

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS

INFORMATIKOS KATEDRA

Verslo informacinių sistemų studijų programa

Kodas _____

MICHAIL BREDICHIN

MAGISTRO BAIGIAMASIS DARBAS

**KLIENTŲ LOJALUMO VERTINIMAS NAUDOJANT NEURONINIUS
TINKLUS**

Kaunas, 2008

VILNIAUS UNIVERSITETAS
KAUNO HUMANITARINIS FAKULTETAS

INFORMATIKOS KATEDRA

MICHAIL BREDICHIN

MAGISTRO BAIGIAMASIS DARBAS

**KLIENTŲ LOJALUMO VERTINIMAS NAUDOJANT NEURONINIUS
TINKLUS**

Leidžiama ginti _____
(parašas)

Magistrantas _____

Darbo vadovas _____
(parašas)

Doc. Virginijus Sakalauskas

(darbo vadovo mokslo laipsnis, mokslo
pedagoginis vardas, vardas ir pavardė)

Darbo įteikimo data _____

Registracijos Nr. _____

Kaunas, 2008

TURINYS

TURINYS.....	3
SANTRUMPŲ SĄRAŠAS	4
LENTELIŲ SĄRAŠAS.....	5
PAVEIKLSŲ SĄRAŠAS.....	6
ĮVADAS.....	7
1. Kliento gyvenimo ciklo ir lojalumo samprata.....	10
1.1 Lojalumas- kas tai? Tradiciniai lojalių klientų atrinkimo būdai ir rodikliai	10
1.2 Vartotojo gyvenimo ciklo vertė (CLV).....	12
1.3 Vartotojo gyvenimo ciklo vertės apskaičiavimas.....	15
1.3.1 CLV apskaičiavimas	15
1.3.2 Taisyklė 55-15.....	16
1.3.2 CLV veiksmų prognozė	17
2. Duomenų analizės metodai	19
2.1 Informacijos išgavimo iš duomenų procesas	19
2.2 Duomenų gavybos metodai.....	21
2.2.1 Statistiniai paketai	22
2.2.2 Sprendimų medžiai.....	22
2.2.3 Taisyklių indukcija.....	22
2.2.4 Daugiaregresinė analizė	22
2.2.5 Neuroniniai tinklai.....	23
2.2.6 Duomenų analizės metodų palyginimas.....	24
2.3 Neuroniniai tinklai.....	25
2.3.1 Principinė struktūra	25
2.3.2 Duomenų analizė naudojant Neuroninius tinklus.	26
3. Klientų lojalumo tyrimas neuroninių tinklų pagalba.....	32
4. Lojalių klientų atpažinimas remiantis CVL pokyčio analize	35
4.1 Duomenų surinkimas.....	35
4.2 Duomenų paruošimas eksperimentui, apdorojimas.	37
4.2.1 Neuroninių tinklų formavimas:	38
4.2.2 Neuroninių tinklų mokymas, kontrolinės aibės analizė	39
5. Lojalių klientų atpažinimas remiantis kliento potraukį pirkti apibūdinančiais rodikliais.....	43
5.1 Duomenų paruošimas eksperimentui, apdorojimas.	43
5.2 Neuroninių tinklų formavimas, mokymas ir rezultatų analizė.....	45
5.3 Gautų klientų lojalumo tyrimo rezultatų analizė.....	46
IŠVADOS.....	49
LITERATŪROS SĄRAŠAS.....	52
SANTRAUKA	53
PRIEDAI	54

SANTRUMPŲ SĄRAŠAS

- CRM** (Customer Relationship Management) Ryšio su klientais valdymo sistemos
- aCRM** (analytical Customer Relationship Management) analitinio pobūdžio ryšio su klientais valdymo sistemos
- CLV** (Customer life value) – vartotojo gyvenimo ciklo vertė
- RFM** (Recency- Frequency- Monetary)- Naujumas, apsilankymo dažnumas ir vidutinis perkamumas

LENTELIŲ SĄRAŠAS

1 lentelė. Duomenų analizės metodų palyginimas	24
2 lentelė. Eksperimento duomenys: pardavimų operacijos	36
3 lentelė Apmokymo ir testavimo rezultatai	40

PAVEIKLSŲ SĄRAŠAS

1 Pav. Kliento gyvenimo ciklo etapai	13
2 Pav. Duomenų gavyba CRM kontekste.	19
3 Pav. Lojalūs i neloyalūs klientai	20
4 Pav. Lojalumo išskyrimo išvados.....	20
5 Pav. Duomenų gavybos metodų tipai.....	21
6 Pav. „Neuroninio tinklo principinė struktūra.”	25
7 Pav. „7 žingsnių metodologija“	27
8 Pav. Vartotojo gyvenimo ciklas ir klasifikavimo poreikio modelis.....	32
9 Pav „Lankomumas, pinigine vertė ir lojalumas“	33
10 Pav. Duomenų bazės schema: Duomenų šaltinių lentos	36
11 Pav. „Eksperimento duomenų sudėtis“	36
12 Pav. Neuroniniai tinklai ir importuoti duomenys matlab aplinkoje	39
13 Pav. „Kontrolinės aibės pateikimas tinklui“	40
14 Pav Grafikas: „Network 8“ išeigos duomenys ir kontrolinė aibė.	41
15 Pav. Grafikas: „Network3“ išeigos duomenys ir kontrolinė aibė.	42
16 Pav. Duomenų paruošimas eksperimentui	44
17 Pav. IPS vedlio paleidimas.....	45
18 Pav. Įeigos ir išeigos kintamųjų pasirinkimas.....	45
19 Pav. Tyrimo trukmės pasirinkimas.....	46
20 Pav. Kuriamų tinklų nustatymai.....	46
21 Pav. Sukurtų neuroninių tinklų struktūros abibendrinimas.....	47
22 Pav. Principinė neuroninio tinklo struktūra.....	47
23 Pav. Klasifikavimo statistika.....	48

IVADAS

Bet kokiai įmonei klientų lojalumas yra gyvybiškai svarbus. Per pastaruosius du dešimtmečius, dėmesys ryšio su klientais valdymo sistemoms (angl. CRM- customer relationship management) labai išaugo. Tai galima interpretuoti kaip įmonių siekimą individualiai pažinti savo klientus. Pasikeitė ir marketingo orientyras. Vietoj orientuoto į transakcijas, jis patapo orientuotu į klientą, kadangi praktika parodė, jog tai yra optimalus ryšys su klientu. (Daug pigiau pardavinėti prekes esamiems klientams, nei naujiems, mažesni bendradarbiavimo kaštai) Būtent todėl klientų lojalumo sąvoka interpretuojama kaip viena svarbiausių koncepcijų šiuolaikiniame marketinge.

Ryšio su klientais valdymo teorijose aprašoma daug kriterijų bei skaičiavimo metodų kliento lojalumui įvertinti.¹

Egzistuojanti vadybinė taisyklė 55-15² teigia jog bandant klasifikuoti geriausius klientus, maždaug 55 % yra klasifikuojami klaidingai ir dėl to įmonė praranda investicijas (klientams neskiriamas atinkamas dėmesys: nuolaidos, ir t.t.) Pagrindinė tokio netikslaus klasifikavimo priežastis- besikeičianti verslo aplinka. Lojalūs klientai neišlieka tokiais visada.

Tyrimo problema: Lojalių ir nelojalių klientų atpažinimas

Tyrimo objektas- egzistuojantys vadybiniai lojalių klientų atpažinimo metodai

Tyrimo tikslas: Remiantis konceptualia duomenų apie klientus analizę atrinkti kriterijus, bei metodus, kurių pagalba būtų įmanoma atpažinti lojalų klientą ankstyvose ryšio su klientu stadijose. Pritaikyti duomenų analizės modelį lojalių klientų atpažinimui remiantis istoriniais bendradarbiavimo su klientu metu surinktais duomenimis.

Siekiant iškelto tikslo buvo keliami tokie **uždaviniai:**

1. Atlikti ryšio su klientais valdymo teorijų, skirtų lojaliems klientams atpažinimui bei lojalumui įvertinti analizę.
2. Išnagrinėti duomenų gavybos įrankius lojalumo nustatymui ir klasifikavimui.
3. Išnagrinėti neuroninių tinklų struktūrą bei išanalizuoti jų taikymo galimybę duomenų apie klientus analizei klasifikavimui pagal lojalumo požymį.
4. Pasiūlyti naują lojalių klientų atpažinimo ankstyvuose bendradarbiavimo su klientu stadijose modelį.
5. Atlikti eksperimentą su pardavimų duomenimis ir nustatyti pasiūlyto modelio tinkamumą lojalių klientų atpažinimui.

Darbo struktūra: darbas yra sudarytas iš penkių skyrių.

¹ Edward C.Malhouse ir Robert C. Blattberg (2005m.)

² Edward C.Malhouse And Robert C. Blattberg, (2006) „Can we predict customer lifetime value?“

1. *Kliento gyvenimo ciklo ir lojalumo sampratos* skyriuje atskleidžiama lojalumo samprata, apžvelgiami naudojami šiuolaikinių ir klasikiniai klientų vertinimo metodai. Detaliau nagrinėjami kliento gyvenimo ciklo etapai bei vertės paskaičiavimai. Suformuojamas lojalumo tyrimo uždavinys. Remiantis suformuotu uždaviniu 2. *Duomenų analizės metodų dalyje* apžvelgiami informacijos išgavimo iš duomenų procesai, metodai, jų taikymo specifika bei modeliai. Nagrinėjamas jų tinkamumas šiam uždaviniui spręsti. Palyginimo būdu pasirenkamas tinkamiausias metodas. Detaliau nagrinėjama jo sudėtis bei taikymas. 3. *Klientų lojalumo tyrimas neuroninių tinklų pagalba* skyriuje aprašomi siūlomi sprendimai nagrinėjamai problemai spręsti. 4. *Lojalių klientų atpažinimo remiantis CVL pokyčio analize* ir 5. *Lojalių klientų atpažinimo remiantis kliento potraukį pirkti apibūdinančiais rodikliais* skyriuose eksperimento pagalba, vertinamas siūlomų sprendimų tinkamumas apibrėžtai problemai spręsti.

Galiausiai apibendrinami gauti rezultatai ir pabrėžiama jų teorinė bei praktinė reikšmės

Darbe naudoti literatūros šaltiniai

Siekiant išsiaiškinti suformuotos problemos ištyrimo lygį, atlikta klasikinės tradicinių lojalių klientų atrinkimo metodų ir rodikliai analizė: Pareto, Oliver, klasikiniai Newman ir Werbel apibrėžimai. Detaliau nagrinėtas mokslininkų Wouter Buckinx, Geert Verstraeten, Dirk Van den Poel pasiūlyti kliento „norą pirkti“ apibūdinantys rodikliai.

Nagrinėjant vartotojo gyvenimo ciklo vertę (toliau CLV) detaliai apžvelgtas kompleksinis E. Ksevelonakiso pasiūlytas CLV etapų modelis. O kiekybinam CLV įvertinimui nagrinėta Bolton, Lemon, ir Verhoef (2004) pasiūlyta formulė

Tyrimo metodai. Siekiant išsiaiškinti pagrindinius rodiklius, bei metodus, kurių pagalba galima apibrėžti ir nustatyti kliento lojalumo statusą, buvo atliekama mokslinės literatūros loginė analizė; egzistuojančių duomenų gavybos metodų analizei naudotas mokslinės literatūros analizės bei palyginimo metodai. Siūlomam metodui suformuoti naudojamas apibendrinimo bei modelio kūrimo metodai.

Darbo rezultatų teorinė ir praktinė reikšmė

Remiantis CLV pokyčio analize ir Wouter Buckinx pasiūlytais kliento „norą pirkti“ apibūdinančiais rodikliais sudaryti 2 modeliai. Teoriškai šie modeliai yra veiksmingesni nei E.Ksevelonakiso pasiūlytas „CLV ir įmonės tikslinės strategijos sąveikos modelis“, kadangi laiku gavus informaciją apie potencialią kliento klasę (lojalus/nelojalus), įmonė gali nedelsiant taikyti tikslines strategijos priemones ir laiku pateikti klientui atitinkamą pasiūlymą jau bendradarbiavimo stadijoje, sumažinant jo atsako laiką (lag) ir tikimybę atsitraukti.

