_{i=0}^n x_i w_{ij}\right)$$

Šioje funkcijoje j yra generinis neuronas, x – įvesties signalai, w_j yra svoriai, P_j – potencialas, y_j – išėjimo signalas.

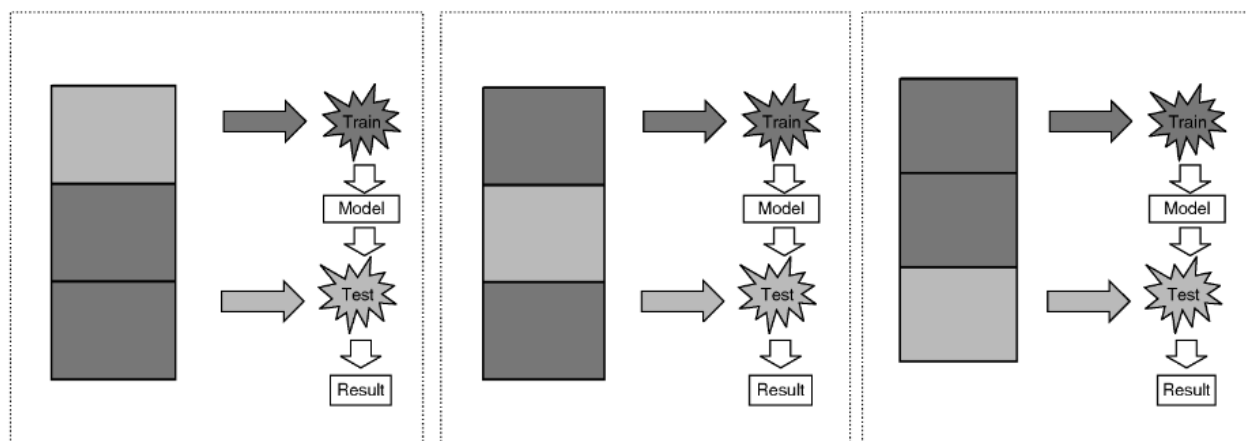
Naudojant šiuos algoritmus galima įvertinti lojalumą, tačiau pasirinkti, kuris algoritmas yra geresnis gali padėti kryžminė patikra, apie kurią rašoma sekančiame poskyryje.

2.5. Detekcijos gerumo įvertinimas

Klasifikavimo algoritmų rezultatams įvertinti dažniausiai naudojamas k -blokų kryžminės patikros metodas. Kryžminė patikra (angl. *Cross validation – CV*) yra statistinis mokymosi vertinimo ir lyginimo metodas, padalijantis duomenis į atskiras imtis. Vienas imtis atidedama modelio testavimui, o kitos naudojamos apmokyti modelį. Įprastai kryžminės patikros metu apmokymo ir testavimo rinkiniai turi pereiti per eilę iteracijų, kad visi duomenys būtų patikrinti. Pagrindinė kryžminės validacijos forma yra k -blokų (angl. *k-fold*) kryžminis patikrinimas. Kitos formos yra specialios k -blokų arba pakartotinės k -blokų kryžminės validacijos (Mudry & Tjellström, 2011).

K -blokų kryžminės validacijos metu duomenys suskaidomi į k vienodo dydžio segmentus arba blokus. Po to k apmokymo ir tvirtinimo iteracijos yra vykdomos taip, kad per kiekvieną iteraciją

vienas duomenų blokas yra saugomas tikrinimui, kol kiti $k-1$ blokai yra naudojami apmokymui. Toliau pateikiamas trijų blokų pavyzdys (žr. 6 pav.), kuriame dvi tamsesnės sekcijos naudojamos apmokymui, o viena šviesesnė naudojama patvirtinimui. Duomenų tyryboje dažniausiai naudojamas 10 blokų kryžminė patikra ($k = 10$) (Mudry & Tjellström, 2011).

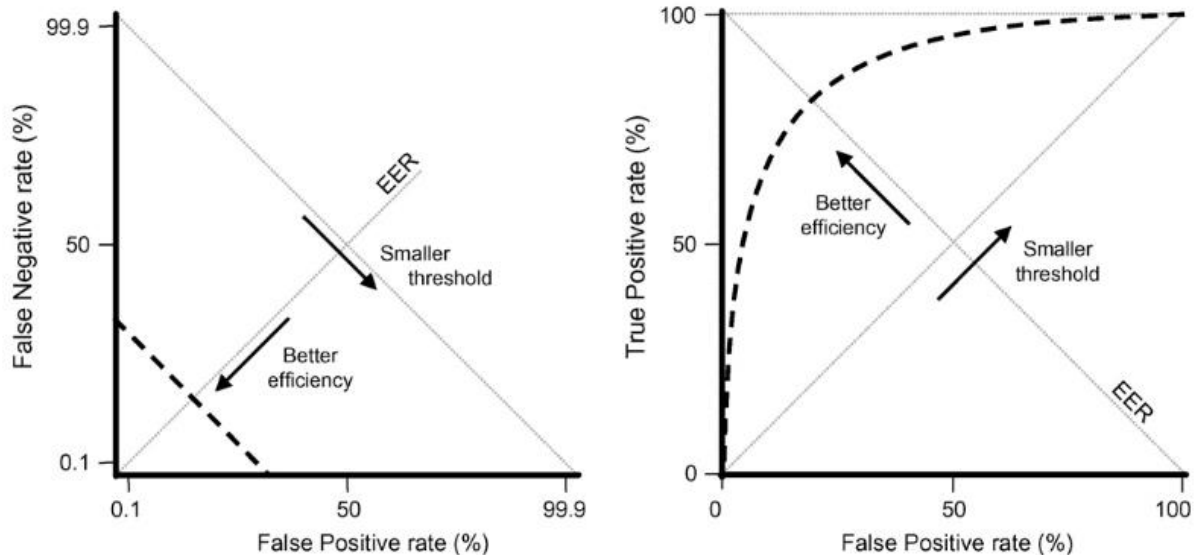


6 pav. Trijų blokų kryžminės validacijos pavyzdys (Mudry & Tjellström, 2011)

Kryžminė patikra naudojama įvertinti ar palyginti mokymosi algoritmus. Kiekvienoje iteracijoje vienas ar daugiau apmokymo algoritmų naudoja $k-1$ duomenų blokų apmokyti vieną ar daugiau modelių, o tada apmokytų modelių pareikalaujama numatyti duomenis esančius patvirtinimo bloke. Stratifikuotos kryžminės patikros atveju blokai parenkami taip, kad kiekvienas būtų panašus į visą imtį ir išlaikytų pusiausvyrą tarp grupių, pagal kurią bus atliekama detekcija, pavyzdžiui tarp lojalių ir nelojalių. Kiekvieno apmokymo algoritmo veiksmus, kiekviename bloke, galima stebėti naudojant tam tikrą stebėjimo metriką, pavyzdžiui tikslumą (angl. *accuracy*). Po to k vykdymo metrikos pavyzdžiai bus prieinami kiekvienam algoritmui. Šio metodo naudojimo pagrindiniai tikslai (Mudry & Tjellström, 2011):

- Įvertinti pasiekto duomenų modelio rezultatus naudojant vieną algoritmą;
- Palyginti dviejų ar daugiau algoritmų efektyvumą ir rasti geriausią algoritmą turimiems duomenims.

Testavimo metu pateikiamas rezultatas, kuris atspindi įvesties vektoriaus tikimybę priklausyti pageidaujama klasei. Šis rezultatas turi būti lyginamas su ribine verte, kad būtų galima apskaičiuoti sumaišymo matricą. Po to gaunama kreivė, kuri atspindi klaidingai teigiamus (arba klaidingai priimtus) sprendimus lyginant su klaidingai neigiamais rezultatais (klaidingai atmetais) skirtingomis ribinėmis reikšmėmis (žr. 7 pav. kairėje) – tokia kreivė vadinama DET (Sáenz-Lechón, Godino-Llorente, Osmá-Ruiz, & Gómez-Vilda, 2006).



7 pav. DET (kairėje) ir ROC (dešinėje) kreivės (Sáenz-Lechón et al., 2006)

Kitas atvejis skirtas atvaizduoti klaidingai teigiamus rezultatus atsižvelgiant į tikrai teigiamus rezultatus, tam skirta ROC kreivė (žr. 7 pav. dešinėje) (Sáenz-Lechón et al., 2006). ROC (angl. *Receiver Operating Characteristic*) kreivė – tai grafikas vaizduojantis klasifikatoriaus, atskiriančio dvi klases, jautrumo ir specifiškumo santykį (Rapečka, 2015). Dažniausiai naudojamas ploto po ROC kreive (angl. *AUC – area under ROC curve*) įvertis, kuris nusako eilės nustatymo kokybę. Jei modelis veikia prastai, tada AUC bus artimas 0,05, jei gerai, tada AUC artimas 1. Ši sritis taip pat leidžia apskaičiuoti kitas statistikas, siekiant nustatyti ar egzistuoja statistinis reikšmingumas tarp skirtingų kreivių gautų iš tų pačių duomenų. Dar vienas būdas nustatyti pasikliautinąjį intervalą – apskaičiuoti standartinę AUC paklaidą (Sáenz-Lechón et al., 2006).

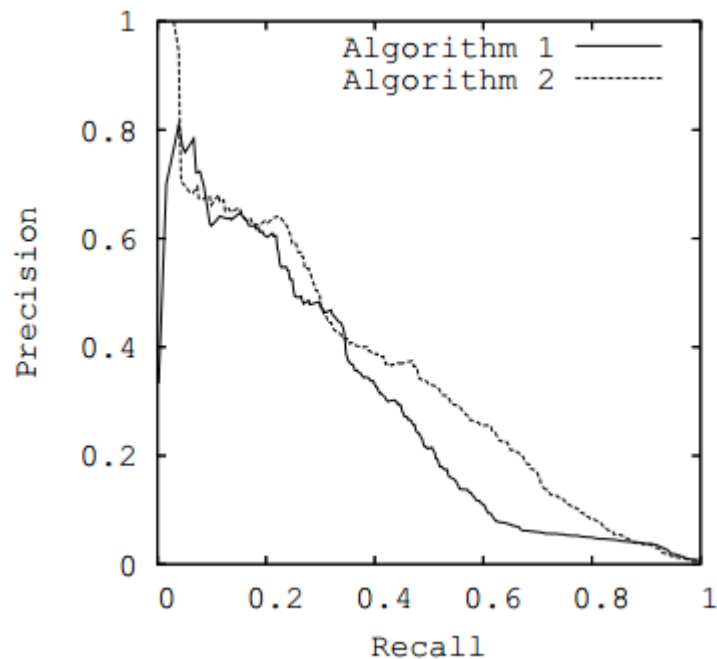
Detektoriaus klaidos (angl. *DET – detector error trade off*) kreivė yra ROC modifikacija naudojant logaritmines ašis. Lygių klaidų įvertis (angl. *EER – equal error rate*) yra detektoriaus veikimo taškas (gaunamas parinkus tinkamą slenkstį) ant ROC / DET kreivės, kada jautrumas tampa lygus specifiškumui – tai nutinka kreivei kertant įstrižainę (žr. 7 pav.). Kuo geresnis detektorius, tuo kreivė bus arčiau apatinio kairiojo kampo. ROC kreivė schemeje (žr. 7 pav. kairėje) artėja prie viršutinio kairiojo kampo, nes parodo modelio efektyvumą. Kitas svarbus skirtumas tarp kreivių – DET brėžiama įprastai nukreipta skale, todėl jei duomenų pasiskirstymas iš dviejų klasių yra artimas normaliam, kreivė bus tiesinė. Įprastai tai bus atvejis, kada duomenys gaunami iš kryžminio patikrinimo apskaičiuojant skirtingų blokų vidurkius (Sáenz-Lechón et al., 2006).

Preciziškumo – prisiminimo (angl. *PR – Precision – recall curve*) kreivė rodo santykį tarp preciziškumo (angl. *precision*) (teigiamos prognozės vertės) ir prisiminimo (angl. *recall*) (atšaukimo) kiekvienai galimai ribinei vertei. Toliau pateikiamos formulės, kaip gaunami šie dydžiai (Ekelund, 2017):

$$\text{Taikinio preciziškumas} = \text{Teigiama prognostinė vertė} = \frac{TT \text{ skaičius}}{TT \text{ skaičius} + KT \text{ skaičius}}$$

$$\text{Taikinio prisiminimas} = \text{Jautrumas} = \frac{TT \text{ skaičius}}{TT \text{ skaičius} + KN \text{ skaičius}}$$

Grafikuose x ašis rodo preciziškumą, o y ašis – prisiminimą (žr. 8 pav.). Paprastai preciziškumas ir prisiminimas yra atvirkščiai susiję, kada padidėja preciziškumas, prisiminimas sumažėja ir atvirkščiai (Tolay, 2018). Taigi kiekvienas PR kreivės taškas parodo pasirinktą ribą. Pasirinkus apribojimą galima pamatyti prisiminimą ir preciziškumą. Ši kreivė dažnai būna zigzago formos, kuri gali leisti ir kilti, todėl kreivės dažniau kerta viena kitą nei ROC kreivės, o tai gali sukelti sunkumų lyginant kreives tarpusavyje. Tačiau kreivės esančios aukščiau turi geresnį našumo lygį. Pagrindinis skirtumas tarp PR ir ROC kreivių tai, kad tikrai neigiami rezultatai nėra naudojami PR kreivei sudaryti (Ekelund, 2017).



8 pav. Preciziškumo – prisiminimo kreivės pavyzdys (Tolay, 2018)

Klasifikavimui naudojamos dvi klasės. Kiekvienas atvejis yra susietas su vieno elemento rinkinio {T;N} teigiamos ir neigiamos klasės žymėmis. Norint atskirti faktinę ir prognozuojamą klasę naudojamos {T;K} žymės klasės modelių prognozėms. Suteikus klasifikatoriams žymes yra keturi įmanomi rezultatai. Jei atvejis yra teigiamas ir jei jis klasifikuojamas teigiamai, tada laikomas tikrai teigiamu (angl. *true positive*), jei klasifikuojamas neigiamai, tada jis laikomas klaidingai neigiamu (angl. *false negative*). Jei atvejis yra neigiamas ir jis klasifikuojamas neigiamai, tada laikomas tikrai neigiamu (angl. *true negative*), o jei klasifikuojamas teigiamai, tada laikomas klaidingai teigiamu

(angl. *false positive*). Atspindėti gautus rinkinius padeda sumaišymo matrica (Fawcett, 2004) (žr. 8 lentelę). Šiame variante sumaišymo matrica vaizduoja nubyrijimo aptikimą, kada nelojalumas priskiriamas teigiamai klasei, o lojalumas neigiamai.