Praktiškai skirtingai nei kitos šio metu siūlomos priemonės, šiame darbe pasiūlytas metodas išsiskiria savo lankstumu. Jis nereikalauja griežto lojalių klientų apibrėžimo, kadangi gali mokytis iš pateiktų pavyzdžių. Kadangi tai klasifikavimas remiantis istoriniu potraukį apibūdinančių

rodiklių pokyčiu (prognozavimas), įmonė gali laiku (užlaikymo stadijos pradžioje) išgauti informaciją klientus, kurie yra linkę atsitraukimai, o kurie potencialiai Lojalūs.

Neuroninis tinklas sudarytas CLV vertės pokyčio pagrindu pripažintas neobjektyviu dėl įeigos kintamųjų tarpusavio priklausomybės

Eksperimento metu, atliekant neuroninio tinklo, sudaryto remiantis „norą pirkti apibūdinančiais rodikliais“, vertinimą iš 154 lojalių klientų atpažino 141, suklydo atitinkamai dėl 13 klientų. Atpažįstant neaiškius, iš 119 pateiktų jis nustatė 82 o suklydo dėl 37 klientų. Atrenkant lojalūs, iš 147 klientų jis atrinko 102 teisingai, o dėl 45 suklydo. Nežinomų variantų nepasitaikė.

Tinklo klasifikavimo rezultatai iš dalies paneigia vadybinę taisyklę 55-15 kuri teigia, jog įmonės bandant klasifikuoti geriausius (20 % pelningiausių) klientus, maždaug 55 % yra klasifikuojami klaidingai. Šio tyrimo metu tinklas suklydo tik dėl 13 klientų, kas sudaro 8%.

Darbo apribojimai ir sunkumai

Su pagrindiniais sunkumais susidurta atliekant eksperimentinę darbo dalį. Kadangi eksperimentuojama su realiais pardavimų duomenimis šiame darbe neskelbiamas įmonės pavadinimas. Siekiant užtikrinti konfidencialumą, eksperimento duomenys buvo koduojami, o kainos buvo pakeistos išlaikant priklausomybę tarp operacijų duomenų (visoms operacijų kainoms buvo parinktas bendras daugiklis)

Dėl didelio duomenų kiekio, atmestos fiskalinių kasų (POS), nutolusių padalinių ir firminių parduotuvių pardavimo operacijos

Darbo struktūra ir apimtis

Darbas susideda iš 66 puslapių, jame yra sukurtos 3 lentelės, 21 paveikslai, 13 puslapių skirti priedams.

1. Kliento gyvenimo ciklo ir lojalumo samprata

Bet kokiai įmonei klientų lojalumas yra gyvybiškai svarbus. Per pastaruosius du dešimtmečius, dėmesys ryšio su klientais valdymo sistemoms (angl. CRM- customer relationship management) labai išaugo. Tai galima interpretuoti kaip įmonių siekimą individualiai pažinti savo klientus. Pasikeitė ir marketingo orientyras. Vietoj orientuoto į transakcijas, jis patapo orientuotu į klientą, kadangi praktika parodė, jog tai yra optimalus ryšys su klientu. (Daug pigiau parduoti prekes esamiems klientams, nei naujiems, mažesni bendradarbiavimo kaštai) Būtent todėl klientų lojalumo sąvoka interpretuojama kaip viena svarbiausių koncepcijų šiuolaikiniame marketinge.

Šio skyriaus paskirtis- išnagrinėti pagrindinius modelius lojaliems klientams atskirti bei lojalumui įvertinti. Pradžioje apžvelgiami lojalumo apibrėžimai, atrinkimo būdai, klasikiniai rodikliai remiantis kuriais galima įvertinti klientus ir galiausiai analizuojami kliento gyvenimo ciklo samprata ir etapai.

1.1 Lojalumas- kas tai? Tradiciniai lojalių klientų atrinkimo būdai ir rodikliai

I.Pikturnienė (2003) kliento lojalumą įvardina kaip „(...) prisirišimą prie prekės ženklo, parduotuvės arba tiekėjo, paremtą stipriu teigiamu požiūriu ir pasireiškiančiu nuosekliu kartotiniu prekės pirkimu arba naudojimu paslauga“. Dar prieš keletą metų, Oliver (1999) lojalumą apibrėžė kaip „(...) vidinį vartotojo įsipareigojimą ateityje pakartotinai įsigyti prekę (paslaugą) arba pakartotinai apsilankyti prekybos vietoje, nepaisant situacijos įtakos ar marketingo pastangų pakeisti vartotojo elgseną“. Pasak Newman ir Werbel (1973), lojalūs vartotojai yra tie, kurie: pakartotinai įsigyja prekės (paslaugos) ženklą; prieš įsigydami svarsto galimybę įsigyti būtent tą prekės (paslaugos) ženklą; neišitraukia į su šiuo prekės (paslaugos) ženklu susijusią informacijos paiešką.

Tai yra keletas daugelio sukurtų apibrėžimų. Jie tik iš dalies apibūdina lojalius klientus, ir neleidžia įmonei vienareikšmiškai įvertinti lojalumą. Naujų visuomenės informavimo priemonių augimas, nesibaigiantys technologiniai patobulinimai ir išradimai stipriai padidino galimybes sąveikauti tiesiogiai, t.y. vienas su vienu ryšiu. Pastaraisiais metais buvo sukurta daug įrankių, kurie padeda įmonėms sustiprinti esamą ryšį su klientais. Šie įrankiai yra daugiau operatyvinio pobūdžio (kontaktų saugojimas, akcijos, pasiūlymai) nei analitinio (pvz. Lojalių klientų tyrimas)

Retos kompanijos naudoja faktinį klientų lojalumą įmonių vykdomųjų marketingų strategijų įvertinimui. Ko gero pagrindinė šio reiškinio priežastis- informacijos nepasiekiamumas. Įmonės, palaikančios tradicines transakcines duomenų bazes, kurios saugo informaciją apie klientų kontaktus, negali apskaičiuoti prekių, kurias klientas perka iš konkurentų, kiekio. Mokslinis Vorhoef (2002) darbas parodė, kad 7,5 % įmonių renka ir saugo tokia informaciją.

Duomenų apie klientus kaupimas ir statistinės informacijos gavyba yra šiuolaikinės verslo sprendimų srities CRM kompetencijoje. Šiose sistemose plačiai naudojami tokie rodikliai:

RFM (angl. Recency- Frequency- Monetary)- Naujumas, apsilankymo dažnumas ir vidutinis perkamumas (Lt.). Naujumas ir apsilankymo dažnumas parodo ar klientas yra atsitiktinis ar pastovus. Vidutinis perkamumas parodo kiek klientas yra pirkęs iš įmonės vidutiniškai per vieną apsilankymą (Lt.) Bendrai šie rodikliai neparodo ar klientas yra lojalus, neutralus ar atsitiktinis. Jie tik statistiškai apibūdina kliento nešamą naudą įmonei.

Pareto 80/20³- Remiasi fenomenu kad 20 % geriausių įmonės klientų duoda 80% viso įmonės pelno. Ryšio su 20- ties % geriausių klientų palaikymas yra kritiškai svarbus momentas bet kokiai įmonei. Todėl jiems dažniausiai siūlomos įvairios nuolaidos. (pvz. „New York Mets” ir “New York Metropolitan Opera” siūlo sezoninius bilietus savo geriausiems klientams patrauklia kaina)

Analizė remiantis banginio kreivė. Tai Pareto principu paremtas metodas. Sudaroma kreivė, kurioje parodomas sukauptas pelnas (Lt. arba %) su lyg kiekvienu klientu (%) Gaunama kreivė yra banginio formos (iš čia ir kilęs pavadinimas) kur sparčiai kyla pradedant 1 iki 30 % (parėto taisyklė) geriausių klientu ir sustoja kilti (arba leidžiasi žemyn) formuodama būdingą banginio uodegą, kuri nustato nuostolingus klientus. Šis metodas yra vaizdinis

Remiantis aukščiau aprašytais rodikliais, įmonė gali atskirti savo gyvybiškai svarbius klientus ir klientus, kurie įmonei yra nuostolingi.

Straipsnio „Predicting customer loyalty using internal transactional database“ autoriai (Wouter Buckinx, Geert Verstraeten, Dirk Van den Poel) siūlomame klientų vertinimo modelyje naudoja sekančius rodiklius, kiekybiškai apsakantčius kliento „norą pirkti“:⁴

- Kliento pirkimai per pastarąsias 5 kalendorines mėnesio dienas
- Kliento pirkimai per pastarąsias 10 kalendorinių mėnesio dienų
- Suminiai kliento pirkimai per nagrinėjamą mėnesį
- A kategorijos prekių (aukšta kaina, mažas pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- B kategorijos prekių (vidutinė kaina, vidutinis pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- C kategorijos prekių (žema kaina, didelis pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- D kategorijos prekių (labai žema kaina, didelis pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- Suminis nupirktų prekių per nagrinėjamą mėnesį skaičius
- Nupirktų A kategorijos prekių skaičius per nagrinėjamą mėnesį

³ W. G. ZIKMUND, (2003) *CRM Integrating Relationship Management and Information Technology*. p. 15

⁴ Wouter BUCKINX, Geert VERSTRAETEN, Dirk Van den POEL, (2007) „Predicting customer loyalty using internal transactional database“

- Nupirktų B kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų C kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų D kategorijos prekių skaičius per nagrinėjamą mėnesį
- Paskutinio apsilankymo naujumas (dienų skirtumas po paskutinio apsilankymo)
- Standartinis apsilankymo naujumo nuokrypis (standartinis nuokrypis dienų skirtumo tarp apsilankymų)
- Vidutinis apsilankymo naujumas (vidutinis dienų skirtumas tarp apsilankymų)

Pabrėžtina, jog prie rodiklių paminėti periodai gali būti ilgesni arba trumpesni, priklausomai nuo operuojamų duomenų. (Vietoj mėnesio gali būti metai, vietoj dienų-savaites ir pan.) Straipsnio autoriai teigia, kad nustatčius tinkamą priklausomybę tarp šių rodiklių ir lojalumo- galima sėkmingai klasifikuoti lojalius klientus.

Mokslinėje literatūroje taip pat yra aprašomi modeliai, remiantis kuriais galima nustatyti kliento patrauklumą įmonei:

- Vartotojų reakcijos (atsako) modeliavimas yra vienas šios srities atstovas, kadangi šie modeliai tarnauja kaip įrankis ryšiui su klientu valdyti. Modelio esmė- parinkti siūlomas paslaugas/prekes taip, kad atitektų vartotojų poreikius su didžiausią tikimybę. Kitaip tariant, optimalaus produkto/paslaugos ar jų kombinacijos paieška duotam vartotojui.
- Churn analizės reprezentuoja įmonių baimę prarasti esamus klientus. Jos tikslas- nustatyti klientus, kurie turi didžiausią potencialą nutraukti esamą ryšį su įmone.
- CLV(customer life time value- kliento gyvavimo ciklo vertė) analizė yra plačiausiai naudojamas metodas, siekiant nustatyti klientų ateities potencialą ir tokiu būdu skirti dėmesį tik perspektyviems.

Kadangi iš minėtų modelių tik CLV analizė pasižymi kliento ateities potencialo apibūdinimo savybėmis, pastarąjį apžvelgsime plačiau.

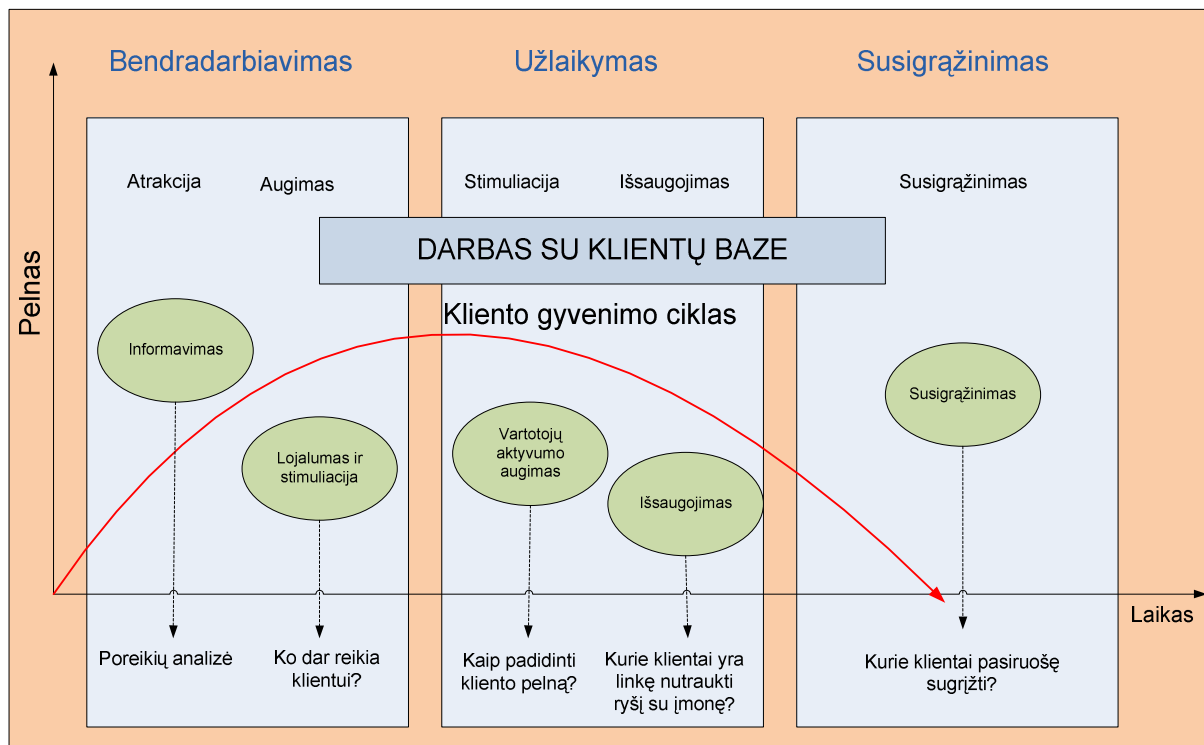
1.2 Vartotojo gyvenimo ciklo vertė (CLV)

CLV pirmą kartą paminėtas marketingo literatūroje 1989 metais. Apibrėžiamas kaip „<...> esamų ir būsimų transakcijų grynoji vertė dabar.“ (Gupta, Lehmann ir Stuart 2004) CLV yra išskiriamas kaip vertingas rodiklis klasifikuojant ir suteikiant prioritetus (remiantis pelnu) klientams bei apskaičiuojant marketinginių investicijų atsipirkimą. (Rust, Lemon ir Zeithaml 2004)

Ryšio su klientais valdymo marketingas (kartais taip vadinamas CRM) yra naujas požiūris į marketinginės veiklos vykdymą. Jis orientuotas į klientų lojalumą ir sąveikos su jais valdymą.

Apžvelgiant šią koncepciją, būtina išskirti eilę raktinių ypatumų, kurie lemia strateginius šios sąveikos aspektus bei operacinio funkcionavimo principus

Svarbus tikslinio marketingo (targeted marketing) principas yra sąveikos su klientu individualizacija, kas operaciniame lygyje pasireiškia kaip įmonės komunikacinės strategijos bei individualių pasiūlymų formavimo naudojimą. Siekiant užtikrinti maksimalių tikslinės sąveikos su klientais efektyvumą, svarbu stebėti kliento vertę visame nagrinėjamame sąveikos su tuo klientu laiko periode. Iš to seka gyvenimo ciklo etapų formalizavimo įmonės atžvilgiu būtinumas. Pav. 1 yra pateiktas kompleksinis E. Ksevelonakiso (E. Ksevelonakis, 2007) pasiūlytas modelis. Jis atspindi kliento gyvavimo ciklo ir įmonės tikslinės strategijos sąveiką.



Šaltinis: Sukurta autoriaus pagal (E. Ksevelonakis, 2007)

1 Pav. Kliento gyvenimo ciklo etapai

Įmonės atžvilgiu bendradarbiavimas susideda iš 3jų pagrindinių stadijų: **Bendradarbiavimas** (susipažinimas su klientu, pradiniai prikimai ir pan.), **Užlaikymas** (siekiama išsiaiškinti kurie klientai linkę atsitraukimui) ir **Susigrąžinimas** (nutraukusių ryši su įmone susigrąžinimas)

Pagal E.Ksevelonakisą, gyvavimo ciklą galima suskirstyti į sekančius etapus:

1. Susipažinimas su įmone. Tikslinės sąveikos rėmuose, šiame etape pagrindė naudojami įrankiai, skirti klientų apie įmonės pasiūlymus informavimui.

2. Pirmo įsigijimo etapas charakterizuojamas kaip įmonės pastangos analizuoti papildomus kliento poreikius ir siekimas juos patenkinti- taip stimuliuojant klientą aktyviai naudotis siūlomomis paslaugomis ar vartoti produktą.
3. Augimo etapas, charakterizuojamas įmonės pastangų koncentracija ties kliento lojalumo ir jo produktų ar paslaugų aktyvumo stimuliavimu.
4. Kliento sprendimo dėl ryšio su įmone nutraukimo priėmimo etapas. Šiame etape tikslinės sąveikos užduotis yra išsaugoti klientą.
5. Išėjusio kliento susigrąžinimas. Net jeigu klientas nutraukė bendradarbiavimą su įmone, ji gali pabandyti jį susigrąžinti. Tam taip pat gali būti naudojami tikslinio marketingo metodai.

E. Ksevelonakiso siūlomam modelyje taip pat yra pavaizduota pelno kreivė (raudonai). Ji atstovauja pelno, gaunamo iš kliento, pokytį visame kliento gyvavimo ciklo periode. Kaip galime matyti iš (1 Pav.) ji sparčiai kyla nuo bendradarbiavimo su klientu pradžios ir pasiekia piką tarp užlaikymo ir bendradarbiavimo etapų. Paskutinis etapas yra susigrąžinimas, kurio metu įmonė bando susigrąžinti klientus, kurie nutraukė ryšį su įmone. Natūralu, kad šiame etape kreivė tolygiai mažėja (klientas mažai perka, investuojami papildomi pinigai į susigrąžinimą)

Šis modelis detalai apibūdina įmonės tikslinės strategijos sąveiką su individualiu klientu. Tačiau remiantis marketingine praktika, reta įmonė yra pajėgi skirti tokio detalumo dėmesį

Iš čia gimsta poreikis klientus klasifikuoti (ir atitinkamai suteikti prioritetus). Remiantis klasifikavimo rezultatais įmonė gali pasirinkti kurių klientų lojalumą skatinti labiau, kuriems suteikti ypatingus pasiūlymus, o kuriems neskirti dėmesio visiškai (analogiškai neinvestuojame į visas pasaulio biržose prekiaujamas akcijas, bet pasirenkam tas, kurių pelningumas yra labiausiai tikėtinas atsižvelgiant į tai kiek pinigų galime skirti investavimui)

Taikant šį modeli klientų visumai (o ne individualiems klientams) galime pastebėti pagrindinį trūkumą: klausimas „kurie klientai yra linkę atsitraukimui“ gimsta tik užlaikymo stadijos pabaigoje. Nustačius kurie klientai yra linkę atsitraukti įmonė pereina prie susigrąžinimo stadijos. Tuo tarpu Reinartz and Kumar (2002) teigia “let butterflies fly.” (leiskime drugeliams išskristi) ir koncentruotis ties potencialiai pelningais klientais.

Įmonės galimybė klasifikuoti turimą klientų bazę yra svarbus tikslinės sąveikos aspektas. Mokslinėje literatūroje siūloma išskirti klasifikavimo procesą į tokius etapus: klasifikavimas, vykdomas siekiant apibrėžti sąveikos su klientu strategiją ir segmentaciją, siekiant išanalizuoti kliento poreikius. (o vėliau, gal būt, tiekti individualius pasiūlymus). Sąveikos su kiekviena išskiriama klase strategijos nustatymas priklauso nuo raktinių marketinginių charakteristikų, įeinančių į šias klases. (pvz. vertė įmonei, lojalumo lygis, potencialas nutraukti ryšius su įmone ir

t.t...). Klasifikavimo etapas gali naudoti skirtingus parametrus, padedančius padaryti apibendrinančias išvadas dėl individualaus pasiūlymo tiekimo.

K.Balašovas siūlo išskirti tokius klasifikavimo parametrus⁵:

- Elgesio klasifikavimas
- Psichografinis klasifikavimas
- Klasifikavimas remiantis lojalumo ir naudos požymiais
- Ekonominis klasifikavimas
- kt..

Jis taip pat pabrėžia, jog klasifikavimo parametrų spektras, o taip pat klientų pasiskirstymo (formuojamų klasės viduje) pagal šiuos parametrus, specifika yra stipriai įtakojami verslo aplinkos ir turi būti nustatomi kiekvienu konkrečiu atveju

Iš aukščiau minėtų parametrų, klasifikavimas remiantis lojalumo ir naudos požymiais šio darbo temai yra aktualiausias.

1.3 Vartotojo gyvenimo ciklo vertės apskaičiavimas

Kaip buvo minėta CLV yra vienas iš būdų pamatuoti santykinį kliento arba klientų segmento patrauklumą. Besivystančio (ankstyvo) ryšio su klientais stadijoje CLV vertinimas yra labai svarbus rodiklis. Įmonės gali identifikuoti vartotojus su aukštu CLV net prieš pasirodant aukštos vertės kliento požymiams (lojalus, potencialiai lojalus klientas). Straipsnyje „Lifetime Value Prediction at Early Customer Relationship Stages“, autorius Wangenheim apibūdina savo pasiūlytą modelį, kuris pasak jo gali prognozuoti transakcijų skaičių duotam laiko momentui ir CLV, remiantis duomenimis, kurie yra prieinami įmonei ankstyvoje ryšio su klientu vystymosi stadijoje.

1.3.1 CLV apskaičiavimas

Skirtinguose straipsniuose siūlomi skirtingi CLV modeliavimo metodai priklausomai nuo tiriamos srities. Pvz. Daugumoje literatūros, kur ryšys su klientu apibūdinamas kaip turto valdymas, siūlomos 3 dimensijos, kurios, pasak straipsnio autorius, turi būti įtrauktos į CLV modeliavimą (Bolton, Lemon, ir Verhoef, 2004): santykio su klientais *ilgis*, *gylis* ir *plotis*. Daromos sekančios prielaidos:

- modelis koncentruojasi ties viena įmonės tiekama preke/paslauga. Tai reiškia kad santykio su klientais plotis neįtraukiamas (nėra pasirinkimo);

⁵ Balašovas K., 2007, „Маркетинговая Стратегия Целевого Взаимодействия С Клиентами“, pusl. 5

- santykio ilgis (tęstinumas) interpretuojamas kaip nuolat besitęsiantis. Kitaip tariant, jei klientas kurį laiką neperka paslaugos/prekės- jis neinterpretuojamas kaip „prarastas visam laikui“. Daroma prielaida, kad jis gali sugrįžti. Ši prielaida yra realistiškesnė.
- Santykio plotis. Kuris apibrėžiamas kaip paslaugų naudojimo arba prekių įsigijimo intensyvumo lygis tam tikram laiko periodui, kuris nusako vartotojo elgesį (pvz. Kai vartotojas perka brangesnes aukštos vertės prekes vietoj baziniu (pigėsniu mažesnės vertės)

Bendruoju atveju vartotojo CLV apskaičiuojamas pagal formulę (1):

$$CLV_i = \sum_{t=1}^T \frac{(CM_{NRT} * NT_{it} + CM_{Up} * UP_{it})}{(1+d)^t} \quad (1)$$

kur

CLV_i	vartotojo <i>i</i> gyvenimo ciklo vertė
NT_{it}	bazinių (pvz. pigėsniu mažesnės vertės) transakcijų, vykdytų vartotojo <i>i</i> per laiko periodą <i>t</i> , skaičius.
CM_{NRT}	vidutinis bazinių transakcijų skaičiaus pokytis
UP_{it}	didelės vertės (pvz. brangesnių prekių aukštesnės vertės) transakcijų, vykdytų vartotojo <i>i</i> per laiko periodą <i>t</i> , skaičius
CM_{Up}	vidutinių pirkinio tipo pakeitimų (iš bazinio į didelės vertės) skaičiaus pokytis
<i>d</i>	nuolaida
<i>t=1..T</i>	tiriamų laiko periodų skaičius

Šis aukščiau apibrėžtas modelis interpretuoja bazines ir didelės vertės transakcijas kaip nepriklausomus įvykius.