8 lentelė. Sumaišymo matricos klasės (sudaryta autorės, remiantis Fawcett, 2004)

		Lojalumas (tikros klasės)		
		Lojalus (Neigiamas)	Nelojalus (Teigiamas)	
Hipotezės klasės	Lojalus	<i>Tikrai Neigiamas (TN)</i>	<i>Klaidingas Neigiamas (KN)</i>	Neigiama prognozė (TNV) = TN/(TN + KN)
	Nelojalus	<i>Klaidingai Teigiamas (KT)</i>	<i>Tikrai Teigiamas (TT)</i>	Teigiama prognozė (TTV) = TT/(TT + KT)
		<i>Visi lojalūs = KT + TN</i>	<i>Visi nelojalūs = TT + KN</i>	
		<i>Specifiškumas = TN / (TN + KT)</i>	<i>Jautrumas = TT / (TT + KN)</i>	

Klasifikatoriaus jautrumas parodo, kiek lojalių klasės elementų buvo suklasifikuota teisingai su bendru faktiniu lojalių klasės skaičiumi. Kuo jautrumas didesnis, tuo daugiau lojalių klientų buvo teisingai priskirti lojalių klasei. Specifiškumas parodo nelojalių klasės klientų priskyrimą nelojalių klientų klasei. Taip pat remiantis 8 lentele galima apskaičiuoti klasifikavimo kokybę, remiantis papildomomis keliomis formulėmis (Paulauskienė & Kurasova, 2013; Rapečka, 2015):

$$\text{Bendras klasifikavimo tikslumas} = \frac{TT \text{ skaičius} + TN \text{ skaičius}}{\text{Visų objektų skaičius}}$$

$$\text{Klasifikavimo klaida} = 1 - \text{bendras klasifikavimo tikslumas}$$

Labai svarbus sumaišymo matricos įvertis yra Kappa patikimumo įvertinimui, kuris įvertina visas gautas reikšmes. Jei įvertis mažesnis arba lygus 0 – modelis nepatikimas, 0,01 – 0,20 patikimumas labai silpnas, 0,21 – 0,40 patikimumas neblogas, 0,41 – 0,60 vidutiniško patikimumo modelis, 0,61 – 0,80 stiprus modelis, 0,81 – 1,00 beveik puikus modelis (Mchugh, 2012). Kappa apskaičiuojamas pagal formulę (Ben-David, 2008):

$$Kappa = \frac{P_0 - P_c}{1 - P_c}$$

P_0 yra bendras klasifikavimo tikslumas, o P_c – atsitiktinis tikslumas, kuris gaunamas iš sekančios formulės:

$$P_c = \frac{(TN + KT) * (TN + KN) + (KN + TT) * (KT + TT)}{(TT + TN + KT + KN) * (TT + TN + KT + KN)}$$

$$P_c = \frac{Visi nelojalūs * Nelojalių prognozė + Visi lojalūs * Lojalių prognozė}{Visų objektų skaičius * Visų objektų skaičius}$$

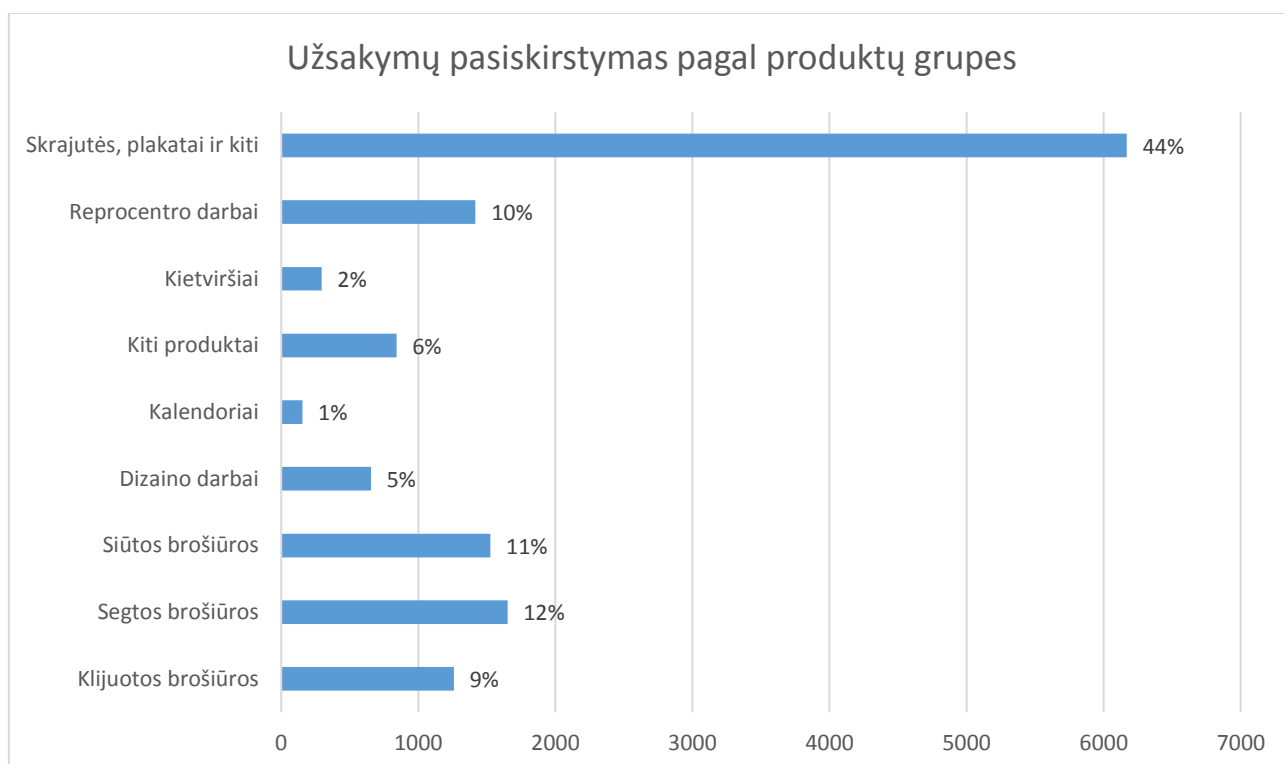
Taigi lojalumo detekcijai naudojamų algoritmų įvertinimui naudojama kryžminė patikra, kada duomenų testavimo metu išvedami rodikliai, kurie gali būti atvaizduojami per ROC, DET ir PR kreives, taip pat naudojant sumaišymo matricą, kurioje atsispindi tokie svarbūs įverčiai kaip bendras klasifikavimo tikslumas ir Kappa.

3. Poligrafijos įmonės klientų segmentavimas ir lojalumo patikra

3.1. Įmonės ir jos klientų analizė

Šis poskyris skirtas apžvelgti poligrafijos įmonės veiklą, kuo ji užsiima, bei jos klientus. Susipažinus su įmonės klientais ir jos veikla bus galima paprasčiau atlikti klientų segmentavimą bei lojalumo vertinimą ir pagrįsti gautus rezultatus.

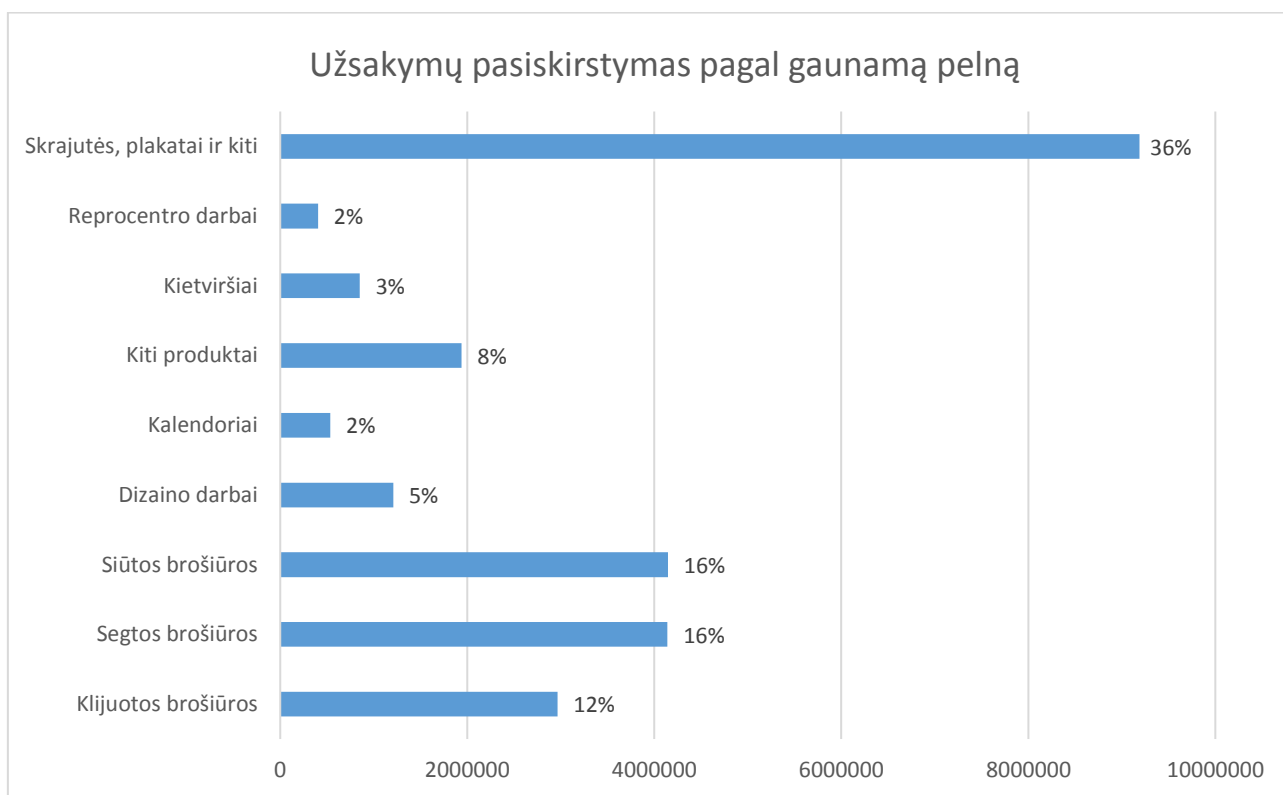
Poligrafijos įmonė vykdo savo veiklą Lietuvoje, Kauno mieste, 27 metus. Anksčiau agentūrine veikla užsiėmusi įmonė jau daug metų siūlo dizaino, leidybos ir poligrafijos darbus. Toliau pateikiamas grafikas (žiūrėti 9 pav.), kuriame matoma paslaugų ir produktų pardavimai per 5 metus. Visa įmonės klientų analizė pateikiama nuo 2013.01.01 iki 2017.12.31 laikotarpio duomenų. Per šį laikotarpį 1793 klientai atliko 22 930 užsakymų.



9 pav. Paslaugų pasiskirstymas įmonėje pagal populiarumą

Iš grafiko matoma, kad įmonės teikiamos dizaino paslaugos sudaro tik 5% visų užsakymų, o spaudos paruošiamųjų darbų skyriaus (toliau – reprocentras) paslaugos – 10%, taigi likę 85% įmonės užsakymų susidaro iš spaudos darbų. Patys populiariausi produktai: skrajutės, lankstinukai, plakatai ir kiti paprasti, mažos vertės gaminiai – 44%, segtos, siūtos ir klijuotos brošiūros – 32%, o visi kiti produktai sudaro tik 9% užsakymų. Tačiau pagal gaunamą pelną (žiūrėti 10 pav.), didžiausią pelną įmonė gauna iš brošiūrų – 44% ir iš skrajučių bei kitų paprastų gaminių – 36%, kitų paslaugų ir produktų užsakymai sudaro tik 20% įmonės pelno. Taigi iš pateiktų įmonės užsakymų matoma, kad

įmonės pagrindas yra brošiūrų ir paprastų gaminių (t. y. skrajutės, vizitinės, plakatai, lankstinukai) gamyba, taip pat įmonė teikia savo klientams papildomas reprocentro ir dizaino paslaugas, kurių poreikis yra, bet sukuriama vertė nėra didelė.



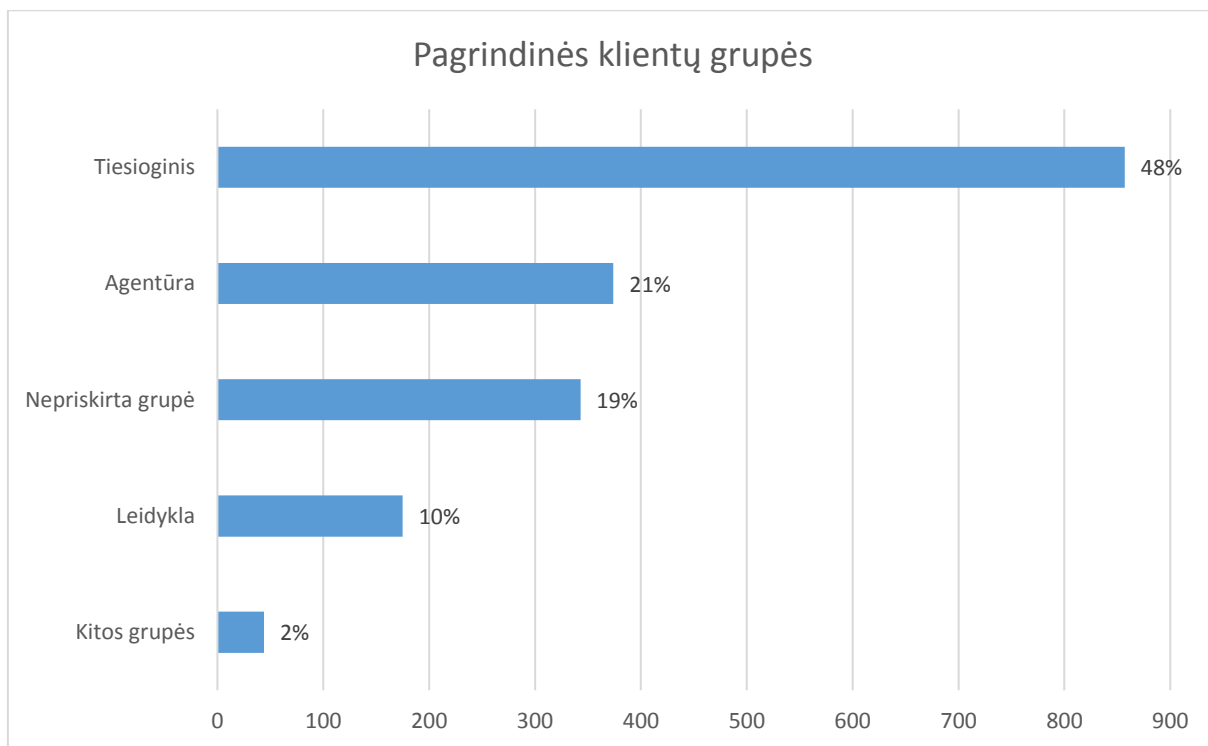
10 pav. Paslaugų pasiskirstymas pagal gaunamą pelną

Šios įmonės klientų poreikiai dažniausiai nėra standartiniai, dauguma užsakymų yra išskirtiniai dėl produktų sudarančių medžiagų, matmenų, spaudos ar apdailos darbų, todėl išskirtos tik apibendrintos produktų grupės, o detalesnio jų išskaidymo nėra.