Šiuo metu CLV prognozavimas ir kitų paminėtų modelių taikymas sulaukia nedaug dėmesio. Straipsnyje „Can we predict customer lifetime value?“ (Edward C.Malthouse And Robert C. Blattberg) yra paminėtos 2 taisyklės dėl kurių tradiciniai klientų klasifikavimo metodai yra netikslingi

1.3.2 Taisyklė 55-15⁶

Vadybine prasme 80-15 taisyklės išvados yra jog įmonės leidžia didelius pinigus klaidingiems (ne tiems, kam turėtų) klientams.

Bandant klasifikuoti geriausius klientus, maždaug 55 % yra klasifikuojami klaidingai (ir jiems nebus skirtas atitinkamas įmonės dėmesys: nuolaidos, ir tt.)

Bandant klasifikuoti vidutinius klientus, maždaug 15 % yra klasifikuojami klaidingai (jiems nebus skirtas atitinkamas įmonės dėmesys)

⁶ Edward C.Malthouse And Robert C. Blattberg, (2006) „Can we predict customer lifetime value?“

Taip pat aišku, jog geriausi klientai neišlieka tokiais visada. Reinartz and Kumar (2002) turėjo tą patį omenyje sakydami „leiskime drugeliams išskristi (“let butterflies fly.”) Kitaip tariant jei įmonė aptinka, jog klientas nebėra lojalių tarpe ir nebeteikia jam nuolaidų ir kitų paskatų, ji turi būti pasiruošusi tas paskatas suteikti kitam „potencialiai lojaliam“ klientui, kuris atneštų tapusio nelojaliu kliento pelno dalį. (kitap pagal pareto taisyklę įmonės pelnas ženkliai sumažės) Klaidingas klasifikavimas sąlygoja potencialų nuostolį. „Geriausias klientas“, kuris klaidingas klasifikuojamas kaip „įprastas“ ieškos alternatyvų t.y. kreipsis pas konkurentus, arba išvystys neigiamą nuomonę apie įmonę. Iš kitos pusės, įprastas klientas klasifikuojamas kaip „geriausias“ nepatenkins įmonės lūkesčių (neatneš įmonei tikėtų rezultatų.)

Kita priežastis- CLV pokytį įtakojantys veiksniai, kuriuos apžvelgsime toliau.

1.3.2 CLV veiksnių prognozė

Nuo 2001 metų CLV apibrėžimų ir išvedamų formulių tipai sparčiai išaugo (pvz. Hitt ir Frei 2002; Thomas2001; Reinartz ir Kumar 2003; Villanueva, Yoo, ir Hanssens, 2004; ir kt...) Atsirado poreikis teorinėje struktūroje, kurios pagalba galima būtų nustatyti ryšį tarp CLV ir jos įtakojančių veiksnių. (Bolton, Lemon ir Verhoef 2004)

(Florian V. Wangenheim, (2006) siūlo išskirti tokius pagrindinius veiksniai, įtakojančius CLV:

Kliento sąveika su įmone. Naudojamas kreipimosi kanalas

Šis veiksnys gali būti stebimas pačioje ryšio su klientu užmezgimo pradžioje (koku būdu vartotojas nusiperka prekę, t.y. kreipiasi į įmonę). Bolton, Lemon ir Verhoef (2004) šioje srityje iškėlė svarbų klausimą „ar vartotojo pasirinktas kanalas įtakoja CLV dimensijas. Pastaruoju metu atlikti tyrimai ir analizės parodė, jog vartotojai gali pasirinkti iš daugybės kanalų, ir jų pirminis pasirinkimas indikuoja ateities vertę. Keane ir Wang (1995) parodė, kad vartotojų išlaikymo lygis svyruoja priklausomai nuo kanalo, o Thomas (2001) nustatė, kad kanalo pasirinkimas įtakoja numatomą įmonė-klientas ryšio trukmę. Logiška, jog klientai, kurie aptarnaujami internetu yra didesnės vertės nei tie, dėl kurių samdomas papildomas personalas ir naudojami papildomi resursai (kasos, eilių terminalai, ir t.t)

Tiesioginio komunikacijos kanalo užtikrinimas teigiamai įtakoja ateities transakcijas.

Parodomasis elgesys. Vartotojo pasirinkimo kompetencija

Žiūrint iš įmonės pusės, būti pasirinktam dėl aukštu konkurencingumu pasižyminčių paslaugų- yra svarbesnis pirmenybės iš kliento pusės teikimo argumentas, nei būti pasirinktam dėl monopolijos. Praktika parodė, kad vartotojai, kurie pasirenka įmonę iš visų konkurentų, yra labiau linkę pakartoti transakcija ateityje, nei transakcijų su ribotomis

pasirinkimo galimybėmis atvejais. Pasirinkimo konkurencingumas pozityviai įtakoja ateities transakcijas. Ekonominiai vartotojo elgesio modeliai daro prielaidą, jog praeities elgesys įtakoja ateities elgesį. Venkatesan ir Kumar (2004) empiriškai parodė, jog praeities laiko tarpas tarp įsigijimų yra geras ateities tarpų tarp įsigijimų prognozavimo rodiklis. Venkatesan tai pavadino „inercija“ Laiko tarpas tarp įsigijimų yra atvirkščias transakcijų dažnumo požūriu.

Demografinė informacija

Yra prieinama ankstyvoje santykio su klientu stadijoje (amžius ir lytis). Labai tikėtina, jog amžiui didėjant, „inercijos“ poveikis bus didesnis dėl mažėjančio noro eksperimentuoti. Tačiau jokių tiesioginių hipotezių nėra iškelta, dėl teorijos trūkumo. Taip pat potencialus efektas gali skirtis priklausomai nuo industrijos.

Klasikiniai lojalumo apibrėžimai tik iš dalies apibūdina lojalius klientus, ir neleidžia įmonei vienareikšmiškai įvertinti lojalumą.

Apžvelgus pagrindinius modelius , o taip pat (pareto 20/80, RFM ir Banginio kreivės) lojaliems klientams atskirti bei lojalumui įvertinti paaiškėjo, kad plačiausiai paplitęs metodas yra Pareto 20/80, kuris teigia jog 20 procentų geriausių (pelningiausių) įmonės klientų duoda 80 procentų įmonės pelno. Remiantis šiais rodikliais, įmonė gali atskirti savo gyvybiškai svarbius klientus ir klientus, kurie įmonei yra nuostolingi.

Pasak Wouter Buckinx kliento „norą pirkti“ galima aprašyti tokiais rodikliais kaip: Kliento pirkimai per pastarąsias 5, 10 kalendorinių mėnesio dienų, suminiai kliento pirkimai; A,B,C,D kategorijos prekių pirkimai bei suminis nupirktų prekių per nagrinėjamą mėnesį skaičius

Iš apžvelgtų kliento patrauklumo įmonei įvertinimo modeliu pasirinktas ir apžvelgtas plačiau CLV analizės metodas, kadangi jis pasižymi kliento ateities potencialo apibūdinimo savybėmis. Nagrinėjant E. Ksevelonakiso modelį (CLV modelis) pastebėtas pagrindinis trūkumas: klausimas „kurie klientai yra linkę atsitraukimui“ gimsta tik užlaikymo stadijos pabaigoje. Nustačius kurie klientai yra linkę atsitraukti, įmonė pereina prie susigrąžinimo stadijos. Tuo tarpu Reinartz and Kumar (2002) teigia “let butterflies fly.” (leiskime drugeliams išskristi) ir koncentruotis ties potencialiai pelningais klientais.

Kaip matome, lojalumo tyrimas susiveda į sekančius uždavinius: klasifikavimas remiantis kliento norą pirkti apibūdinančiais rodikliais ir klasifikavimas remiantis CLV vertės pokyčio analize.

Sekančiame skyriuje apžvelgsime pagrindinius statistinius metodus, ir išnagrinėsime jų tinkamumą šiems uždaviniams spręsti.

2. Duomenų analizės metodai

Beveik visos šiuolaikinės įmonės kaupia duomenis apie klientus. Tačiau turėti duomenis ir informacija yra skirtingi dalykai. Informacijos išgavimas iš duomenų yra sudėtinis procesas, kurio komplikotumas priklauso nuo norimos išgauti informacijos pobūdžio. Lojalumo tyrimas susiveda į sekančius uždavinius: klasifikavimas remiantis kliento norą pirkti apibūdinančiais rodikliais ir klasifikavimas remiantis CLV vertės pokyčio analize.

Šiame skyriuje apžvelgsime informacijos išgavimo iš duomenų procesą, metodus, jų taikymo specifiką, ir tinkamumą šiam uždaviniui spręsti

2.1 Informacijos išgavimo iš duomenų procesas

Duomenų apie klientus analizei atlikti dažniausiai naudojami Data-mining metodai ir įrankiai. Jų esmė CRM kontekste- nustatyti kliento tipą (arba prognozuoti jo ateities elgesį) remiantis patyrimu arba iš anksto apibrėžtais kriterijais ir taisyklėmis. Pav. Yra pavaizduota bendra duomenų apie klientus analizės stadijos. Duomenys dažniausiai saugomi įmonės duomenų saugykloje.



Šaltinis: sukurta autoriaus

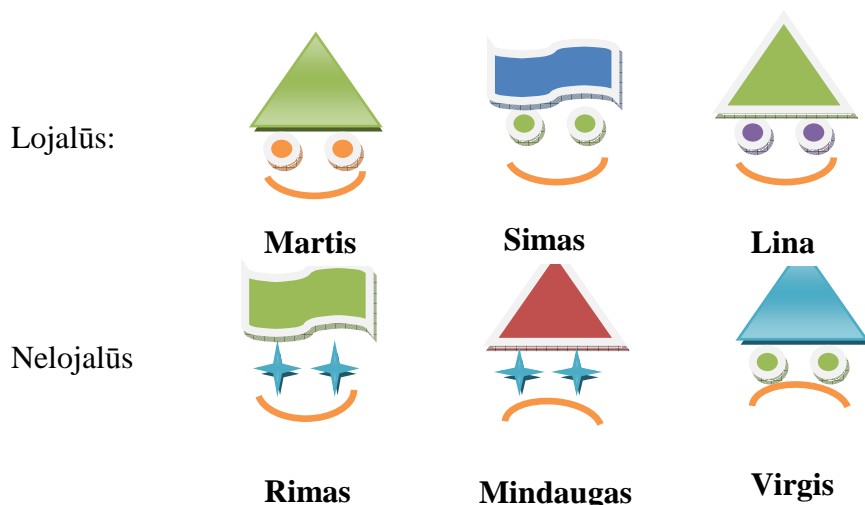
2 Pav. Duomenų gavyba CRM kontekste.

Čia jie susistemunami pagal turinio pobūdį. Duomenų gavybos proceso įeiga – Vartotojo profilis, tai yra duomenų apie klientą rinkinys (uždavinio kontekste tai informacija apie CLV, norą

pirkti apibūdinantys rodikliai). Šio proceso išeiga-nustatytos klientų klasės. Remiantis gauta informacija, įmonė daro atitinkamus sprendimus kliento, arba įmonės kampanijos atžvilgiu (lojalumo skatinimo priemonės, strategija, tikslai ir t.t.). Taip iš duomenų gaunama informacija.

Kaip buvo minėta, norint atskirti lojalius klientus nuo nelojalių, reikia iš anksto apsibrėžti lojalumą (nurodyti požymius, nusakančius lojalumą). Vieni metodai reikalauja apibrėžimo, o kitiems pakanka duomenų egzempliorių (apmokymo). Kadangi mes neturime a-priorių žinių apie lojalumą, pastaruosius apžvelgsime plačiau.

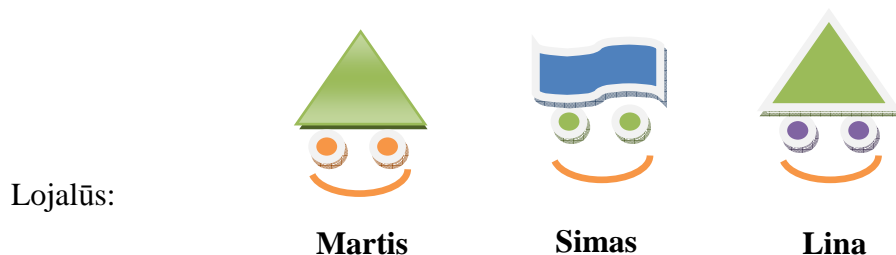
Tarkime jog turime lojalius įmonei klientus (Martis, Simas ir Lina). Taip pat turime nelojalius (Rimas, Mindaugas ir Virgis).



Šaltinis: sukurta autoriaus

3 Pav. Lojalūs i nelojalūs klientai

Pateikus šiuos duomenis, apmokama sistema atrenka lojalius klientus ir analizuoja atskiriančius lojalius nuo nelojalių klientų požymius.



Lojalumas = "apvalios akys" IR „šypsena“

Šaltinis: sukurta autoriaus

4 Pav. Lojalumo išskyrimo išvados

Sistema randa bendrus bruožus ir juos apjungia. Šiuo atveju daroma išvada, jog lojalumas tai yra šypsena ir apvalios akys. Realiai vietoj akių ir šypsenos gali būti lankomumas, perkamumas,

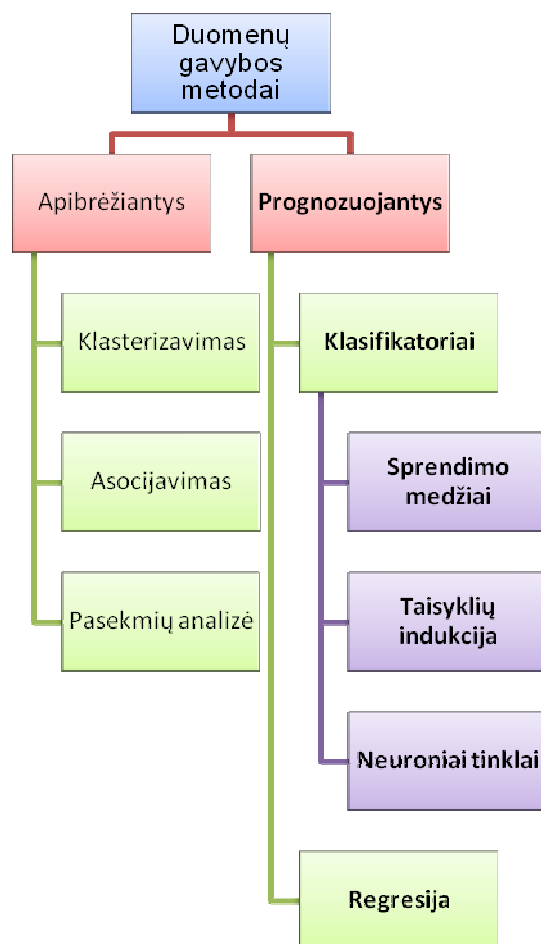
apsilankymo dažnis, CLV, ir kiti požymiai. Svarbiausia, kad sistema juos apjungia į viena požymi, priskiriant svorius kriterijams.

Duomenų gavyboje yra nemažai metodų bei įrankių tokiai analizei atlikti. Sekančiame poskyryje trumpai apžvelgsime ir palyginsime juos.

2.2 Duomenų gavybos metodai.

Duomenų gavybos metodai bendrai skirstomi į apibrėžiančius ir prognozuojančius. Kadangi kliento analizė ankstyvoje ryšio su įmone stadijoje yra prognozuojančio pobūdžio, detaliau nagrinėsime prognozuojančius metodus. Kadangi lojalių klientų atpažinimo uždavinys yra ne apibrėžiančio pobūdžio (asociacijos ir klientų grupavimas nebūtų naudingas), prognozuojančius apžvelgsime detaliau

5 Pav. Yra pavaizduoti metodai, įeinantys į prognozuojančių duomenų gavybos metodų šeimą. Tai yra klasifikatoriai: sprendimo medžiai, taisyklių indukcija ir neuroniniai tinklai; bei regresija.



Šaltinis: sukurta autoriaus

5 Pav. Duomenų gavybos metodų tipai

Toliau apžvelgsime statistinius paketus ir šiuos duomenų gavybos metodus

2.2.1 Statistiniai paketai

Pastaruoju metu beveik visose naujausiose statistinių paketų versijose kartu su tradiciniais statistiniais paketais yra įtraukiami duomenų gavybos elementai. Tačiau pagrindinis jų dėmesys skiriamas klasikinėms metodikoms- koreliacinei, regresinei, faktorinei (ir kitoms) analizėms. Šių sistemų pagrindinis trūkumas yra vartotojo specialaus paruošimo reikalavimai. Be to pabrėžtina, kad šiuolaikiniai statistiniai paketai yra per daug „sunkiasvoriai“ ir dėl to netinka masiniam taikymui finansuose ir versle. Vienas svarbiausių statistinių paketų trūkumų, dėl kurių jie neįtraukiami į duomenų gavybos sistemas, yra tai, kad dauguma metodų, įeinančių į paketo sudėtį remiasi statistinę paradigmą, kuri pasižymi atrinkimo charakteristikų suvidurkinimu. O realiose gyvenimo fenomenuose šios charakteristikos dažnai yra fiktyvūs didžiai. Be to kadangi regresija ir kiti plačiai taikomi metodai iš prigimties skirti tiesinėms priklausomybėms modeliuoti, šie metodai netinka duomenų apie klientus analizei, kadangi tarp lojalumo ir kliento požymių yra ne tiesinė priklausomybė.

2.2.2 Sprendimų medžiai

Sprendimų medžiai yra bene pats populiariausias duomenų gavybos uždavinių sprendimas. Jie sudaro hierarchinę klasifikavimo „jeigu ... tai“ (if-then) taisyklių struktūrą medžio pavidale. Priimant sprendimą, kuriai klasei priskirti tam tikrą objektą arba situaciją, reikia atsakyti į klausimus, kurie yra šio medžio mazguose. , pradedant nuo šaknies. Klausimai turi „ar parametro A reikšmė didesnė nei X“ pavidalą. Jeigu atsakymas yra teigiamas, pereinama į dešinę sekančio lygio mazgą, o jei neigiamas- į kairįjį. Toliau seka kito mazgo(prie kurio pereita) klausimas.

Šio metodo populiarumas susijęs su vaizdumu ir aiškumu. Priimtą sprendimą lengva atsekti ir interpretuoti. Tačiau sprendimo medžiai nesugeba atrasti „geriausias“ (mažiausios paklaidos) priklausomybes arba taisykles duomenyse. Jos tik realizuoja naivią požymių sekos peržiūrą ir priskiria faktines realių priklausomybių dalis.

2.2.3 Taisyklių indukcija.

Taisyklių indukcija yra labai panaši į sprendimų medžius. Klasifikacijos taisyklė gali būti apibrėžta sekančiai: Duoti apmokymo pavyzdžiai (atvejai, kuriems klasifikacija yra žinoma). Rasti klasifikavimo taisyklių rinkinį, kuris gali būti naudojamas klasifikuojant naujus atvejus (tuos, kurie nebuvo pristatyti apmokymui) Naudojama speciali data description kalba, kurios pagalba suformuluota hipotezė įprogramuojama kaip taisyklių rinkinys.

2.2.4 Daugiaregresinė analizė

Daugiaregresinė analizė tai yra metodas, kuris įvertina ryšį tarp priklausomo kintamojo ir vieno ar daugiau nepriklausomų kintamųjų. Duomenų apie klientus atveju daugiaregresinė analizė gali būti panaudota įvertinant klientų perkamumą atsižvelgiant į amžių, lytį, plaukų spalvą ir kitus kriterijus. Bendruoju atveju gautume:

$$Y=f(X_1,X_2,X_3,\dots X_n) \quad (1)$$

kur: Y = Klientų perkamumo galia;
X₁ amžius
X₂ lytis
X₂ kreipimosi dažnis
X₃ pirkimo dažnis
X_n n-tasis kintamasis

Norint sudaryti lygtį, reikia būtinai žinoti a priori lygties pavidalą. Paprastai to niekas nežino. Kadangi lygties sudarinėtojas turi tik pasirinkti regresijos kreivę, kuri labiausiai atitinka duotus duomenis, gautos lygties tinkamumas skaičiuojamas klaidų terminais. Dažnai sudarant lygtis daro palengvinančią prielaidą, kad duomenų struktūroje vyrauja tiesinė priklausomybė. O tai yra minusas, nustatant lūžio taškus. Kadangi duomenys apie klientus yra ne tiesio pobūdžio, duomenų eilutės labiau linkusios turėti ne tiesinius trendus. Duomenų eilutės taip pat gali būti chaotiškos.⁷ Neskaitant šių akivaizdžių trūkumų, daugiaregresinė analizė taip pat turi keletą metodologinių problemų⁸. Tai „<...>multi- kolinearumas (multicollinearity) ir sisteminis poslinkis (heteroscedasticity). Pav. ankščiau minėti kintamieji X₁ ir X_n=“apsilankymo dažnumas“ yra stipriai koreliuoti ir veda prie multi- kolinearumo. Tai savo ruožtu padaro lygtį nestabilia dėl nestabilių koeficientų.“⁹

Tradicinė daugiaregresinė analizė tam tikru laiko momentu gali apdoroti tik vieną priklausomą kintamąjį. Tarkim vadybininko tikslas- pardavimų bei grynųjų pinigų cirkuliacijos prognozė. Šiuo atveju abu atvejai turi būti tiriami atskirai. Savo darbe Proctor (Proctor, 1992) teigia kad šiuo atveju kanonė regresija gali išspręsti šią problemą, tačiau tai yra labai sudėtinga interpretuoti tokios analizės rezultatus ir „<...>vargu ar tokia sistema gali leisti prognozuoti.“¹⁰

2.2.5 Neuroniniai tinklai

Pagal paprasčiausią apibrėžimą neuroninis tinklas tai yra duomenų apdorojimo sistema susidedanti iš daugelio elementarių tai sistemai vienetų – ląstelių (vadinamų neuronais), kurios tarpusavyje jungiasi svoriniais ryšiais (sinapsė). Turėdamas daug panašumu su žmogaus smegenų

⁷ Thall, 1992m.

⁸ Proctor, 1992m.

⁹ V. Venugopal, W.Baets, 2002m., p. 32

¹⁰ Proctor (1992m.)

veikimo ypatumais, neuroninis tinklas paveldėjo tokias svarbias savybes kaip gebėjimą mokytis, mokėjimą apibendrinti išmoktas taisykles.

Pagrindinis neuroninių tinklų trūkumas yra tai, kad jų apmokymui reikia turėti didelį duomenų kiekį. Kitas trūkumas yra tai, kad netgi gerai treniruotas neuroninis tinklas yra kaip „juodoji dėžė“. Žinios, kurios užfiksuojamos kaip kelių šimtų neuronų ryšių svoriai, visiškai nepasiduoda analizei ir žmogaus interpretacijai. (Žinomi bandymai interpretuoti neuroninio tinklo nustatymus pvz. KINOsuite-PR yra neįtikintys)

2.2.6 Duomenų analizės metodų palyginimas

Žemiau esančioje 1 lentelėje yra pateikiami apžvelgtų metodų privalumai ir trūkumai. Atsižvelgiant į nagrinėjama sritį (duomenų apie klientus analizė, lojalumo tyrimas, lojalumo požymių išskyrimas ir prognozavimas) neuroniniai tinklai geriausiai atitinka analizuojamų duomenų prigimtį.

1 lentelė.