Įmonėje klientai yra skirstomi į grupes pagal kliento vykdomą veiklą (žr. 11 pav.):

1. agentūros – tai agentūros ir tarpininkai;
2. tiesioginiai – tai menininkai, fotografai, dizaineriai ir kiti tiesioginiai klientai;
3. leidyklos;
4. kitos grupės: asociacijos, mokyklos, universitetai, galerijos, muziejai, spaustuvės ir kitos mažos grupės.

Šios grupės padeda suprasti klientų poreikį, kokių produktų, kokio sudėtingumo, kokios vertės ir kaip dažnai užsakymai gali būti pateikiami įmonei. Iš viso per penkerius metus aktyvių klientų, kurie pateikė po vieną ar daugiau užsakymų, buvo 1793, iš kurių beveik pusė klientų – 48% yra tiesioginiai, 21% – agentūros.



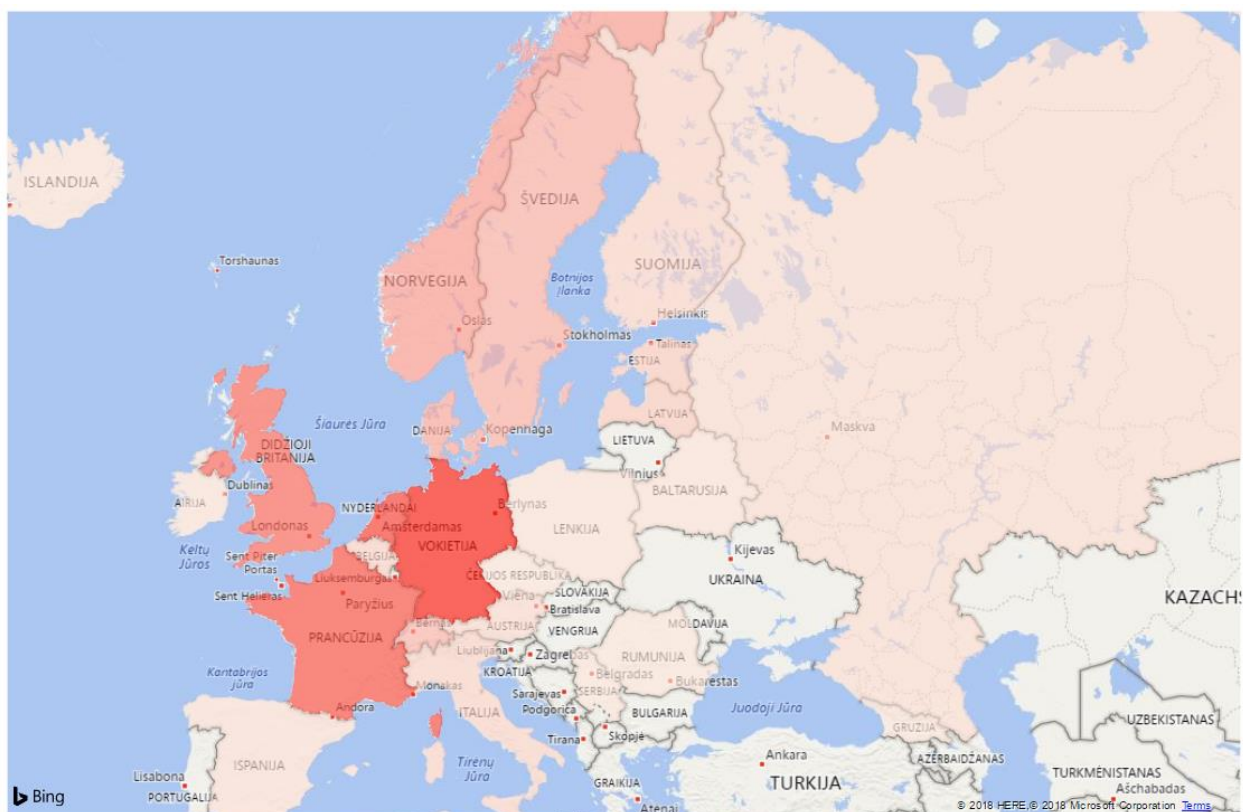
11 pav. Pagrindinės klientų grupės

Nors tiesioginių klientų yra daugiau nei agentūrų, tačiau per 5 metų laikotarpį vidutiniškai viena agentūra pateikė daugiau užsakymų nei visi kiti klientai:

- agentūra – 23 užsakymai;
- tiesioginiai – 12 užsakymų;
- leidyklos – 9 užsakymai;
- klientai be grupės – 6 užsakymai;
- kitos grupės – 5 užsakymai.

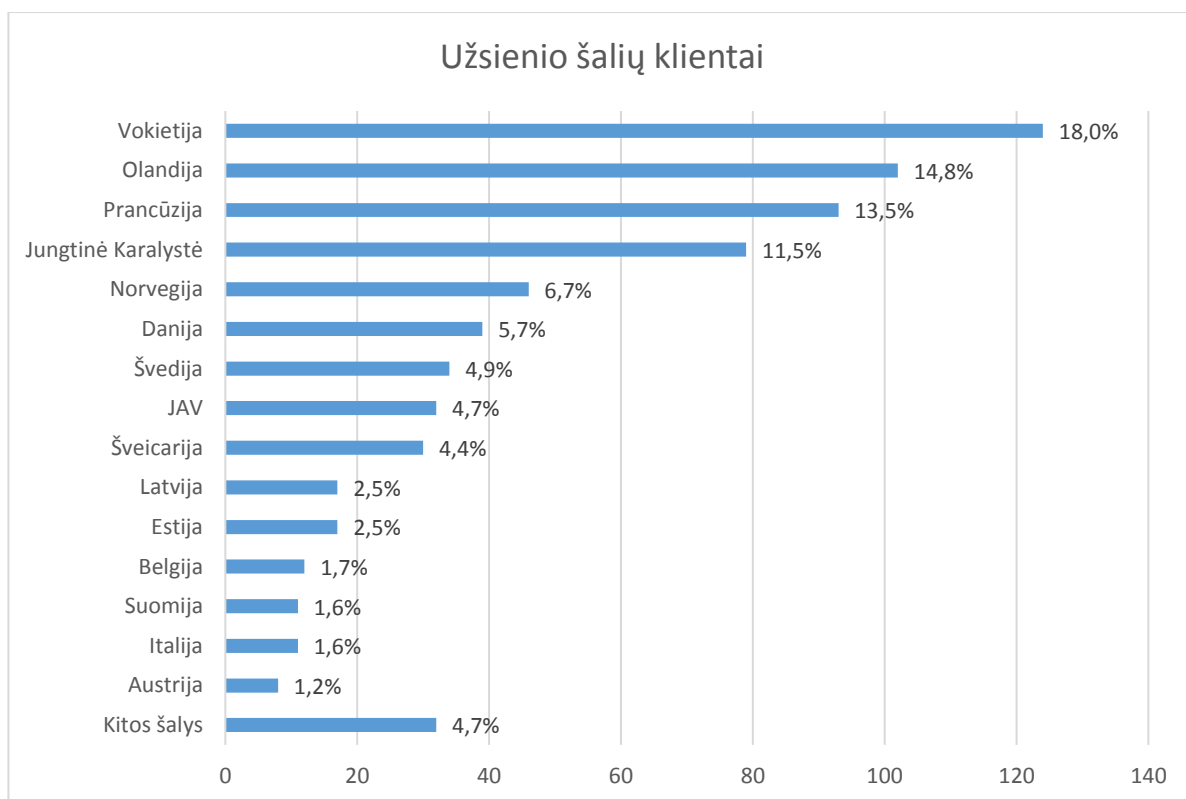
Taip pat dėl savo vykdomos veiklos agentūros užsakymus teikia dažniau nei tiesioginiai ar kitų grupių klientai. Todėl šių klientų vertė yra labai didelė analizuojamai įmonei.

Įmonės vykdoma veikla yra žinoma ne tik Lietuvoje, bet ir užsienyje, net 38% klientų yra iš užsienio. Dauguma Lietuvos pirkėjų (84% Lietuvos klientų) yra iš Kauno ir Vilniaus. Daugiausia užsienio klientų yra iš keturių Europos šalių: Vokietijos, Olandijos, Prancūzijos ir Jungtinės Karalystės, kurie sudaro 58% visų užsienio šalių klientų (žiūrėti 12 pav.).



12 pav. Klientų pasiskirstymas Europoje (spalvos ryškumas parodo klientų tankį šalyje)

Taigi kaip matoma iš paveikslėlio įmonės klientai ne tik iš Europos Sąjungos, bet ir iš kitų šalių (žr. 13 pav.). Kitų šalių grupę sudaro šalys, kuriose yra mažiau nei 1% įmonės klientų.



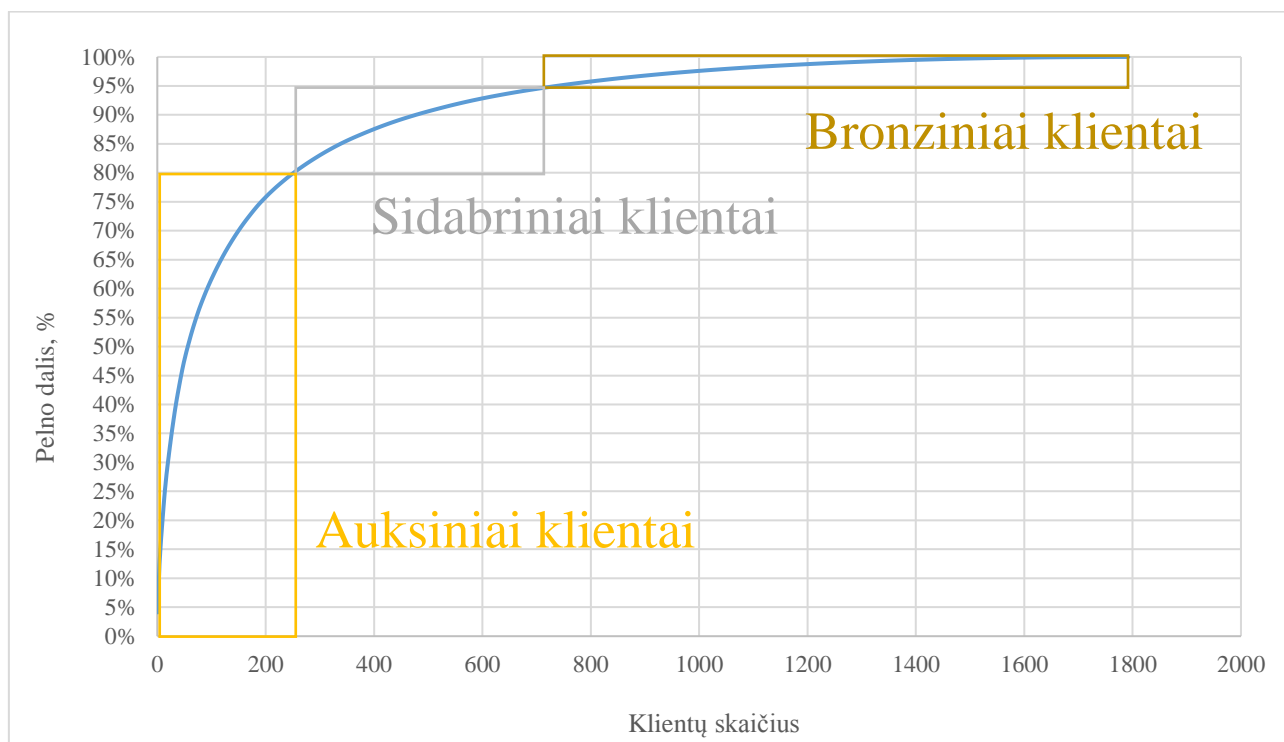
13 pav. Klientų pasiskirstymas tarp užsienio šalių

Taigi įmonės klientai yra iš įvairių šalių, užsiimantys skirtinga veikla, todėl jų poreikiai taip pat skiriasi. Įmonė negali su visais klientais dirbti remiantis viena rinkodaros strategija, nes jie neturi vienodos perkamosios galios. Atliekant klientų užsakymų analizę, galima išvelgti daugiau panašumų tarp klientų, kuriais remiantis galima būtų juos suskirstyti į homogeniškas grupes ir remiantis jomis nuspręsti, koks turėtų būti tolimesnis darbas su esamais ir naujai aptiktais klientais.

3.2. Klientų segmentavimas

Detalesnis grupavimas nei klientų suskirstymas pagal vykdomą veiklą, šalį ar piniginę vertę įmonėje nebuvo atliktas, todėl toliau bus atliekamas klientų segmentavimas remiantis daugiau rodiklių, kurie leistų išvelgti grupes, kurios galėtų padėti vykdyti rinkodaros tikslus.

Naudojant ABC metodą klientai suskirstyti į tris grupes: auksiniai, sidabriniai ir bronziniai, remiantis klientų užsakymų verte, kuri gauta susumavus kiekvieno kliento per penkis metus įvykdytų užsakymų pelną (žr. 14 pav.).



14 pav. Klientų segmentavimas ABC metodu

Gautas įmonės klientų pasiskirstymas pagal pelną yra panašus į literatūroje aprašytus A, B ir C segmentų dydžius:

- A segmentas – tai auksiniai klientai, kurie sudaro 14,40% visų klientų ir jie sukuria 80% įmonės pelno;

- B segmentas – tai sidabriniai klientai, kurie sudaro 28,96% visų įmonės klientų ir 15% įmonės pelno;
- C segmentas – bronzinių klientų grupę sudaro daugiau nei pusė įmonės klientų – 56,64%, kurie įmonei atnešė tik 5% pelno.

Toliau pateikiama kiekvieno segmento duomenų statistika (žr. 9 lent.). Iš imties intervalo matoma, kad kiekviename segmente skirtumas tarp didžiausios ir mažiausios pelno sumos yra labai didelis, o tai parodo, kad šiose grupėse klientų vertė nėra panaši, todėl ABC metodo nepakanka ir reikia detalesnio klientų segmentavimo.