Duomenų analizės metodų palyginimas

Nr.	Metodo/Įrankio pavadinimas	Privalumai	Trūkumai
1.	Statistiniai paketai	Turi savyje klasikinius statistinius duomenų analizės metodus. Yra galimybė juos kombinuoti.	dauguma metodų, įeinančių į paketo sudėtį remiasi statistinę paradigmą, kuri pasižymi atrinkimo charakteristikų suvidurkinimu
2.	Sprendimo medžiai/ taisyklių indukcija	Vaizdumas ir aiškumas. Priimtą sprendimą lengva atsekti ir interpretuoti. Plačiai naudojamas daugelyje srityčių.	Nesugeba atrasti „geriausias“ (pilnesnės bei teisingės) priklausomybes / taisykles duomenyse realizuoja naivią požymių sekos peržiūrą ir priskiria faktines realių priklausomybių dalis
3.	Daugia regresinė analizė	Našus metodus apdorojant duomenis, kuriuose yra tiesinė priklausomybė. Maža paklaida, jeigu duomenis atitinka priorų regresijos modelį	Apdoroja duomenis pakeisdamas pradinius. Reikia žinoti a priori lygties pavidalą. Jei priklausomybė tarp duomenų yra ne tiesinė-daroma didžiulė paklaida nustatant lūžio taškus. Metodologinės problemos: multi- kolinearumas (multicollinearity) ir sisteminis poslinkis (heteroscedasticity) Tam tikru laiko momentu gali apdoroti tik vieną priklausomą kintamąjį. Duomenų trūkumas stipriai įtakoja analizės rezultatus. Lygtis skirtingai parenkama prognozavimui ir klasifikavimui.
4.	Neuroniniai tinklai	Nereikalauja a priori žinių apie duomenis. Klaidas toleruojantis mechanizmas. Duomenų trūkumas sintetiškai pašalinamas tinklu imituojant reikšmes. Vienas algoritmas gali būti panaudotas kaip prognozavimui taip ir klasifikavimui. Paklaida gali būti sumažinta padidinus iteracijų skaičių apmokymo metu. Pakankamai lankstus apibendrinti bet kokią testinę funkciją.	Iteracijų skaičius tiesiogiai įtakoja laika užgaišta tinklo treniravimui. Nėra tiksliai apibrėžtas iteracijų skaičius bei konfigūracijos ypatumai sudarant/apmokant tinklą. Tai daroma bandymo būdu, lyginant gautas paklaidas

Tarp kliento lojalumo ir duomenų apie jį vyrauja ne tiesinė priklausomybė. Lojalumo sąvoka bei lojalių klientų vertinimo kriterijai priklauso nuo įmonės, jos iškeltų tikslų, lojalumo interpretavimo. Dėl duomenų nepakankamumo, atsiranda a- priorių žinių apie nagrinėjamą objektą (klientą) trūkumas. Iš anksto apibrėžti lojalumą sąlygojančius veiksnius neįmanoma.

Kadangi Neuroninis tinklas nereikalauja a priorių žinių apie duomenis, pakankamai lankstus apibendrinti bet kokią testinę funkciją galime daryti prielaida, kad šis metodas yra tinkamiausias duomenų apie klientus analizei.

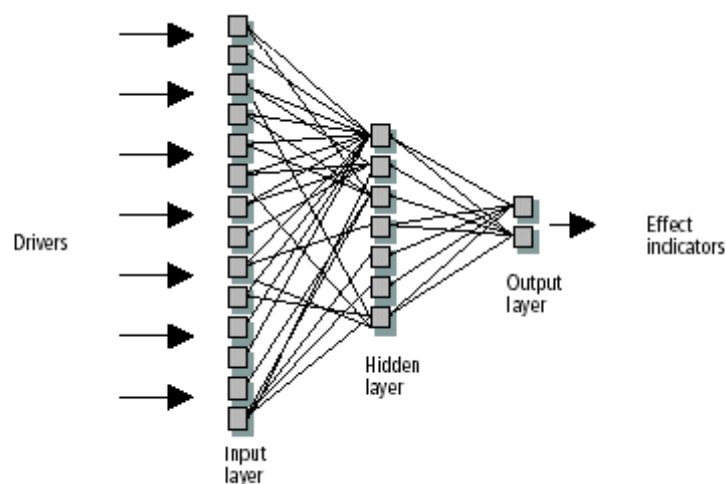
Toliau apžvelgsime neuroninių tinklų struktūrą bei duomenų analizės, naudojant neuroninius tinklus, eigą.

2.3 Neuroniniai tinklai

Pastaruoju metu pastebimas vis augantis interesas neuroninių tinklų naudojime sprendžiant plataus spektro problemas. Dėl jų apibendrinimo (prisitaikymo) savybės, jie plačiai naudojami medicinoje, nustatant pacientų diagnozes, plačiai taikomi vertybinių popierių bei akcijų kurso prognozavime.

2.3.1 Principinė struktūra

Pagal paprasčiausią apibrėžimą neuroninis tinklas tai yra duomenų apdorojimo sistema susidedanti iš daugelio elementarių tai sistemai vienetų – ląstelių (vadinamų neuronais), kurios tarpusavyje jungiasi svoriniais ryšiais (sinapsė). Turėdamas daug panašumu su žmogaus smegenų veikimo ypatumais, neuroninis tinklas paveldėjo tokias svarbias savybes kaip gebėjimą mokytis, mokėjimą apibendrinti išmoktas taisykles.



Šaltinis: Lars Grønholdt, Anne Martensen (2002m.)

6 Pav. "Neuroninio tinklo principinė struktūra."

Neuroninis tinklas sudaromas taip, kad neuronų skaičius įėjimo sluoksnyje lygus nepriklausomų kintamųjų skaičiui, o neuronų skaičius išėjimo sluoksnyje lygus priklausomų kintamųjų skaičiui. Pav.6 yra pavaizduota tokio tinklo principinė schema (Lars Grønholdt, Anne Martensen , 2002m.). Paslėptų sluoksnių bei neuronų juose skaičius gali būti pasirenkamas atsitiktinai arba laisvai. Neuroninis tinklas su bent vienu paslėptu sluoksniu naudoja apdoroja duomenis, paduodamus per įėjimo sluoksnį ir suformuoja vidinę ryšių reprezentaciją tarp nepriklausomų kintamųjų taip, kad visas procesas nereikalauja priorų žinių apie parametrų pasiskirstymą ir priklausomybes tarp jų. Šis mechanizmas toleruoja klaidas, kadangi iš jų mokinasi. Naudojant neuroninius tinklus rezultatai yra žymiai geresni, kai ryšiai tarp kintamųjų neatitinka priimta šabloną arba modelį. (regresinės analizės atveju- tai regresijos kreivės tipas). Taip pat pastebėta kad neuroninis tinklas su bent vienu paslėptu sluoksniu yra pakankamai lankstus apibendrinti bet kokią tęstinę funkciją.

Paslėptas (vidutinis) sluoksnis yra dažnai apibrėžiamas kaip „savybių nustatytojai ir apibendrintojai“¹¹, leidžiantys tinklui daryti pagrįstas išvadas. Per didelis kiekis neuronų paslėptame sluoksnyje sukelia problemą- gaunamas tinklas įsimena įėjimo duomenis ir nebesugeba daryti išvadas. Priimta kad neuronų skaičius šiame sluoksnyje privalo būti apie 75% įėjimo kintamųjų skaičiaus.

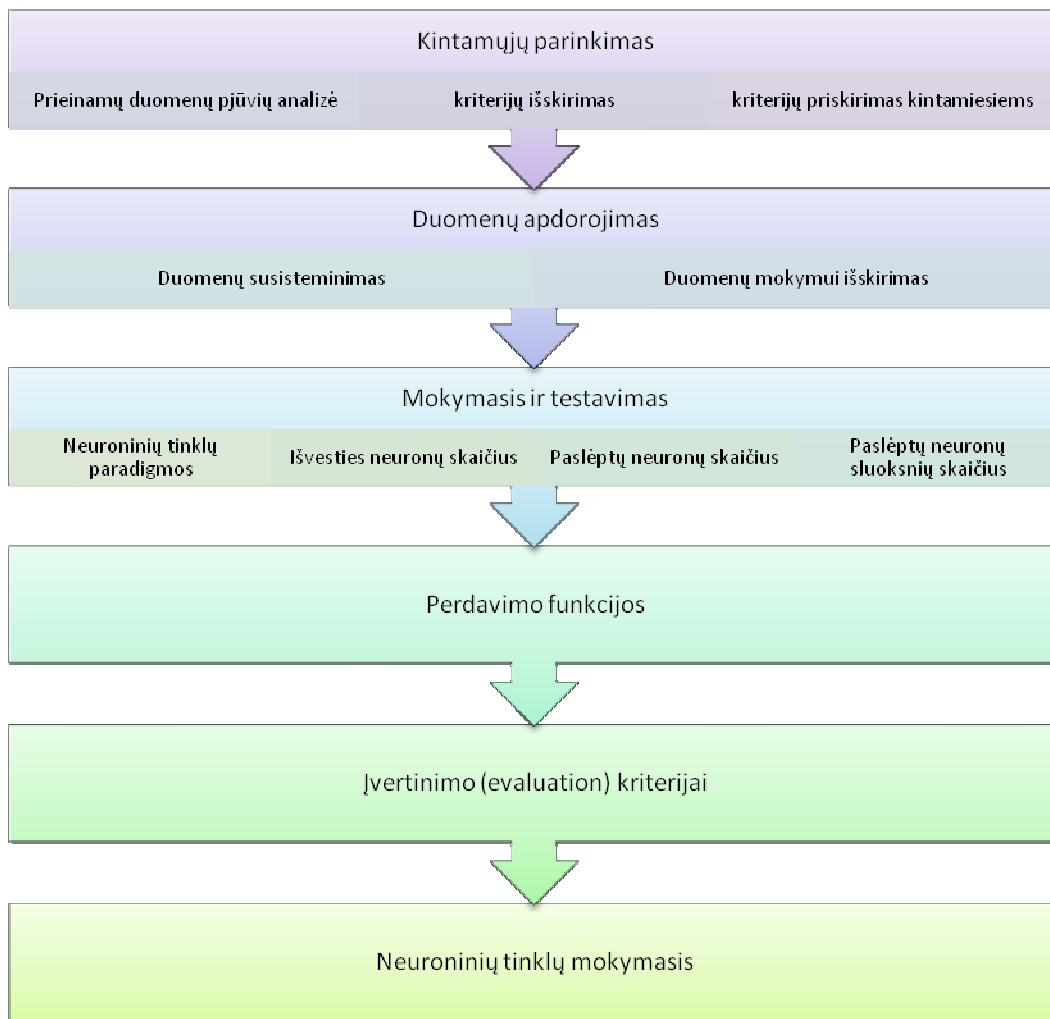
Toliau apžvelgsime Neuroninių tinklų taikymo specifiką.

2.3.2 Duomenų analizė naudojant Neuroninius tinklus.

Yra įvairiausių neuroninių tinklų modelių, sukurtų finansiniam prognozavimui, tačiau jų kūrimo principai iš dalies yra vienodi. Gali skirtis tik naudojami algoritmai, neuronų skaičius, sluoksnių skaičius.

Dažniausiai yra naudojama septynių žingsnių metodologija patiekta (pav. 7)

¹¹ V. Venugopal, W.Baets, 2002m., p.33



Šaltinis: sukurta autoriaus

7 Pav. „7 žingsnių metodologija“

1. Kintamųjų parinkimas. Pirmasis žingsnis yra labai svarbus, kadangi svarbu žinoti, kurie kintamieji klientų lojalumo atžvilgiu yra svarbūs ir prieinami. Šiame etape būtina nuspręsti ar naudoti formalius ar fundamentalius įėjimo duomenis vieno ar daugelio klientų. Formalūs įėjimo duomenys apibrėžti kaip priklausomų kintamųjų ar indikatorių senesnės reikšmės, kurios buvo apskaičiuotos naudojantis šiek tiek senesnius duomenų, o fundamentalūs įėjimo duomenys - tai finansiniai ir kiti apskaičiuojami kintamieji, darantys įtaką priklausomiems kintamiesiems. Svarbu patikrinti duomenų pjūvių prieinamumą. (duomenų pjūvis arba dimensija- požymis, kuris sieja visus duomenis. (Pvz. kiekvieno kliento apyvarta, lytis, amžius ir t.t.). Beprasmiška imti tokius pjūvius, kuriuos turi ne visi duomenys.

Pasirenkami kriterijai turi atspindėti apskaičiuojamus faktorius, darančius įtaką lojalumui. Čia turi būti griežtai apibrėžta lojalumo sąvoka (kaip įmonė suvokia lojalų klientą) ir išskirti lojalumo požymiai arba tų požymių kombinacija.

Etapas užbaigiamas kriterijų priskiriu kintamiesiems.

2. Duomenų apdorojimas. Tai įvedamų ir išvedamų kintamųjų analizė ir transformacija, siekiant sumažinti „triukšmą“, nustatyti svarbiausius ryšius, aptikti kryptingumus (tendencijas).

Šiame etape esami duomenys sutvarkomi taip kad nei viena duomenų eilutė (duomenys apie konkretų klientą) nesiskirtų nuo kitos stulpelių atžvilgiu (kriterijų) Dažnai pasitaiko, kad tikrinant duomenis pastebimas tarpas tarp duomenų nuoseklumo, t.y. tam tikrų reikšmių nėra. Jei pastebima tokia klaida, trūkstamiems duomenims suteikiamos gretimų duomenų reikšmės. Yra ir kitas būdas išspręsti problemą – trūkstamiems duomenims priskirti vieną neuroną, kuriam suteikiama 1 reikšmė.

Norint panaikinti tiesinę duomenų kryptį, naudojami diferenciacija ir kintamojo natūrinis logaritmvimas. Techninei analizei gali būti naudojami slenkamieji vidurkiai, generatoriai, nepastovumo filtrai. Geri rezultatai gaunami naudojant kelis metodus iš karto. Taip sumažinamas kintamųjų dubliavimasis, o neuroniniai tinklai geriau prisitaiko prie kintamų rinkos sąlygų, aišku, nuolatos mokydami. Dažnai yra taikomas įvedamų ir išvedamų duomenų glotninimas naudojant paprastąjį arba eksponentinį slenkamųjų vidurkių metodą.

Išskiriant duomenis apmokymui, svarbu parinkti tokius egzempliorius, kurie aiškiai įmonės supratimu yra lojalūs. Svarbu parinkti kuo daugiau tikslių pavyzdžių, remiantis kuriais neuroninis tinklas galėtų apsimokyti. Svarbu žinoti, jog kuo daugiau duomenų, tuo mažesnę paklaidą darys apmokytas tinklas.

3. Mokymasis ir testavimas. Neuroninių tinklų apmokymus galima apibūdinti kaip jų sugebėjimą atpažinti duomenų struktūras. Testavimas yra naudojamas įvertinti apibendrinimo galimybes tariamai apmokyto neuroninio tinklo. Patvirtinimas – tai paskutinis sistemos įvertinimas, pripažinimas, kad sistema įvykdė uždavinius ir yra gauti norimi rezultatai.

Pats geriausias ir tiksliausias testavimo būdas yra „Ėjimas pirmyn“ (*walk-forward*), taip pat žinomas kaip slenkamųjų ar judančių langelių testavimas.

Dažnas apmokymas leidžia neuroniniam tinklui greičiau adaptuotis prie besikeičiančių kriterijų reikšmių

4. Neuroninių tinklų paradigmos. Neuroninių tinklų sudėčiai nusakyti naudojami 2 terminai: neurodinamika ir architektūra. Neurodinamika apibūdina atskiro neurono savybes, tokias kaip perdavimo funkcija ir kaip apjungiami įvedami duomenys. Neuroninių tinklų architektūra nusako jo struktūrą, tame tarpe neuronų skaičių kiekviename sluoksnyje ir tarpusavio ryšių skaičių.

Paslėptų sluoksnių skaičius. Paslėpti sluoksniai leidžia neuroniniam tinklui apibendrinti duomenis. Tinklas, turinti vieną paslėptą neuronų sluoksnį gana gerai sugeba apskaičiuoti bet kokią tolydžią funkciją. Sluoksnių skaičiaus didinimas nereiškia, kad prognozė bus tikslesnė. Tiesą sakant, sluoksnių didinimas taip pat padidina funkcijos apskaičiavimo laiką ir gali sukelti

nesuderinamumo pavojų. Tokiu atveju prognozės rezultatai būtų visiškai netikę. Taigi norint pasiekti kuo geresnių rezultatų geriau naudoti vieno arba dviejų paslėptų neuronų sluoksnių tinklą.

Paslėptų neuronų skaičius. Nepaisant to, kad neuronų skaičiaus parinkimas yra svarbus faktorius kuriant neuroninį tinklą prognozavimui, nėra nustatyta, koks turėtų būti neuronų skaičius. Tačiau yra keletas pasiūlymų, tokių skaičių galima būtų parinkti pvz. 3 –jų sluoksnių tinklui. Jei turėtume n įvedimo neuronų ir m išvedimo neuronų, tai neuronų skaičius tinkle būtų $\sqrt{n * m}$. Taip pat, neuronų skaičius tinkle gali būti 75% nuo įvedamų neuronų skaičiaus. Kitas būdas - kaskart testuojant neuroninį tinklą padvigubinti neuronų skaičių ir taip daryti iki tol, kol testavimo rezultatai pablogės. Taip pat manoma, kad neuronų skaičius turėtų būti nuo 1/3 iki 3 kartų didesnis nei įvedamas neuronų skaičius. Šiuo atveju akivaizdu, kad nėra vieningos nuomonės koks turėtų būti neuronų skaičius. Optimalus variantas gali būti pasiekiamas tik eksperimentiniu būdu su skirtingais neuronų skaičiais sistemose.

Išvedimo neuronų skaičius. Išvedimui dažniausiai naudojamas vienas neuronas. Tam yra svarbios priežastys. Neuroniniai tinklai, turintys daugiau išvedimo neuronų pateiks labiau supaprastintus duomenis nei tinklai su vienu išvedimo neuronu. Neuroninis tinklas, mokydamasis, vidutinė paklaidą paskirsto ant visų išvedimo neuronų vienodai, taip paklaida yra sumažinama, todėl prognozės rezultatai nėra tikslūs.

5. Perdavimo funkcijos. Perdavimo funkcijos yra matematinės formulės, kurios apibrėžia apdorojamo neurono išvedimą. Jos taip pat gali būti pavadintos transformacija, aktivizacija, arba tai gali būti pradinės funkcijos. Dauguma dabartinių neuroninių tinklų naudoja sigmoidinę (S-formos) funkciją. Perdavimo funkcijos tikslas yra apriboti išvedamų duomenų dydžius, kurie gali paralyžuoti neuroninius tinklus ir tokiu būdu nutraukti neuroniniu tinklų mokymosi procesą.

6. Įvertinimo (pripažinimo, angl. *evaluation*) kriterijai. Dažniausiai neuroniniuose tinkluose naudojama paklaidų minimizavimo funkcija yra kvadratinių paklaidų suma. Taip pat gali būti naudojamos absoliučios deviacijos (nukrypimo), procentų skirtumų funkcijos. Tačiau šios funkcijos nėra paskutinis žingsnis įvertinant sistemos tinkamumą, kadangi kiti prognozavimo įvertinimo metodai, tokie kaip MAPE, paprastai nėra minimizuojami neuroniniuose tinkluose.

Lojalių klientų atpažinimo sistemose neuroninių tinklų prognozės yra paverčiamos į ne lojalus/lojalus signalus. Pavyzdžiui, visos prognozės didesnės nei 0.8 ar 0.9 gali būti traktuojamos kaip lojalumas, o visos prognozės mažesnės nei 0.2 ar 0.1 – nelojalumo.

7. Neuroninių tinklų apmokymas. Neuroninis tinklas su įvestais parametrais negali prognozuoti. Tinklo apmokymas spręsti problemas vyksta nuolatos apmokant tinklą. Neuroniniam tinklui yra parodomi duomenys, tada jis prognozuoja rezultatą. Pirmos prognozės turi dideles paklaidas. Kai tinklas yra nuolat pataisomas. Sekančios prognozės yra tikslesnės. Kaip ir žmogus, neuroninis tinklas apibendrina duomenis. Taip sumažinama paklaida sekantį kartą prognozuojant

duomenis. Reikia nustatyti kokią informaciją suteikti neuroniniam tinklui, o šis atpažįsta struktūras, tendencijas ir paslėptus ryšius tarp elementų.

Neuroninių tinklų apmokymo esmė yra surasti tokius svorius neuronams, kad galima būtų minimizuoti paklaidos funkciją. Neuroniniai tinklai mokymuisi naudoja laipsniško mažinimo algoritmą. Šio algoritmo esmė yra svorių neuronams parinkimas, kad būtų galima nusileisti prie žemiausio paklaidų paviršiaus taškų. Globalaus minimumo radimas ne visada pavyksta, kadangi paklaidos paviršius gali turėti daugybę vietinių minimumų. Taip algoritmas gali tiesiog „pakibti“.

Vyrauja keletas nuomonių, kada reikėtų nutraukti apmokymus. Pirma, jei papuolama į aklavietę vietiniame minimume ir nerandama globalaus minimumo. Apmokymas turėtų baigtis tik tada, kai naudojant atsitiktinai paimtus svorius, nesimato jokio paklaidos funkcijos pagerėjimo. Tokia neuroninių tinklų būklė vadinama konvergencija. Kita nuomonė, kada reiktų baigti tinklų apmokymus, yra apmokymo-testavimo pertraukimas. Apmokymai nutraukiami po tam tikro bandymų skaičiaus. Kai apmokymai pertraukiami, testuojama ir įvertinama neuroninio tinklo galimybė apibendrinti duomenis.

Informacijos išgavimas iš duomenų yra sudėtinis procesas, kurio komplikotumas priklauso nuo norimos išgauti informacijos pobūdžio. Šiame skyriuje apžvelgtas informacijos išgavimo iš duomenų procesas, metodai, ir išnagrinėtas jų tinkamumas lojalumo tyrimo uždaviniams spręsti. Tarp kliento lojalumo ir duomenų apie jį vyrauja ne tiesinė priklausomybė. Lojalumo sąvoka bei lojalių klientų vertinimo kriterijai priklauso nuo įmonės, jos iškeltų tikslų, lojalumo interpretavimo. Dėl duomenų nepakankamumo, atsiranda a- priorių žinių apie nagrinėjamą objektą (klientą) trūkumas. Iš anksto apibrėžti lojalumą sąlygojančius veiksnius neįmanoma.

Kadangi Neuroninis tinklas nereikalauja a priorių žinių apie duomenis, pakankamai lankstus apibendrinti bet kokią testinę funkciją galime daryti prielaida, kad šis metodas yra tinkamiausias duomenų apie klientus analizei.

Minėtam uždaviniui realizuoti labiausiai tinka neuroninis tinklas kadangi:

- *Mes neturime žinių apie priklausomybę tarp kliento elgesio bendradarbiavimo stadijoje ir lojalumo, nelojalumo statuso užlaikymo stadijoje. (Neuroninis tinklas nereikalauja a priorių žinių apie duomenis.)*
- *Mokosi iš savo klaidų: Susigrąžinimo etapo rezultatas, kuomet mes žinome post-factum kuris klientas ištikusių lojalus o kuris ne-gali būti pateiktas tinklui kaip papildoma mokymosi imtis. Taip kiekvieną kartą tinklo klasifikavimo tikslumas didės. (Klaidas toleruojantis mechanizmas.)*

- Užpildo duomenų „spragas“: Jei apie klientą surinkta nevisapusiška informacija: nepavyko identifikuoti atliekant pirkimo operacijas ir pan. (Duomenų trūkumas sintetiškai pašalinamas tinklu imituojant reikšmes.)
- Iš esmės atitinka darbo uždavinį. Tai klasifikavimas ateičiai, kitaip tariant prognozavimas. Vienas algoritmas gali būti panaudotas kaip prognozavimui taip ir klasifikavimui

Yra įvairiausių neuroninių tinklų modelių, sukurtų finansiniam prognozavimui, tačiau jų kūrimo principai iš dalies yra vienodi. Formuojant duomenų analizės modelį neuroninių tinklų pagalba, naudojama 7ių žingsnių metodologija:

1. Kintamųjų parinkimas
2. Duomenų apdorojimas
3. Mokymasis ir testavimas
4. Neuroninių tinklų paradigmos
5. Perdavimo funkcijos
6. Įvertinimo (pripažinimo, angl. evaluation) kriterijai
7. Neuroninių tinklų apmokymas

„Neuroniniai tinklai negali atlikti nieko tokio, ko negali kiti metodai“¹², tačiau jie gali sutaupyti laiką ir pinigus ieškant priklausomybių tarp Klientų lojalumo ir kiekybinių kliento vertinimo rodiklių.