9 lentelė. Visų užsakymų vertės aprašomoji statistika klientų segmentams pagal ABC analizę

	Auksas	Sidabras	Bronza
Imties dydis	258 (14,40%)	519 (28,96%)	1015 (56,64%)
Vidurkis	79132,94	7363,17	1127,30
Mediana	40814,42	6228,95	965,62
Standartinis nuokrypis	114828,58	3804,53	824,51
Imties intervalas	1025137,46	14082,18	2991,19
Minimumas	17270,04	3001,57	5,56
Maksimumas	1042407,50	17083,75	2996,75

Išsamesnei klientų apžvalgai pasirinktas RFM patobulintas metodas, kurio metu papildomai naudojamas T kintamasis. Šia analize klientai apžvelgiami per keturis dydžius:

1. naujumas (R) – nusakantis, kada paskutinį kartą buvo atliktas užsakymas (dienomis);
2. dažnumas (F) – parodo, kiek iš viso klientas yra pateikęs užsakymų;
3. vertė (M) – apibrėžia užsakymo vidutinę vertę;
4. laikas (T) – parodo, nuo kada pirkėjas yra įmonės klientas (dienomis).

Šių rodiklių pasirinkimo tikslas – rasti lojalius, vertingus bei prarastus / blėstančius įmonės klientus, kuriuos verta susigrąžinti. Lentelėje pateikiama bendra visų klientų duomenų statistika pagal RFMT kintamuosius (žr. 10 lentelę).

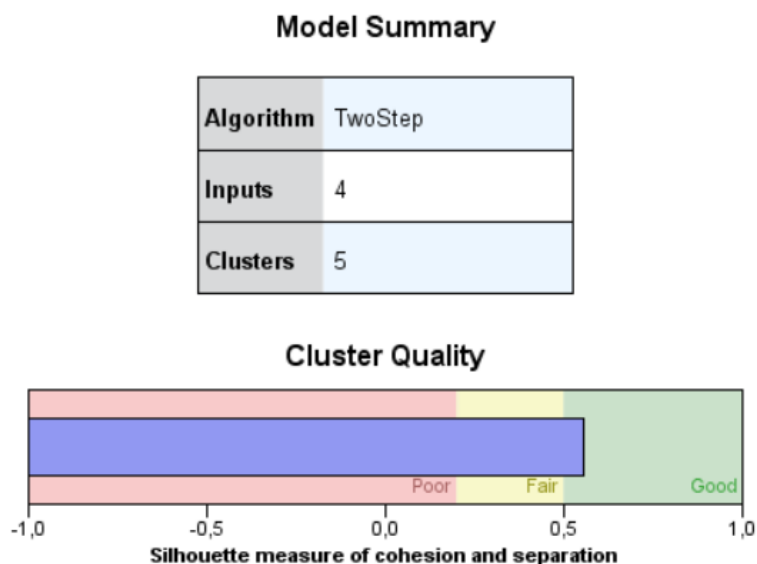
10 lentelė. Visų įmonės klientų RFMT kintamųjų aprašomoji statistika

	R	F	M	T
Imties dydis	1792	1792	1792	1792
Vidurkis	661,46	12,35	1388,11	1092,81
Mediana	579	2	911,92	1147
Standartinis nuokrypis	516,98	47,70	1949,20	561,03
Imties intervalas	1822	817	31563,66	1820
Minimumas	2	1	5,56	4
Maksimumas	1824	818	31569,22	1824

Iš duomenų statistikos pastebima, kad pasirinkus naujus kintamuosius grupavimui pagal mažiausią reikšmę duomenyse bus galima rasti vienkartinį, naujų klientų segmentus, kurių rodikliai bus žemi, o pagal didžiausią reikšmę išvelgiami vertingiausių ir lojalių klientų segmentai, nes jų

rodikliai yra labai aukšti. Tačiau dėl didelių skirtumų tarp reikšmių klientai negali būti diferencijuojami į kelias stambias grupes.

Taigi segmentavimui pasirinktas SPSS programos dviejų žingsnių metodas, kuris automatinio būdu gali parinkti tinkamiausią segmentų skaičių. Panaudojus keturis kintamuosius, rezultato kokybės balas yra 0,6, kuris reiškia, kad parinktas klasterių skaičius 5 yra geras (žr. 15 pav.).



15 pav. RFMT visų duomenų klasterizavimo kokybė naudojant dviejų žingsnių metodą (silueto plotis = 0,6)

Šiame modeliavime iš keturių kintamųjų mažiausią reikšmę grupių išskyrimui turėjo vidutinė užsakymų vertė, visi kiti kintamieji buvo vienodai reikšmingi, kuriais remiantis geriausio segmentavimo metu klientai išskaidyti į penkis segmentus (žr. 11 lentelė).

11 lentelė. Klientų segmentai atsižvelgiant į RFMT reikšmes

Segmento nr.	Segmento pavadinimas	Naujumas (vidurkis)	Dažnumas (vidurkis)	Vidutinė užsakymo vertė (vidurkis)	Laikas (vidurkis)	RFMT modelis
1	Geriausi lojalūs klientai	104,00	248,46	1062,85	1770,05	R↓F↑M↑T↑
2	Pelningiausi blėstantys klientai	344,35	20,20	1204,87	1612,73	R↑F↑M↑T↑
3	Pelningiausi prarasti klientai	1372,58	3,50	2507,54	1570,09	R↑F↓M↑T↑
4	Nauji klientai	189,19	3,54	1203,38	365,11	R↓F↓M↑T↓
5	Mažos vertės prarasti klientai	756,21	2,29	694,46	870,38	R↑F↓M↓T↑

Pavadinimai segmentams priskiriami remiantis RFMT rodikliais ir dėl skirtingų pirkimo charakteristikų, kurios priklauso nuo klientų grupių, kurios yra anksčiau apibrėžtos, nesiremiamos kintamųjų vidurkiais, o atsižvelgiama į dviejų žingsnių suformuotus klasterius. Papildomai atliekant

dviejų žingsnių segmentavimą išvesti ABC ir klientų grupės, kurių dėka galima matyti klientų pasiskirstymą segmentuose (žr. 1 priedą). Taigi sudaryti penki segmentai (žr. 16 pav.):

Clusters

Input (Predictor) Importance
■ 1,0 ■ 0,8 ■ 0,6 ■ 0,4 ■ 0,2 ■ 0,0

Cluster	1	2	3	4	5
Label	Geriausi lojalūs klientai	Pelningiausi blėstantys klientai	Pelningiausi prarasti klientai	Nauji klientai	Mažos vertės prarasti klientai
Description	Klientai dažnai užsakantys didelės vertės užsakymus	Klientai retai darantys užsakymus dėl mažo poreikio, tačiau pastovūs klientai	Pamiršti klientai, kurių užsakymų vidutinė vertė didelė, tačiau seniai atliko paskutinį užsakymą	Nauji klientai, kurie neseniai pradėjo naudotis įmonės paslaugomis	Labai mažos vertės klientai, kurie tik išbandė įmonės teikiamas paslaugas
Size	2,3% (41)	21,3% (381)	24,7% (442)	25,4% (456)	26,3% (472)
Inputs	Frequency 248,46	Frequency 20,20	Frequency 3,50	Frequency 3,54	Frequency 2,29
	Recency 104,00	Recency 344,35	Recency 1 372,58	Recency 189,19	Recency 756,21
	Time 1 770,05	Time 1 612,73	Time 1 570,09	Time 365,11	Time 870,38
	Monetary 1 062,85	Monetary 1 204,87	Monetary 2 507,54	Monetary 1 203,38	Monetary 694,46
Evaluation Fields	Segment Auksas (100,0%)	Segment Sidabras (44,1%)	Segment Bronza (51,1%)	Segment Bronza (66,9%)	Segment Bronza (87,3%)
	Customer Group AGENTŪRA (48,8%)	Customer Group TIESIOGINIS (53,0%)	Customer Group TIESIOGINIS (59,3%)	Customer Group TIESIOGINIS (37,1%)	Customer Group TIESIOGINIS (43,6%)

16 pav. Visų klientų segmentavimas naudojant dviejų žingsnių klasterizavimo metodą (silueto plotis = 0,6)

- geriausi lojalūs klientai – tokių klientų yra labai mažai, tik 41, kurie sudaro 2,3% visų klientų. Visi šie pirkėjai priklauso auksinei grupei, nes ilgą laiką yra įmonės klientai ir atlikę daugiausiai užsakymų dėl šiame segmente populiariausios klientų kategorijos – agentūros poreikio;
- pelningiausi blėstantys klientai – grupės dydis yra panašus į likusių trijų grupių, 44% grupės klientų priklauso sidabrinei grupei ir didžiąją dalį sudaro tiesioginiai klientai, todėl užsakymų skaičius nėra didelis lyginant su geriausiais klientais, o naujumo vidurkis yra labai žemas, nes tokių klientų užsakymų poreikis nėra dažnas;
- pelningiausi prarasti klientai – šių klientų vidutinė užsakymų vertė yra aukšta, tačiau jie trumpai buvo įmonės klientais ir atliko nedaug užsakymų, todėl dauguma pirkėjų priklauso bronzinei grupei, tačiau šios grupės klientus verta susigrąžinti;

4. nauji klientai – dauguma šių klientų per paskutinius metus įvykdė savo pirmą ir paskutinį užsakymą, o užsakymų vidutinis pelnas yra pakankamai aukštas, šių klientų lojalumą verta formuoti pasirenkant tinkamas rinkodaros strategijas;
5. mažos vertės prarasti klientai – dauguma klientų priklauso bronzinei grupei, nes užsakymų vertė yra labai žema ir jų paskutinis užsakymas atliktas daugiau nei prieš 2 metus, todėl jie yra prarasti įmonės klientais, kurių neverta susigrąžinti.

Geriausiams ir vertingiausiems klientams turėtų būti taikoma išlaikymo strategija, kuri didintų klientų vertę įmonėje ir neleistų jiems pabėgti. Dalį pamirštų išlaidžių klientų įmonei reikėtų pabandyti prisivilioti dar kartą, nes jų užsakymų vertė yra didelė. Naujų klientų vertė yra vidutinė, tačiau šie klientai nevisi buvo išlaikyti, reikėtų jiems priminti apie įmonę ir paskatinti juos atlikti daugiau užsakymų. Mažiausiai dėmesio reikėtų skirti pamirštiems mažos vertės klientams, kurie negali sukurti didesnės vertės įmonei.

Tarp geriausių ir vertingų klientų slepiasi lojalūs klientai, kurie ilgą laiką yra įmonės klientai ir atlieka daug užsakymų, tačiau šio segmentavimo metu jie neišsiskyrė. Taip pat šiose grupėse tarp pirkėjų išliko dideli skirtumai, kaip ir pirmo segmentavimo metu, todėl tolimesnis klientų segmentavimas atliekamas sujungiant ABC, RFMT ir dviejų žingsniu metodus.

Apjungus šiuos metodus, dviejų žingsnių segmentavimas taikomas tris kartus, kiekvienam ABC segmentui atskirai. Visu pirma, atliekama kiekvieno segmento apžvalga pagal RFMT kintamuosius, o tada pateikiamas grupės papildomas skaidymas.

Pirmiausia, atliktas auksinės grupės segmentavimas. Šią grupę sudaro 258 įmonės klientai su skirtinga charakteristika (žr. 12 lentelę), imčių intervalai dar vis išlieka platūs, kritinių reikšmių kiekis (žr. 2 priedą) nėra didelis, tačiau atitinkamai pagal tai turėtų susidaryti atskiri segmentai.

12 lentelė. Auksinės klientų grupės RFMT kintamųjų aprašomoji statistika

	R	F	M	T
Imties dydis	258	258	258	258
Vidurkis	418,31	61,84	2837,30	1582,69
Mediana	223	25,50	1617,59	1729,5
Standartinis nuokrypis	455,66	112,77	3752,84	330,15
Imties intervalas	1711	817	31232,95	1637
Minimumas	2	1	336,26	187
Maksimumas	1713	818	31569,22	1824

Naudojant dviejų žingsnių metodą, gautas rezultatas įvertintas 0,5, kurio svarbiausi kintamieji yra naujumas, laikas bei dažnumas, o remiantis jų reikšmėmis išskirti keturi segmentai (žr. 17 pav.):

1. pamiršti geriausi klientai – tai klientai, kurie vieno užsakymo metu sukuria labai aukštą vertę, vidutinė vieno užsakymo vertė 6106 eurai, tačiau jų paskutinis pirkimas įvykdytas seniai,

vidutiniškai prieš 3 metus, didžiausią šių klientų dalį sudaro tiesioginiai klientai. Įmonė turėtų priminti apie save šiems klientams bei pabandyti juos susigrąžinti;

2. buvę lojalūs klientai – tai klientai, kurie ilgą laiką buvo įmonės pirkėjai ir yra atlikę daug užsakymų, kurių vidurkis siekia 43 užsakymus, tačiau paskutinis užsakymas atliktas prieš daugiau nei pusę metų. Tai blėstantys klientai, kuriuos reikia sustabdyti palikti įmonę;
3. lojalūs klientai – tai pastovūs įmonės pirkėjai, atlikę daug užsakymų ir paskutinis užsakymas atliktas neseniai, dažniausiai tokie klientai yra agentūros;
4. lojalūs geriausi klientai – pirkėjai, kurie palyginus su 3 segmentu neilgai yra įmonės klientais, tačiau atlikę daug užsakymų, o vidutinė užsakymo suma aukšta.

Cluster	1	2	3	4
Label	Pamiršti geriausi klientai	Buvę lojalūs klientai	Lojalūs klientai	Lojalūs geriausi klientai
Size	26,4% (68)	54,3% (140)	10,1% (26)	9,3% (24)
Inputs	Recency 1 056,31	Recency 234,60	Recency 42,27	Recency 89,62
	Time 1 625,62	Time 1 664,96	Time 1 801,35	Time 744,25
	Frequency 13,59	Frequency 43,31	Frequency 317,92	Frequency 29,25
	Monetary 6 106,52	Monetary 1 739,57	Monetary 1 024,65	Monetary 1 941,65
Evaluation Fields	Customer Group TIESIOGINIS (60,3%)	Customer Group TIESIOGINIS (50,7%)	Customer Group AGENTŪRA (53,8%)	Customer Group LEIDYKLA (41,7%)

17 pav. Auksinių klientų segmentavimas (silueto plotis = 0,5)

Sidabrinę grupę sudaro 519 klientų, kurių RFMT kintamųjų vertės taip pat stipriai svyruoja, kaip ir tarp auksinės grupės klientų (žr. 3 priedą), taip pat visi kiti rodikliai yra žemesni nei auksinės grupės atveju, labai stipriai skiriasi dažnumo vidurkis (žr. 13 lentelę).

13 lentelė. Sidabrinės klientų grupės RFMT kintamųjų aprašomoji statistika

	R	F	M	T
Imties dydis	519	519	519	519
Vidurkis	637,04	7,36	1943,26	1227,83
Mediana	499	4	1356,91	1406
Standartinis nuokrypis	532,49	9,28	1781,99	521,82
Imties intervalas	1819	60	13162,84	1799
Minimumas	3	1	63,05	25
Maksimumas	1822	61	13225,89	1824

Šiame klasterizavime visi kintamieji yra svarbūs, nes remiantis visais keturiais kintamaisiais dėl didelio skirtumo tarp reikšmių išskiriamos aiškiai apibrėžtos grupės. Atsižvelgiant į klientų skirtumus sudarytas modelis, kuris įvertintas 0,5 ir išskirti keturi segmentai (žr. 18 pav.):

1. vertingi nauji klientai – tai pirkėjai, kurie prieš pora metų tapo klientais, bet jie retai atlieka užsakymus, tačiau vidutinė užsakymo suma aukšta, juos reikėtų paskatinti dažniau atlikti užsakymus;
2. lojalūs mažos vertės klientai – tai seni klientai, atlikę daug užsakymų, tačiau jie yra labai mažos vertės, tokiems klientams vertėtų pasiūlyti didesnės paslaugų;
3. vertingi pamiršti klientai – tai pirkėjai, kurių užsakymų suma yra aukšta, o paskutinis užsakymas atliktas prieš daugiau nei dvejus metus, tokiems klientams reikėtų priminti apie įmonės teikiamas paslaugas;
4. didelės vertės pamiršti klientai – pirkėjai, kurių užsakymai labai aukštos vertės, tačiau jie labai trumpai ir seniai buvo įmonės pirkėjai.

Cluster	1	2	3	4
Label	Vertingi nauji klientai	Lojalūs mažos vertės klientai	Vertingi pamiršti klientai	Didelės vertės pamiršti klientai
Size	35,8% (186)	6,2% (32)	40,8% (212)	17,1% (89)
Inputs	Time 617,82	Time 1 581,12	Time 1 592,56	Time 1 506,83
	Frequency 5,01	Frequency 36,03	Frequency 7,40	Frequency 1,89
	Recency 246,52	Recency 119,56	Recency 748,34	Recency 1 374,11
	Monetary 1 709,70	Monetary 376,03	Monetary 1 232,09	Monetary 4 688,87
Evaluation Fields	Customer Group TIESIOGINIS (35,5%)	Customer Group TIESIOGINIS (34,4%)	Customer Group TIESIOGINIS (63,2%)	Customer Group TIESIOGINIS (67,4%)

18 pav. Sidabrinių klientų segmentavimas (silueto plotis = 0,5)

Bronzinės grupės klientų yra daugiausiai – 1015 (žr. 14 lent.), tačiau gauti tik du klasteriai, kurie labiausiai skiriasi tik laiko atžvilgiu (žr. 4 priedą). Modelio kokybė yra tokia pati, kaip ir pirmų dviejų klientų grupių – 0,5.



14 lentelė. Bronzinės klientų grupės RFMT kintamųjų aprašomoji statistika

	R	F	M	T
Imties dydis	1015	1015	1015	1015
Vidurkis	735,76	2,31	735,88	899,25
Mediana	674	1	564,76	842
Standartinis nuokrypis	503,7	3,87	633,31	530,4
Imties intervalas	1821	51	2918,1	1820
Minimumas	3	1	5,56	4
Maksimumas	1824	52	2923,66	1824

Šiai grupei didžiausią įtaką turi laikas, nuo kada pradėjo klientai pirkti, o antroje vietoje naujumas, kada paskutinį kartą atliko pirkimą. Kiti rodikliai neturi didelio skirtumo, nes jie yra panašaus didžio abiejuose segmentuose. Taigi mažiausią piniginę vertę įmonei sudariusios grupės klientai padalinti į du segmentus (žr. 19 pav.):

1. Mažos vertės pamiršti klientai – pirkėjai, kurie atliko nedaug mažos vertės užsakymų ir seniai;
2. Mažos vertės nauji klientai – pirkėjai atlikę mažai, žemos vertės užsakymus paskutinių metų laikotarpyje.

Šiems klientams neverta taikyti rinkodaros strategijos, nes jų perkamoji galia labai silpna.

Cluster	1	2
Label	Mažos vertės pamiršti klientai	Mažos vertės nauji klientai
Size	 60,4% (613)	 39,6% (402)
Inputs	Time 1 243,86	Time 373,75
	Recency 1 018,98	Recency 303,90
	Monetary 620,64	Monetary 911,61
	Frequency 2,81	Frequency 1,55
Evaluation Fields	Customer Group TIESIOGINIS (49,3%)	Customer Group TIESIOGINIS (38,8%)

19 pav. Bronzinių klientų segmentavimas (silueto plotis = 0,5)

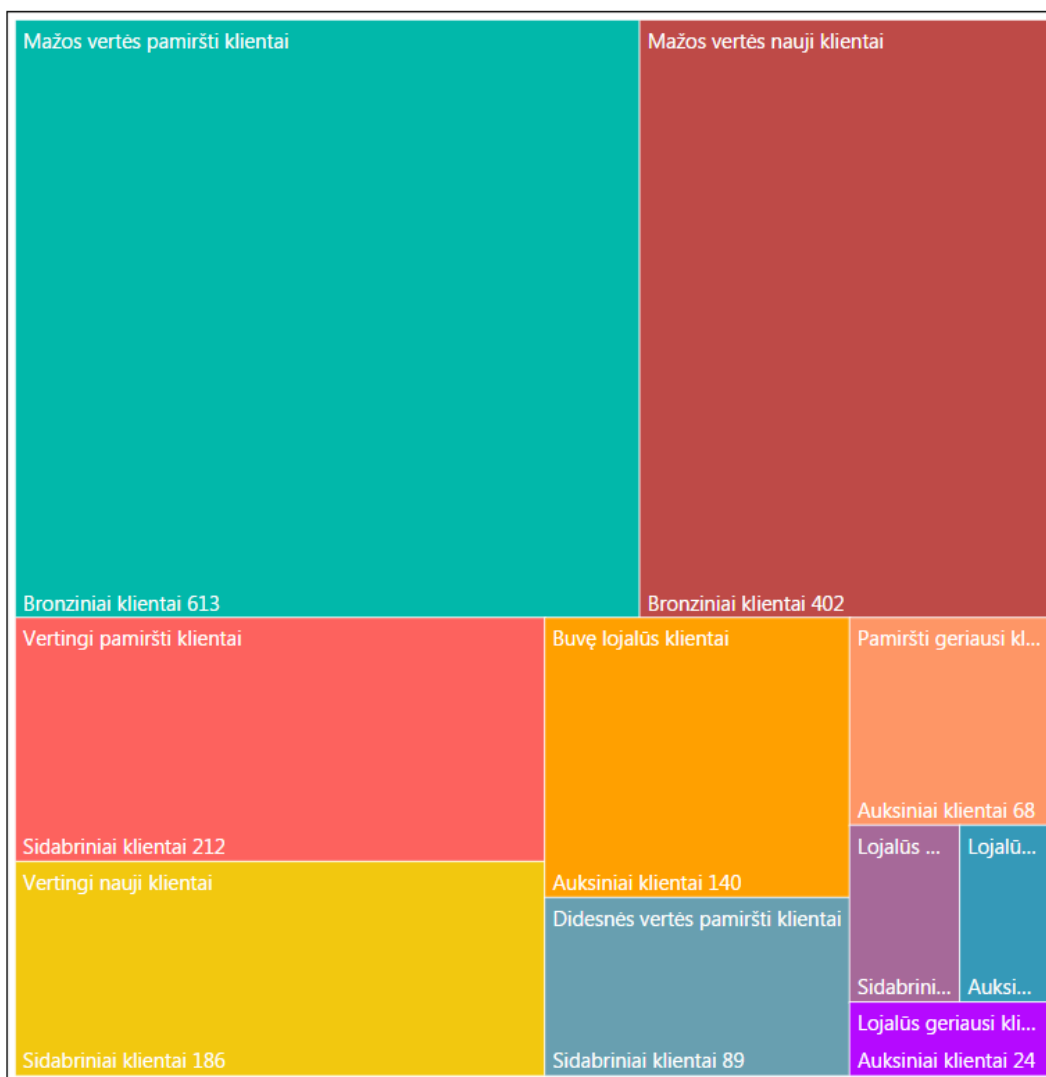
Atlikus atskirus segmentavimus visoms ABC grupėms iš viso išskirta 10 segmentų (žr. 15 lentelę ir 20 pav.). Didžiausius segmentus sudaro mažiausios vertės klientai, kurios priklauso bronzinių klientų grupei. Dabartiniai lojalūs klientai sudaro labai mažą dalį – 4,6%, kuri susideda iš

trijų segmentų. Tarp šių segmentų aptikta daug pamirštų klientų, kuriuos naudinga būtų susigrąžinti įmonei, nes tarp jų yra didelius užsakymus galinčių užsakyti klientų ir buvusių lojalų pirkėjų.

15 lentelė. Visų ABC įmonės klientų segmentų išskaidymas

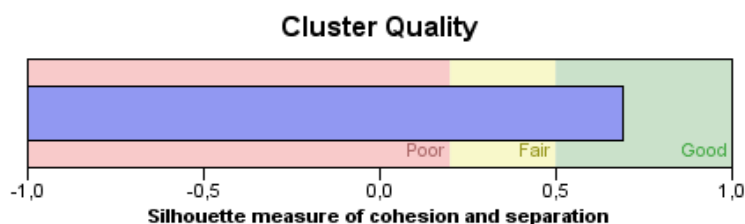
ABC segmentas	Segmento nr.	Segmentas	Klientų sk.	Klientų dalis
Auksiniai klientai (silueto plotis = 0,5)	1	Pamiršti geriausi klientai	68	3,8%
	2	Buvę lojalūs klientai	140	7,8%
	3	Lojalūs klientai	26	1,5%
	4	Lojalūs geriausi klientai	24	1,3%
Sidabriniai klientai (silueto plotis = 0,5)	5	Vertingi nauji klientai	186	10,4%
	6	Lojalūs mažos vertės klientai	32	1,8%
	7	Vertingi pamiršti klientai	212	11,8%
	8	Didesnės vertės pamiršti klientai	89	5,0%
Bronziniai klientai (silueto plotis = 0,5)	9	Mažos vertės pamiršti klientai	613	34,2%
	10	Mažos vertės nauji klientai	402	22,4%

Toliau pateikiama visų dešimties segmentų vizualizacija pagal klientų skaičių (žr. 20 pav.).



20 pav. Visų įmonės klientų segmentų vizualizacija pagal klientų kiekį segmentuose

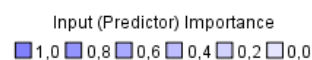
Papildomai atliktas klientų klasterizavimas pagal paslaugų užsakymų populiarumą. Remiantis klientų užsakymų rūšies procentine išraiška buvo gautas modelis, kurio kokybė labai aukšta – 0,7 balo (žr. 21 pav.).



21 pav. Klientų segmentavimo pagal paslaugas modelio kokybė (silueto plotis = 0,7)

Naudojant šešis kintamuosius, gauti šeši klasteriai (žr. 22 pav. ir 5 priedą):

Clusters



Cluster	1	2	3	4	5	6
Label	Simple produktų klientai	Kitų produktų klientai	Kietviršių klientai	Repro paslaugų klientai	Dizaino paslaugų klientai	Brošiūrų klientai
Description	Didelės vertės Lietuvos klientai, agentūros ir tiesioginiai	Lietuvos miestų vidutinės vertės klientai, pagrindiniai - tiesioginiai	Užsienio šalių vidutinės vertės visų grupių klientai	Mažos vertės Kauno miesto klientai	Mažos vertės Lietuvos tiesioginiai klientai	Didelės vertės užsienio šalių visų grupių klientai
Size	25,8% (462)	15,1% (270)	8,5% (152)	3,5% (63)	3,1% (55)	44,1% (791)
Inputs	BROSIUROS, % 7,39 DIZAINAS, % 0,79 KNYG, % 0,73 Product SIMPLE (100,0%) REPRO, % 0,41 SIMPLE, % 83,89	BROSIUROS, % 3,08 DIZAINAS, % 0,34 KNYG, % 2,25 Product KITI (73,0%) REPRO, % 0,00 SIMPLE, % 6,95	BROSIUROS, % 8,56 DIZAINAS, % 0,00 KNYG, % 84,12 Product KNYG (100,0%) REPRO, % 0,33 SIMPLE, % 3,40	BROSIUROS, % 0,07 DIZAINAS, % 0,00 KNYG, % 0,00 Product REPRO (100,0%) REPRO, % 97,07 SIMPLE, % 2,15	BROSIUROS, % 2,84 DIZAINAS, % 81,80 KNYG, % 0,09 Product DIZAINAS (96,4%) REPRO, % 0,00 SIMPLE, % 8,90	BROSIUROS, % 85,08 DIZAINAS, % 0,57 KNYG, % 2,91 Product REPRO, % 0,08 SIMPLE, % 6,04
Evaluation Fields	Group TIESIOGINIS (58,4%) Place Vilnius (43,1%) Segment Bronza (59,5%)	Group TIESIOGINIS (65,2%) Place Vilnius (38,1%) Segment Bronza (50,4%)	Group TIESIOGINIS (34,2%) Place Kitos šalys (40,1%) Segment Bronza (59,2%)	Group Place Kaunas (84,1%) Segment Bronza (66,7%)	Group TIESIOGINIS (74,5%) Place Vilnius (56,4%) Segment Bronza (78,2%)	Group TIESIOGINIS (40,1%) Place Kitos šalys (30,6%) Segment Bronza (54,2%)

22 pav. Klientų segmentai pagal paslaugų užsakymo dažnumą (silueto plotis = 0,7)

1. „simple“ produktų klientai – tai klientai, kurie dažniausiai užsako skrajutes, plakatus, lankstinukus ar kitus mažos vertės gaminius, taip pat didelė dalis klientų linkę kartais užsakyti ir kitos rūšies paslaugas. Daugiausia tokių klientų yra Lietuvos Vilniaus miesto gyventojai ir jie priklauso bronziniam segmentui;
2. kitų produktų klientai – tai nepopuliarių produktų grupių klientai, kurių dauguma yra tiesioginiai pirkėjai iš Vilniaus;
3. kietviršių klientai – tai knygų kietais viršeliais pirkėjai, kurie sudaro mažą dalį visų klientų, tik 8,5%, nes šių produktų vertė yra labai aukšta, tokių pirkėjų didelė dalis yra iš kitų šalių;
4. repro paslaugų (pateiktas sutrumpinimas prieš tai minėtų reprocentro paslaugų) klientai – didžioji dalis klientų yra Kauno miesto gyventojai, kuriuos domina tik viena rūšis įmonės teikiamų paslaugų;
5. dizaino paslaugų klientai – tai mažiausia klientų grupė, kurios 56% yra iš Vilniaus ir dauguma šių klientų priklauso bronzinei grupei – 78%, labai maža dalis šios paslaugos klientų linkę užsakyti kartu ir kitos rūšies gaminių;
6. brošiūrų klientai – šis segmentas yra didžiausias, jį sudaro 44% visų įmonės klientų, grupės pagrindą sudaro užsienio šalių klientai, kurie linkę užsakyti ir kitas įmonės teikiamas paslaugas.

Kaip matoma iš segmentavimo spaustuvės dauguma klientų yra brošiūrų bei paprastų produktų pirkėjai. Mažiausią dalį sudaro reprocentro ir dizaino paslaugų pirkėjai, kurių dauguma priklauso mažos vertės bronziniam klientų segmentui.

Pagal šį segmentavimą remiantis žinomais kriterijais apie klientą: gyvenamoji vieta, kliento grupė bei pagal vidutinę išleidžiamą piniginę sumą galima pasiūlyti papildomus produktus iš tam tikros grupės jau esamiems ir naujiems klientams.

3.3. Klientų lojalumo detekcija

Paskutinis poskyris skirtas atlikti įmonės lojalių ir nelojalių klientų detekciją. Detekcijos metu randamas tiksliausias modelis įmonės klientų lojalumo nustatymui.

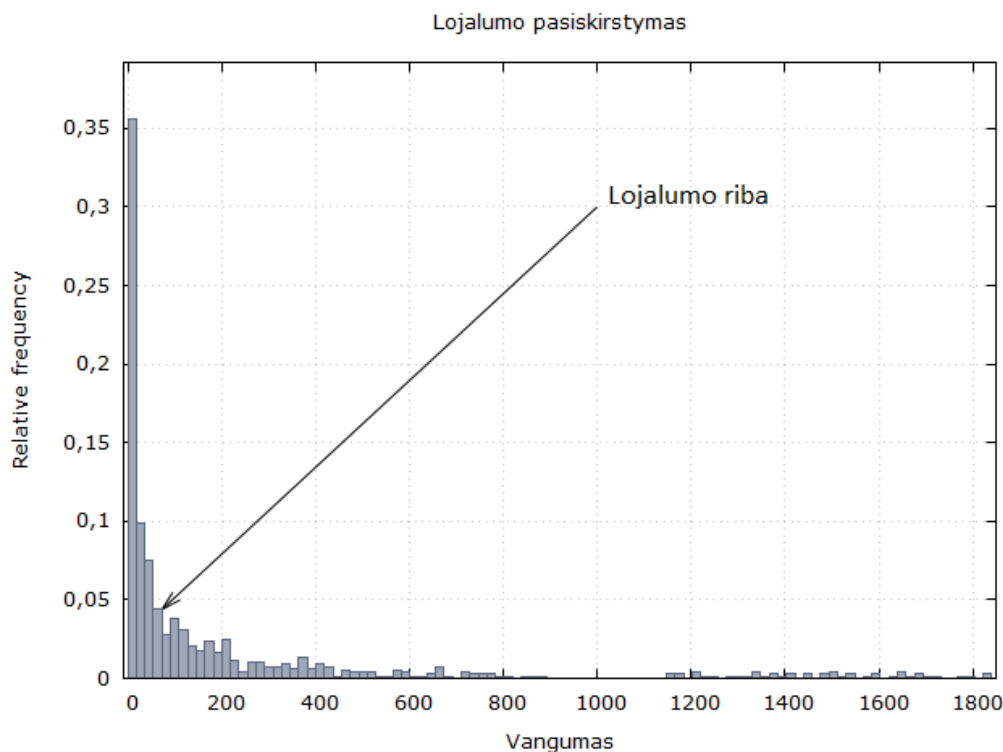
Lojalumo detekcija atliekama tarp vertingiausių įmonės klientų, nes jie yra svarbiausi, kurie išskirti po ABC analizės ir priklauso auksiniam ir sidabriniam segmentams. Šiuos du segmentus sudaro 777 įmonės lojalūs ir nelojalūs klientai, kuriuos reikia išskirti. Lojalumas randamas naudojant vangumo formulę:

$$Vangumas = \frac{Laikas}{Dažnumas}$$

Laiko ir dažnumo santykis parodo vangumą, kaip retai klientas atlieka užsakymus. Siekiant gauti panašaus dydžio imtis tarp lojalių ir nelojalių pirkėjų atlikta dažnių pasiskirstymo analizė, kurios metu atskirtos dvi klientų grupės:

1. lojalūs klientai sudaro 57,40% visų klientų, kurių vangumas mažesnis nei 72,88 dienos;
2. nelojalūs klientai sudaro 42,60%, kurių vangumas lygus arba didesnis nei 72,88 dienos.

Šis lojalių ir nelojalių klientų pasiskirstymas atsižvelgiant į vangumą matomas 23 paveiksle.



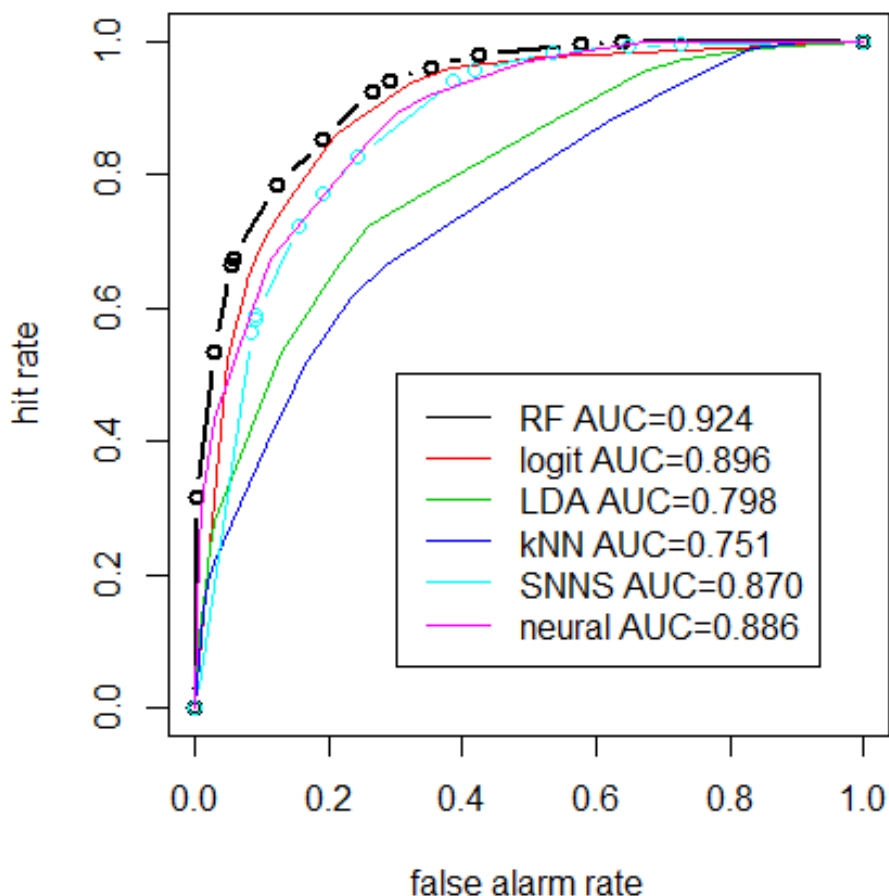
23 pav. Lojalumo atskyrimo riba

Apibrėžus lojalius ir nelojalius klientus, toliau atliekama stratifikuota kryžminė patikra su R statistinių duomenų apdorojimui skirta programa. Naudojami klasifikavimo algoritmai:

1. RF – atsitiktinis miškas;
2. logit – logistinė regresija;
3. LDA – tiesinė diskriminantinė analizė;
4. kNN – k-artimiausių kaimynų klasifikatorius;
5. SNNS – Štutgarto neuroninis tinklas;
6. neural – pagrindinis neuroninis tinklas R kalboje.

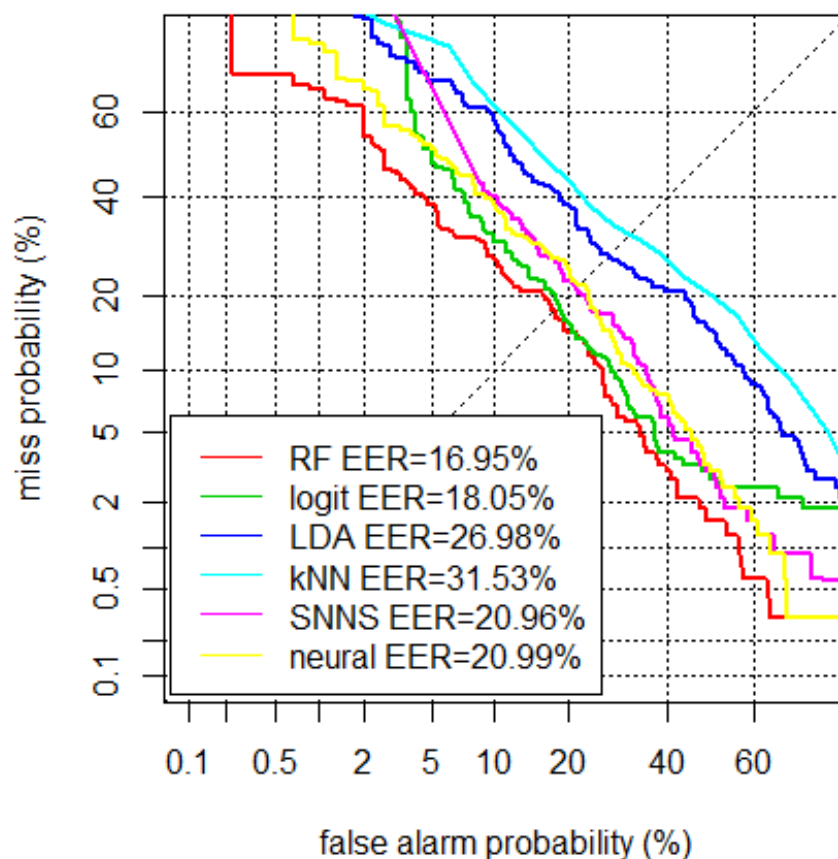
Naudota 10 blokų stratifikuota kryžminė patikra. Atsitiktinių miškų algoritmui naudojama 1000 medžių. Taikant k-artimiausių kaimynų algoritmą numatytas kaimynų skaičius 55. Neuroniniams tinklams naudota 30 vienetų paslėptame sluoksnyje. Iš viso buvo naudojami 57 lojalumo detekcijos

požymiai (žr. 6 priede), į kuriuos patenka klientų charakteristika apie gyvenamąją vietą, pirkėjo grupę, perkamiausius produktus bei užsakymų vertę. Atmesti visi požymiai susiję su laiku, nubyrėjimo kriterijus paremtas laiku. Kaip matoma tolimesniame grafike (žr. 24 pav.) tiksliausias modelis gautas taikant atsitiktinių miškų algoritimą, kuriuo atveju AUC rodiklis yra geriausias 0,924, kuomet kiti algoritmai nesiekia 0,9. Šiuo atveju prasčiausias AUC rezultatas gautas k-artimiausių kaimynų atveju, tik 0,751.



24 pav. ROC tikslumo kreivės

Atitinkamai klaidų dažnis EER atsitiktinių miškų yra mažiausias, tik 16,95%, o k-artimiausių kaimynų didžiausias 31,53% (žr. 25 pav.).



25 pav. DET klaidų dažnio kreivės

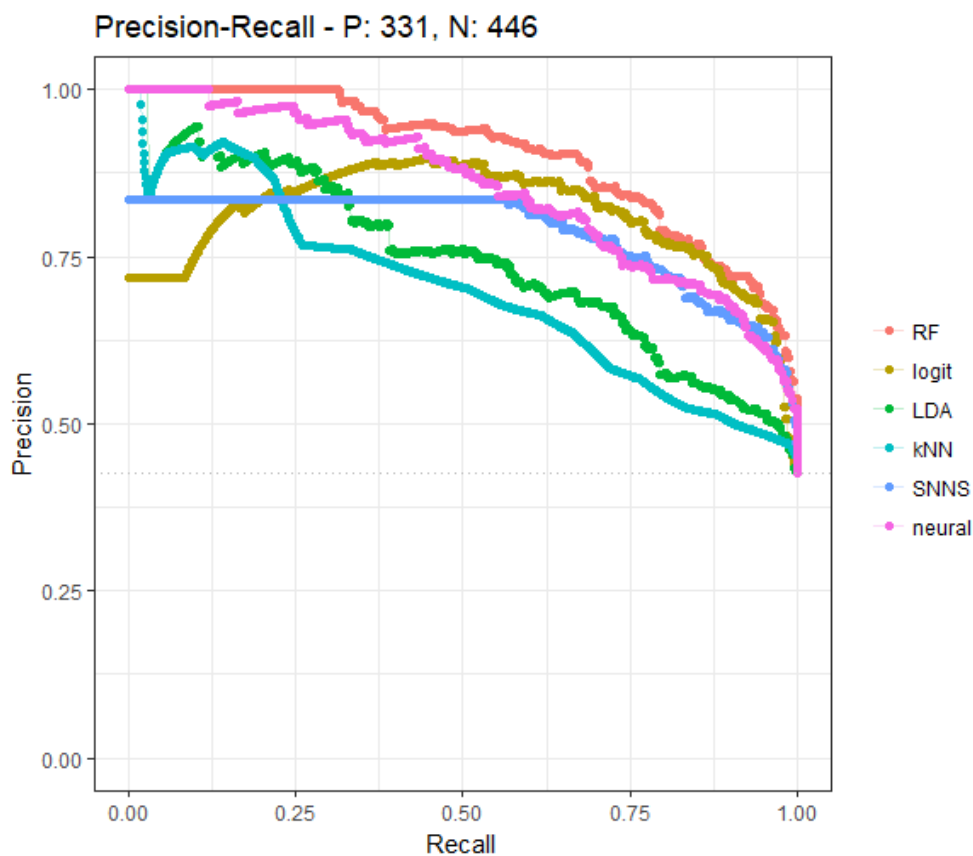
Taigi, nubyrėjimo pagal vangumo rodiklį detekcija kokybiškiausiai realizuota, naudojant atsitiktinių miškų modelį. Pagal visų šešių modelių kokybę, jie yra išsidėstę tokia tvarka:

1. atsitiktiniai miškai;
2. logistinė regresija;
3. neuroniniai tinklai;
4. Štutgarto neuroniniai tinklai;
5. diskriminantinė analizė;
6. k-artimiausių kaimynų.

Antroje vietoje yra logistinės regresijos detektorius su $AUC = 0,896$, bet lyginant šį metodą su atsitiktiniais miškais, kurio $AUC = 0,924$, pagal priklausomų imčių Stjudento t-kriterijaus testą AUC įverčiai reikšmingai skiriasi, nes p -reikšmė (angl. p -value) yra 0,0381, t.y. mažiau nei 0,05.

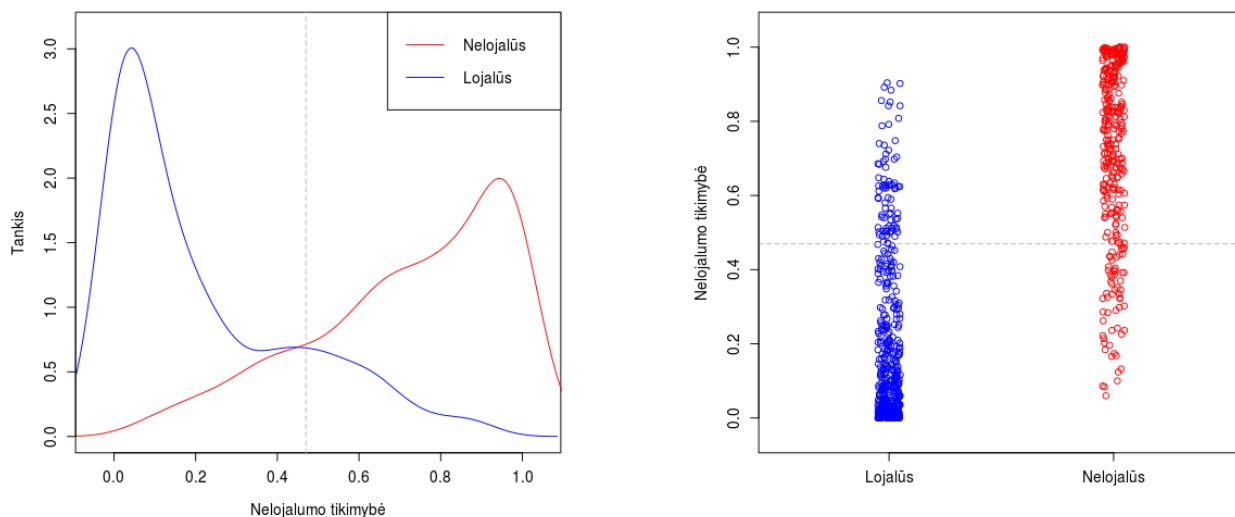
Toliau pateikiamas preciziškumo – prisiminimo grafikas (žr. 26 pav.), atsitiktinių miškų kreivė randasi aukščiausiai, tai reiškia, kad šis detektorius yra geriausias. Šiuo atveju nubyrėjimo preciziškumas 78%, kas reiškia, kad iš visų algoritmo suklasifikuotų nelojalių klientų 78% pavyko nurodyti teisingai. Prisiminimas lygus 82, 48% rodo, kokią dalį iš faktinių nelojalių klientų pavyko suklasifikuoti teisingai, kada ribinė vertė yra 0,47. Kreivėse matosi, kaip kiekvieno detektoriaus

atveju, gerėjant prisiminimo rodikliui, kistų preciziškumas – dažniausiai jis mažėtų, nes algoritmas daugiau duomenų priskirtų į nelojalių klasę neteisingai.



26 pav. Apibendrinti rezultatai

Toliau pateikiami tik atsitiktinių miškų algoritmo rezultatai. Sekančiame 27 paveiksle matomas lojalių (pažymėti mėlyna spalva) ir nelojalių (pažymėti raudona spalva) klientų pasiskirstymas. Grafikuose taip pat išskirtas ribinė vertė, kuri yra 0,47, ši riba atskiria lojalius nuo nelojalių. Taigi iš grafikų matosi, kad dauguma klientų pagal vangumą buvo priskirti teisingai grupei.



27 pav. Suprognuoto lojalumo iš atsitiktinio miško detektoriaus pasiskirstymas

Toliau sumaišymo matricioje (žr. 28 pav.) pateikiama detalesnė informacija apie lojalumo pasiskirstymą. Matoma, kad iš 446 lojaliais nustatytų klientų 369 buvo teisingai nustatyti, o iš 331 nelojaliais nustatytų 273 buvo teisingai nustatyti klientai.

		Truth data			Producer Accuracy (Precision)
		Class 1	Class 2	Classification overall	
Classifier results	Class 1	369	58	427	86.417%
	Class 2	77	273	350	78%
	Truth overall	446	331	777	
	User Accuracy (Recall)	82.735%	82.477%		
Overall accuracy (OA):		82.625%			
Kappa ¹ :		0.647			

28 pav. Sumaišymų matrica

Didesnė tikimybė yra teisingai suklasifikuoti lojalius klientus 82,735% nei nelojalius, kuriuos tikimybė atspėti yra 82,477%. Detektoriumi teigiant, kad klientas yra lojalus, tikimybė, kad tai tiesa 86,417%, didesnė tikimybė nei nelojalių klientų atveju – 78%. Jautrumo ir specifiškumo dydžiai yra panašūs, todėl ir bendras klasifikavimo tikslumo – OA rodiklis yra panašaus dydžio 82,63%. Kappa statistinis rodiklis nurodo atsitiktinio klasifikavimo patikimumą, kuris lygus 0,647, šis dydis yra tarp 0,61 ir 0,8, todėl vertinamas kaip labai aukštas patikimumas.

Taigi remiantis atlikta detekcija pastebėta, kad atsitiktinių miškų detektorius geriausiai tinka nubyrėjimo prognozavimui, remiantis turima informacija apie klientus dėl aukšto AUC, žemo EER ir aukšto Kappa rodiklių.

Išvados

1. Klientų segmentavimas svarbus rinkodaros tikslais, siekiant pažinti ir suprasti klientus bei pritaikyti diferencijuotas strategijas atskiriems klientų segmentams. Svarbu laiku nustatyti klientų pasitraukimą, jį prognozuoti ir adekvačiai reaguoti. Šių tikslų įgyvendinimo strategijoms gali būti taikoma statistinė analizė, klasterizavimo metodai bei mašininio mokymo algoritmai.
2. Spaudos įmonės duomenims naudojama ABC kartu su RFMT klientų analize, kurių rodiklių pagalba atliktas dviejų žingsnių klasterizavimas. Lojalumo patikrai lyginami detektoriai: logistinė regresija, diskriminantinė analizė, neuroniniai tinklai, atsitiktiniai miškai ir k-artimiausių kaimynų, iš kurių atrenkamas tiksliausių rezultatą generuojantis algoritmas.
3. Poligrafijos įmonės klientai yra organizacijos, daugiausiai agentūros ir tiesioginiai klientai, dalis jų iš Lietuvos didžiųjų miestų Vilniaus bei Kauno, taip pat kitų šalių, tokių kaip Vokietijos, Olandijos, Prancūzijos ar Jungtinės Karalystės. Įmonės klientai yra įmonės iš įvairių šalių, daugiausiai Lietuvos, Vilniaus ir Kauno miesto, Vokietijos ir Olandijos įvairių sričių bendrovės, daugiausiai agentūros ir tiesioginiai klientai, kurie įmonėje užsako ne tik produktus, bet ir paslaugas. Naudojant ABC skaidymą dėl tikslesnio segmentavimo ir RFMT atliktas dviejų žingsnių klasterizavimas, kurio metu išskirta dešimt klasterių, kuriuose išsiskiria pamiršti vertingi bei lojalūs klientai, kuriuos reikėtų susigrąžinti, taip pat išryškėja dabartiniai klientai, kurių lojalumą toliau vertėtų stengtis išlaikyti.
4. Po ABC analizės atmetus C segmentą, palikti tik vertingiausi įmonės A ir B segmentų klientai, kuriems spręstas lojalumo detekcijos uždavinys. Nelojaliais nustatyti tie, kurių vidutiniška trukmė tarp gretimų užsakymų ~3 ir daugiau mėnesių. Geriausi rezultatai gauti naudojant atsitiktinių miškų klasifikavimo algoritmą, kurio AUC = 0,92, Kappa = 0,65 ir EER = 16,95%.
5. Kiekvienai įmonei periodiškai reikia atlikti klientų segmentavimą ir lojalumo įvertinimą ir stebėti ar prognozuojamas lojalumo įvertis nepakito bendrovei nepalankia linkme. Analizuojamos įmonės atveju, derėtų daugiau dėmesio skirti blėstantiems ar jau prarastiems klientams susigrąžinti, kurie turi potencialą sukurti didelę vertę – tokie klientai sudaro net 28% visų įmonės klientų. Be to įmonėje didelę dalį – 22% klientų sudaro nauji mažos vertės pirkėjai, kuriuos reikėtų skatinti įsigyti pakartotinai didesnės vertės gaminių. Daugiau nei pusė vertingoms grupėms priklausančių klientų įvardijami kaip lojalūs, tokius klientus reikėtų puoselėti naudojant atitinkamas rinkodaros strategijas.
6. Tyrime naudoti metodai patogūs, tačiau RFM analizė netinka šios įmonės klientų segmentavimui, todėl pasirinktas RFMT metodas. Dviejų žingsnių analizė efektyviai veikia naudojant tik skaitinius kintamuosius, metodo trūkumas, kad jis jautrus kategorinių kintamųjų rinkiniui, todėl spaudos įmonės atveju siūloma vengti kategorinių kintamųjų.

Literatūros sąrašas

- Auruškevičienė, V., & Šalčiuvienė, L. (2003). Paslaugų įmonių klientų lojalumo dimensijų tyrimas. *Transformacijos Rytų Ir Centrinėje Europoje : žurnalo „Tiltai“ Priedas : Mokslo Darbai, 1*(13), 18–23.
- Bafadal, A. K. (2011). Two Step Cluster Analysis.
- Bakanauskas, A., & Pilelienė, L. (2009). *Vartotojų lojalumas: Teoriniai ir praktiniai aspektai : Mokomoji knyga*. Kaunas: Vytauto Didžiojo universitetas.
- Barladean, V. (2017). 7 Key Steps to Reduce Customer Churn Rate. Gauta <https://www.mycustomer.com/community/blogs/cloudsoft1/7-key-steps-to-reduce-customer-churn-rate>
- Ben-David, A. (2008). About the relationship between ROC curves and Cohen's kappa. *Engineering Applications of Artificial Intelligence, 21*(6), 874–882.
- Birant, D. (2016). *Knowledge-Oriented Applications in Data Mining. Data Mining Using RFM Analysis*. InTech.
- Chiang, W. Y. (2011). To mine association rules of customer values via a data mining procedure with improved model: An empirical case study. *Expert Systems with Applications, 38*(3), 1716–1722. <https://doi.org/10.1016/j.eswa.2010.07.097>
- Chu, C. W., Liang, G. S., & Liao, C. T. (2008). Controlling inventory by combining ABC analysis and fuzzy classification. *Computers and Industrial Engineering, 55*(4), 841–851.
- Decision Support Systems, LP. (2018). Gauta 2018 m. kovo 7 d., <https://www.dssresearch.com/Solutions/StrategyResearchSolutionsGroup/MarketSegmentation/MarketSegmentationTechniques.aspx>
- Dehghan, A., & Trafalis, T. (2012). Examining Churn and Loyalty Using Support Vector Machin. *Business and Management Research, 1*(4), 153–161.
- Dursun, A., & Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism Management Perspectives, 18*, 153–160.
- Ekelund, S. (2017). Precision-recall curves – what are they and how are they used? Gauta <https://acutecaretesting.org/en/articles/precision-recall-curves-what-are-they-and-how-are-they-used>
- Fawcett, T. (2004). ROC Graphs : Notes and Practical Considerations for Researchers. *ReCALL, 31*(HPL-2003-4), 1–38.
- Gajdhar, S. (2017). How To Use RFM Analysis & Segmentation To Improve Customer Lifecycle Marketing. Gauta 2018 m. kovo 27 d., <https://monk.webengage.com/how-to-use-rfm-analysis-segmentation-to-improve-marketing/>
- Glady, N., Baesens, B., & Croux, C. (2009). Modeling churn using customer lifetime value. *European Journal of Operational Research, 197*(1), 402–411.
- Glinskienė, R., Kvedaraitė, N., & Kvedaras, M. (2010). Vartotojų lojalumo stiprinimas - reikšmingiausia šiuolai- kinių įmonių siekiamybė. *Ekonomika ir vadyba: aktualijos ir perspektyvos, 4*(20), 40–52.
- Hwang, H., Jung, T., & Suh, E. (2004). An LTV model and customer segmentation based on customer value: A case study on the wireless telecommunication industry. *Expert Systems with Applications, 26*(2), 181–188.
- Janita, M. S., & Miranda, F. J. (2013). The antecedents of client loyalty in business-to-business (B2B) electronic marketplaces. *Industrial Marketing Management, 42*(5), 814–823. Gauta <http://dx.doi.org/10.1016/j.indmarman.2013.01.006>
- Kadiyala, S. S., & Srivastava, A. (2011). Data Mining For Customer Relationship Management. *International Business & Economics Research Journal (IBER), 1*(6), 61–70.
- Kumar, V., & Shah, D. (2004). Building and sustaining profitable customer loyalty for the 21st century. *Journal of Retailing, 80*(4), 317–330.
- Lemmens, A., & Gupta, S. (2013). *Managing Churn to Maximize Profits*. SSRN Electronic Journal.
- Levin, N., & Zahavi, J. (2001). Predictive modeling using segmentation. *Journal of Interactive Marketing, 15*(2), 2–22.
- Li, Y., & Li, R. (2011). *Research on Risk Aversion of Human Resources Outsourcing*. International Symposium Advances in Applied Economics, Business and Development.
- Magenta. (2013). Gauta 2018 m. kovo 6 d., http://wwwowwww.co.za/wp-content/uploads/2013/09/An_Introduction_to_Customer_Segmentation_ebook.pdf
- MarketingSolutions.lt. (2015a). Cross selling. Gauta <http://www.marketingsolutions.lt/cross-selling>
- MarketingSolutions.lt. (2015b). Down selling. Gauta <http://www.marketingsolutions.lt/down-selling>
- MarketingSolutions.lt. (2015c). Up selling. Gauta <http://www.marketingsolutions.lt/up-selling>
- Maroco, J., Silva, D., Rodrigues, A., Guerreiro, M., Santana, I., & De Mendonça, A. (2011). Data mining methods in the prediction of Dementia: A real-data comparison of the accuracy, sensitivity and specificity of linear discriminant analysis, logistic regression, neural networks, support vector machines, classification trees and random forests.

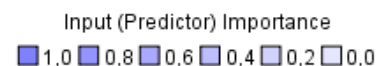
- Mchugh, M. L. (2012). Lessons in biostatistics Interrater reliability : the kappa statistic. *Biochemia Medica*, 22(3), 276–282.
- Milošević, M., Živić, N., & Andjelković, I. (2017). Early churn prediction with personalized targeting in mobile social games. *Expert Systems with Applications*, 83, 326–332.
- Mudry, A., & Tjellström, A. (2011). Historical background of bone conduction hearing devices and bone conduction hearing aids. *Advances in Oto-Rhino-Laryngology*, 71, 1–9.
- Mutanen, T. (2006). *Customer churn analysis—a case study*. Technical Research Centre of Finland (VTT).
- Nabareseh, S. (2017). *Predictive analytics : a data mining technique in customer churn management for decision making*. Tomas Bata University.
- Paulauskienė, K., & Kurasova, O. (2013). Duomenų tyrybos sistemų galimybių tyrimas įvairių apimčių duomenims analizuoti. *Informacijos mokslai*, 65(1392–0561), 85–95.
- Probst, P., & Boulesteix, A.-L. (2017). To tune or not to tune the number of trees in random forest?, (2001), 1–20.
- Rapečka, A. (2015). Rekomendacinių sistemų socialiniuose tinkluose efektyvumo didinimas.
- Raulerson, P., Malraison, J.-C., & Leboyer, A. (2009). *Building Routes to Customers*. New York: Springer.
- Risselada, H., Verhoef, P. C., & Bijmolt, T. H. A. (2010). Staying Power of Churn Prediction Models. *Journal of Interactive Marketing*, 24(3), 198–208.
- Sáenz-Lechón, N., Godino-Llorente, J. I., Osma-Ruiz, V., & Gómez-Vilda, P. (2006). Methodological issues in the development of automatic systems for voice pathology detection. *Biomedical Signal Processing and Control*, 1(2), 120–128.
- SPSS. (2001). The SPSS TwoStep Cluster Component: A scalable component enabling more efficient customer segmentation, 9.
- Sutevski, D. (2018). Improve Your Business Performances with ABC Analysis or 80/20 Rule. Gauta 2018 m. kovo 21 d., <http://www.entrepreneurshipinbox.com/147/abc-analysis-or-8020-rule/>
- Thomas, J. W. (2016). Decision Analyst, Inc. Gauta 2018 m. kovo 6 d., <https://www.decisionanalyst.com/whitepapers/marketsegmentation/>
- Tolay, A. (2018). What is Precision-Recall (PR) curve? Gauta <https://www.quora.com/What-is-Precision-Recall-PR-curve>
- Tsiptsis, K., & Chorianopoulos, A. (2010). *Data Mining Tehniques in CRM: Inside Customer Segmentation*. Mining, 2002. ICDM 2002. Wiley Publishing.
- Urbonavičius, S., & Ivanauskas, R. (2006). Parduotuvių tinklų pirkėjų segmentavimas pagal pirkimo progas. *Verslas: teorija ir praktika*, 7(1), 37–44.
- Vaitkevičius, R. (2012). SPSS programa: klasterinė analizė, 1–6.
- Wei, J. T., Lin, S. Y., Weng, C. C., & Wu, H. H. (2012). A case study of applying LRFM model in market segmentation of a children's dental clinic. *Expert Systems with Applications*, 39(5), 5529–5533.
- Westreich, D., Lessler, J., & Funk, M. J. (2010). Propensity score estimation: neural networks, support vector machines, decision trees (CART), and meta-classifiers as alternatives to logistic regression. *Journal of Clinical Epidemiology*, 63(8), 826–833.
- Willan, D. (2014). B2B Market Segmentation. Gauta 2018 m. kovo 14 d., <https://www.circle-research.com/report/b2b-market-segmentation/>
- Wollen, R. (2017). *A Modern Approach to RFM Segmentation*.
- Xie, Y., Li, X., Ngai, E. W. T., & Ying, W. (2009). Customer churn prediction using improved balanced random forests. *Expert Systems with Applications*, 36(3 PART 1), 5445–5449.
- Ziafat, H., & Shakeri, M. (2014). Using Data Mining Techniques in Customer Segmentation. *International Journal of Engineering Research and Applications*, 4(9), 70–79.

Priedai

1 priedas

ABC klientų pagal RFMT modelį dviejų žingsnių segmentavimas

Clusters



Cluster	1	2	3	4	5
Label	Geriausi lojalūs klientai	Pelningiausi blėstantys klientai	Pelningiausi prarasti klientai	Nauji klientai	Mažos vertės prarasti klientai
Description	Klientai dažnai užsakantys didelės vertės užsakymus	Klientai retai darantys užsakymus dėl mažo poreikio, tačiau pastovūs klientai	Pamiršti klientai, kurių užsakymų vidutinė vertė didelė, tačiau seniai atliko paskutinį užsakymą	Nauji klientai, kurie neseniai pradėjo naudotis įmonės paslaugomis	Labai mažos vertės klientai, kurie tik išbandė įmonės teikiamas paslaugas
Size	2.3% (41)	21.3% (381)	24.7% (442)	25.4% (456)	26.3% (472)
Inputs	Frequency	Frequency	Frequency	Frequency	Frequency
	Recency	Recency	Recency	Recency	Recency
	Time	Time	Time	Time	Time
	Monetary	Monetary	Monetary	Monetary	Monetary
Evaluation Fields	Segment	Segment	Segment	Segment	Segment
	Customer Group	Customer Group	Customer Group	Customer Group	Customer Group

Auksinio klientų segmento pagal RFMT modelį dviejų žingsnių klasterizavimas

Clusters

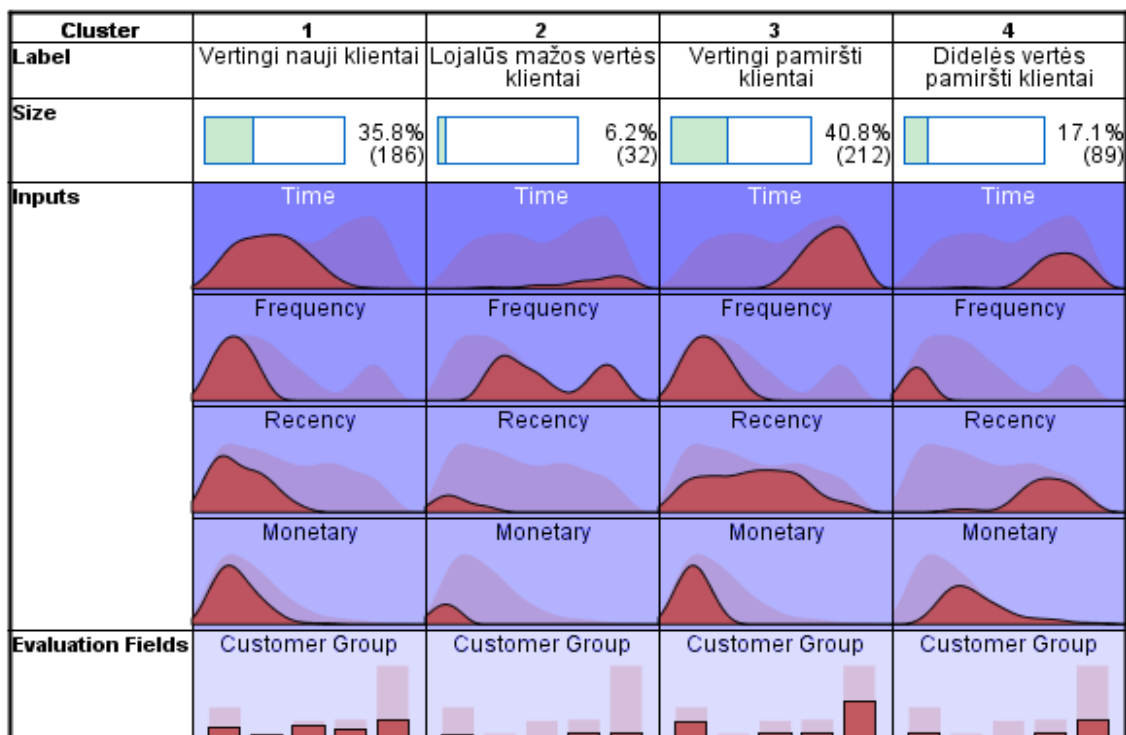
Input (Predictor) Importance
 ■ 1,0 ■ 0,8 ■ 0,6 ■ 0,4 ■ 0,2 ■ 0,0



Sidabrinio klientų segmento pagal RFMT modelį dviejų žingsnių klasterizavimas

Clusters

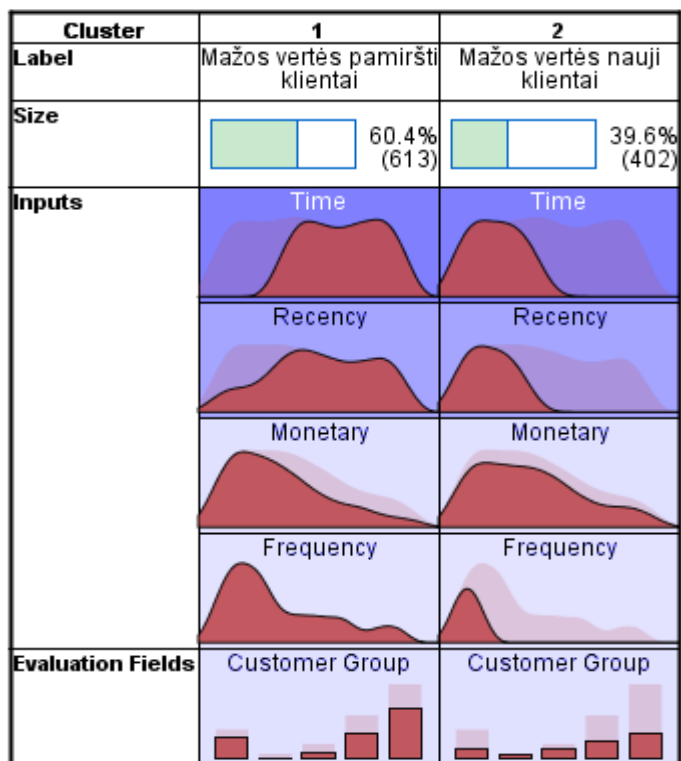
Input (Predictor) Importance
 ■ 1,0 ■ 0,8 ■ 0,6 ■ 0,4 ■ 0,2 ■ 0,0



Bronzinio klientų segmento pagal RFMT modelį dviejų žingsnių klasterizavimas

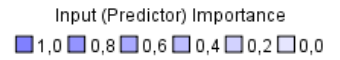
Clusters

Input (Predictor) Importance


 1,0 0,8 0,6 0,4 0,2 0,0


Visų klientų dviejų žingsnių klasterizavimas atsižvelgiant į paslaugų grupių perkamumą

Clusters



Cluster	1	2	3	4	5	6
Label	Simple produktų klientai	Kitų produktų klientai	Kietviršių klientai	Repro paslaugų klientai	Dizaino paslaugų klientai	Brošiūrų klientai
Description	Didelės vertės Lietuvos klientai, agentūros ir tiesioginiai	Lietuvos miestų vidutinės vertės klientai, pagrindiniai - tiesioginiai	Užsienio šalių vidutinės vertės visų grupių klientai	Mažos vertės Kauno miesto klientai	Mažos vertės Lietuvos tiesioginiai klientai	Didelės vertės užsienio šalių visų grupių klientai
Size	25.8% (462)	15.1% (270)	8.5% (152)	3.5% (63)	3.1% (55)	44.1% (791)
Inputs						
Evaluation Fields						

Požymiai lojalumo detekcijai

Nr.	Pavadinimas	Reikšmė
1	Churn	Nubyrėjimas
2	CountofID	Užsakymų skaičius
3	AvgQuotedPrice	Vidutinė užsakymo kaina
4	AvgUnitPrice	Vidutinė vieneto kaina
5	AvgNetProfitAmount	Vidutinis grynasis pelnas
6	AvgMaterials	Vidutinė medžiagų savikaina
7	AvgPriceCalculated	Vidutinė paskaičiuota kaina
8	Profit	Pelnas
9	AvgProfit	Vidutinis pelnas
10	AvgProfitHour	Vidutinis pelnas per valandą
11	Discount	Nuolaida
12	AvgDiscount	Vidutinė nuolaida užsakymui
13	AvgQty	Vidutinis produktų kiekis užsakyme
14	ROI	Investicijų pelningumo rodiklis
15	AGENTURA	Agentūros grupės klientai
16	LEIDYKLA	Leidyklos grupės klientai
17	KITOS	Nepopuliarių grupių klientai
18	NEPRISKIRTA	Klientai su nenurodyta veiklos grupė
19	Uzsienis	Užsienio klientai
20	DE	Vokietijos klientai
21	FR	Prancūzijos klientai
22	NL	Olandijos klientai
23	UK	Jungtinės Karalystės klientai
24	Kaunas	Kauno miesto klientai
25	KitiLTmiestai	Kitų Lietuvos miestų klientai
26	KitosSalys	Kitų užsienio šalių klientai
27	ES	Europos Sąjungos klientai
28	NEES	Ne Europos Sąjungos klientai
29	BROSI	Siūtų brošiūrų klientai
30	BROKL	Klijuotų brošiūrų klientai
31	BROSEG	Segtų brošiūrų klientai
32	DIZAINAS	Dizaino paslaugų klientai
33	KALEND	Kalendorių klientai
34	KITI	Kitų nepopuliarių produktų klientai
35	KNYG	Knygų kietais viršeliais klientai
36	REPRO	Reprocentro paslaugų klientai
37	SIMPLE	Mažos vertės produktų klientai
38	CountBROKL	Klijuotų brošiūrų užsakymų skaičius
39	BROKLproc	Klijuotų brošiūrų užsakymų procentas
40	CountBROSEG	Segtų brošiūrų užsakymų skaičius
41	BROSEGproc	Segtų brošiūrų užsakymų procentas
42	CountBROSI	Siūtų brošiūrų užsakymų skaičius
43	BROSIproc	Siūtų brošiūrų užsakymų procentas

Nr.	Pavadinimas	Reikšmė
44	CountDIZAINAS	Dizaino paslaugų užsakymų skaičius
45	DIZAINASproc	Dizaino paslaugų užsakymų procentas
46	CountKALEND	Kalendorių užsakymų skaičius
47	KALENDproc	Kalendorių užsakymų procentas
48	CountKITI	Kitų produktų užsakymų skaičius
49	KITIproc	Kitų produktų užsakymų procentas
50	CountKNYG	Knygų kietais viršeliais užsakymų skaičius
51	KNYGproc	Knygų kietais viršeliais užsakymų procentas
52	CountREPRO	Reprocentro paslaugų užsakymų skaičius
53	REPROproc	Reprocentro paslaugų užsakymų procentas
54	CountSIMPLE	Mažos vertės produktų užsakymų skaičius
55	SIMPLEproc	Mažos vertės produktų užsakymų procentas
56	CountBROSIUROS	Brošiūrų užsakymų skaičius
57	BROSIUROSproc	Brošiūrų užsakymų procentas