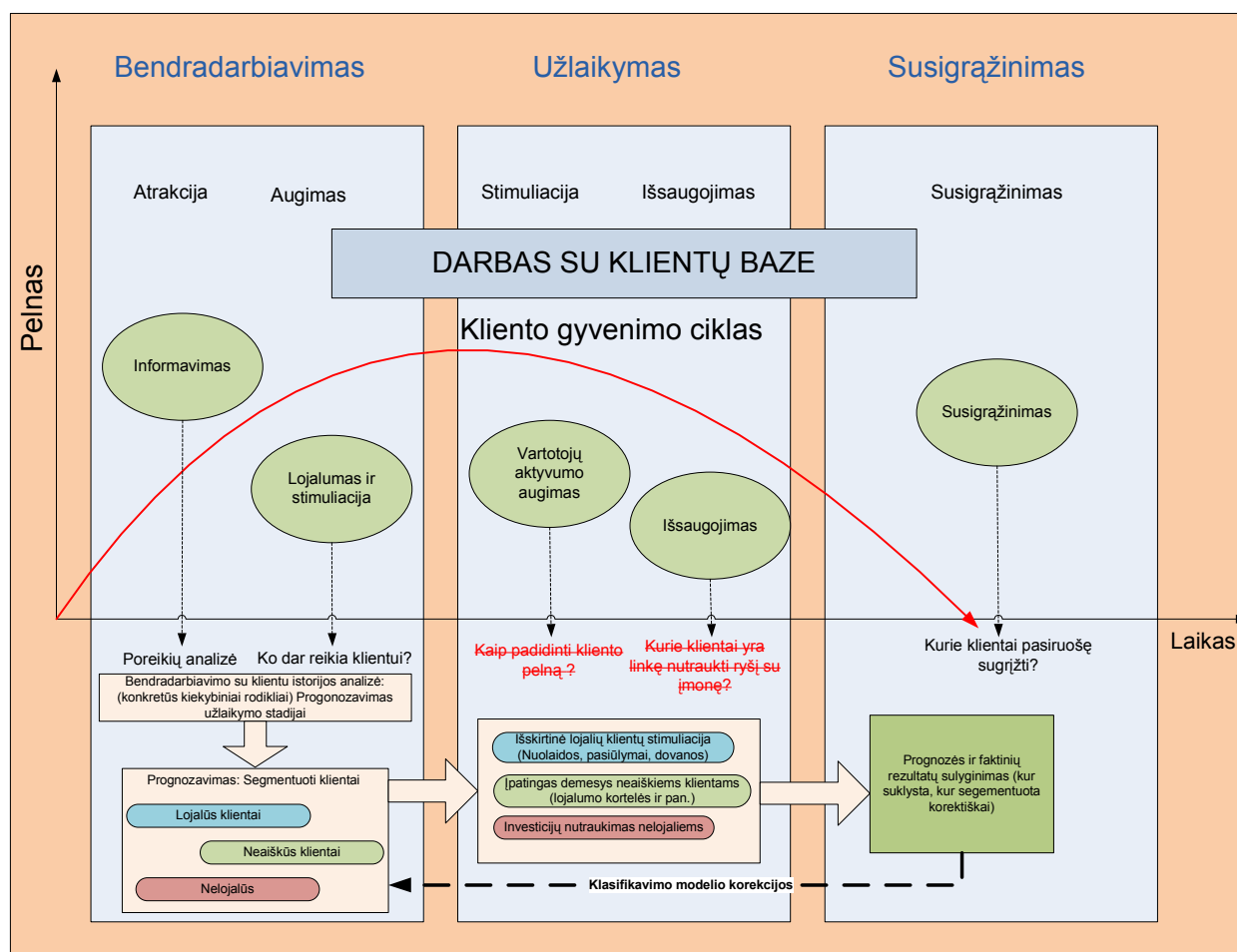
Remiantis klientų lojalumo tyrimo metodais bei parinktu duomenų analizės metodu, sekančiame skyriuje sudarysime kliento lojalumo tyrimo modelį.

12 (Habil. dr. Rimvydas Simutis, Neuroninių tinklų paskaitos, 2006)

3. Klientų lojalumo tyrimas neuroninių tinklų pagalba

Kaip jau buvo minėta, E.Ksevelonakino pasiūlytas modelis, turi pagrindinį trūkumą: klientai, linkę atsitraukimu nustatomi užlaikymo stadijos pradžioje. Iš čia išplaukia poreikis klientus klasifikuoti užlaikymo stadijos pradžia. Kitaip tariant- remiantis bendradarbiavimo stadijoje surinktais duomenimis nustatyti klientus, kurie užlaikymo stadijoje išliks lojaliais arba atsitrauks (nelojalūs). Tai iš esmės prognozavimo/ klasifikavimo uždavinys, kuriam labiausiai tinka neuroninis tinklas (pagal metodų apžvalgos rezultatus.)

Sekančiame 8 Pav. Pateiktas E.Ksevelonakino modelis ir klasifikavimo poreikio modelis.



Šaltinis: sukurta autoriaus pagal E.Ksevelonakino modelį

8 Pav. Vartotojo gyvenimo ciklas ir klasifikavimo poreikio modelis.

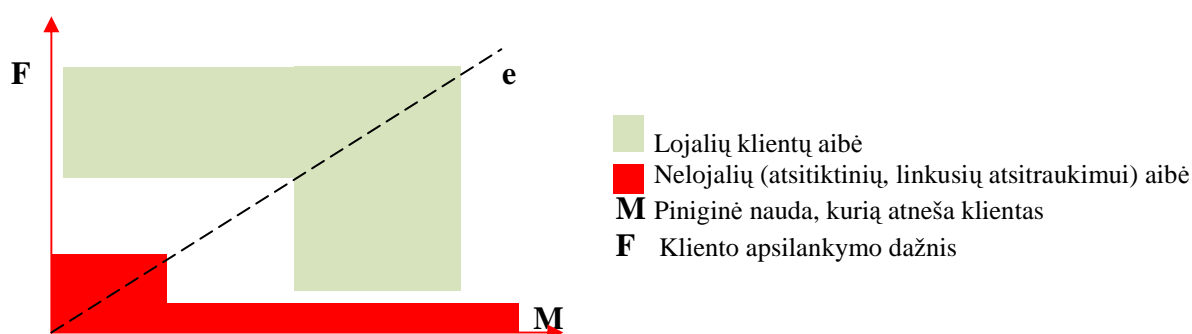
Bendradarbiavimo su klientu stadijoje, surenkama informacija apie jį. Šios stadijos pabaigoje atliekamas klasifikavimas. T.y. sistema atrenka kuris klientas sekančiame laiko periode (užlaikymo stadijoje) bus Lojalus, neaiškus arba nelojalus.

Užlaikymo stadija: priklausomai nuo klasifikavimo rezultatų, įmonė gali taikyti tikslines priemones (pvz. Lojaliems klientams pasiūlyti pažintinę kelionę į Budapeštą, neaiškiems klientams pasiūlyti lojalumo korteles, o nelojaliems klientams nutraukti „privilegijas“) T.y. vietoj klientų

poreikio nustatymo užlaikymo stadijoje, įmonė turi būti pasiruošus taikyti savo iš anksto paruošas priemones atitinkamai kliento klasei.

Susigrąžinimo etapas: atlikus klasifikavimą ir pritaikius tikslesnes priemones, analizuojami gauti rezultatai: kokie klientai išliko lojaliais, kurie prarado lojalumo statusą, t.y. buvo suklysta. Kitaip tariant sulyginamas faktinis klientų statusas su klasifikavimo metu pasiūlytu. Susigrąžinimo stadijoje įmonė taip pat gali palyginti prognozuotas reikšmes ir faktinį klientų lojalumą. Turint „korektiškai „ ir „klaidingai“ klasifikuotų atvejų aibes, galima patikslinti prognozavimo modelį, įtraukiant naujai gautas faktines reikšmes į patyrimą. Tai iteracinis procesas, kurio metu modelis tampa vis tikslesnis.

Norint sudaryti tokį modelį, visų pirmą reikia atrinkti lojalius klientus. Tai galima padaryti remiantis klasikiniiais metodais (Parėto 20/80) arba subjektyviai atrenkant lojaliu klientų aibę (ir kt.)



Šaltinis: sukurta autoriaus

9 Pav. „Lankomumas, piniginė vertė ir lojalumas“

Kaip matome iš pav. 8 lojalūs klientai dažnai lankosi ir dažnai perka, o nelojalūs – atvirkščiai mažai lanko ir mažai perka.

Ar pirkimas apsilankymo dažnis yra didelis nusakomas kiekvienos tariamos įmonės atskirai, kadangi tai priklauso nuo įmonės veiklos specifikos, jos masto ir t.t. Kitaip tariant tai yra subjektyvūs rodikliai.

Atrinkus lojalius klientus, reikia apskaičiuoti jų patrauklumą įmonei apibūdinančius rodiklius už praeitą periodą. Kitaip tariant turime kiekybiškai įvertinti įmonės bendradarbiavimą su tuo klientų. Šioje stadijoje galime apskaičiuoti tokius rodiklius¹³:

- Kliento pirkimai per pastarąsias 5 kalendorines mėnesio dienas
- Kliento pirkimai per pastarąsias 10 kalendorinių mėnesio dienų
- Suminiai kliento pirkimai per nagrinėjamą mėnesį
- A kategorijos prekių (aukšta kaina, mažas pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- B kategorijos prekių (vidutinė kaina, vidutinis pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- C kategorijos prekių (žema kaina, didelis pozicijų skaičius) pirkimai per nagrinėjamą mėnesį

¹³ Wouter BUCKINX, Geert VERSTRAETEN, Dirk Van den POEL, (2007) „Predicting customer loyalty using internal transactional database“

- D kategorijos prekių (labai žema kaina, didelis pozicijų skaičius) pirkimai per nagrinėjamą mėnesį
- Suminis nupirktų prekių per nagrinėjamą mėnesį skaičius
- Nupirktų A kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų B kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų C kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų D kategorijos prekių skaičius per nagrinėjamą mėnesį
- Paskutinio apsilankymo naujumas (dienų skirtumas po paskutinio apsilankymo)
- Apsilankymo naujumo standartinis nuokrypis (standartinis nuokrypis dienų skirtumo tarp apsilankymų)
- Vidutinis apsilankymo naujumas (vidutinis dienų skirtumas tarp apsilankymų)

Kita vertus pasak Bolton, Lemon, ir Verhoef, klientų lojalumas gali būti apibūdinamas CLV verte, kuri gali būti taip pat paskaičiuota šiame etape. Tai rodiklis, kuris skaičiuoja kliento (brangių/ pigių prekių) pirkimus, tų pirkimų pokyčius duotai bendravimo su klientu dienai. CLV galime apskaičiuoti pagal Bolton, Lemon pasiūlytą formulę (formulė 1). Skaičiavimai vykdomi kiekvienai pasirinktai laikotarpio atkarpai (pvz. Jei nagrinėjame metus- atliekame CLV skaičiavimus kiekvienam mėnesiui). Taip gaudami CLV pokyčio istorija per visą nagrinėjamą laikotarpį. Tai labiau primena akcijų kurso prognozavimo uždavinį.

Kadangi iš esmės mūsų uždavinys- klientų klasifikavimas, turime paskaičiuoti minėtus rodiklius ne tik lojalių klientų klasei, bet ir nelojalių. Taip pat labai svarbu atskirti dideles klaidos tikimybės zoną (t.y. klientai, kuriems remiantis esamais duomenimis negalime vienareikšmiškai nustatyti ar jie bus lojalūs ar ne.) Šią klientų klasę pavadinsime „neaiškiais“

Turėdami 3jų klasių egzempliorius (lojalūs, neaiškūs ir nelojalūs), o taip pat istoriškai paskaičiuotus rodiklius galima sudaryti neuroninį tinklą.

Kadangi CLV analizės pagrįstas klientų lojalumo vertinimas ir vertinimas remiantis kliento patrauklumo įmonei analizės nėra palygintos mokslinėje literatūroje, eksperimentinėje dalyje atliksime 2 eksperimentus ir ištirsime jų tinkamumą klientų lojalumo vertinimo uždaviniui spręsti.

4. Lojalių klientų atpažinimas remiantis CVL pokyčio analize

CLV pokyčio analizės eksperimento eigoje planuojama išanalizuoti klientų lojalumo požymius. Tam eksperimentas buvo suskaidytas į sekančias pagrindines dalis:

- Duomenų surinkimas: surenkami pardavimų duomenys už 2 mėnesius
- Duomenų apdorojimas, paruošimas eksperimentui: surinkti duomenys apdorojami formuojant 2 eksperimentines aibes: Pirmo mėnesio pardavimai ir antro mėnesio pardavimai. Abi aibės skaidomos į lojalių klientų pirkimų istorijas ir nelojalių klientų pirkimų istorijas.
- Eksperimentinės aplinkos (duomenų analizės įrankių) paruošimas. Duomenų analizė
- Gautų ir kontrolinių rezultatų sulginimas, apibendrinimas.

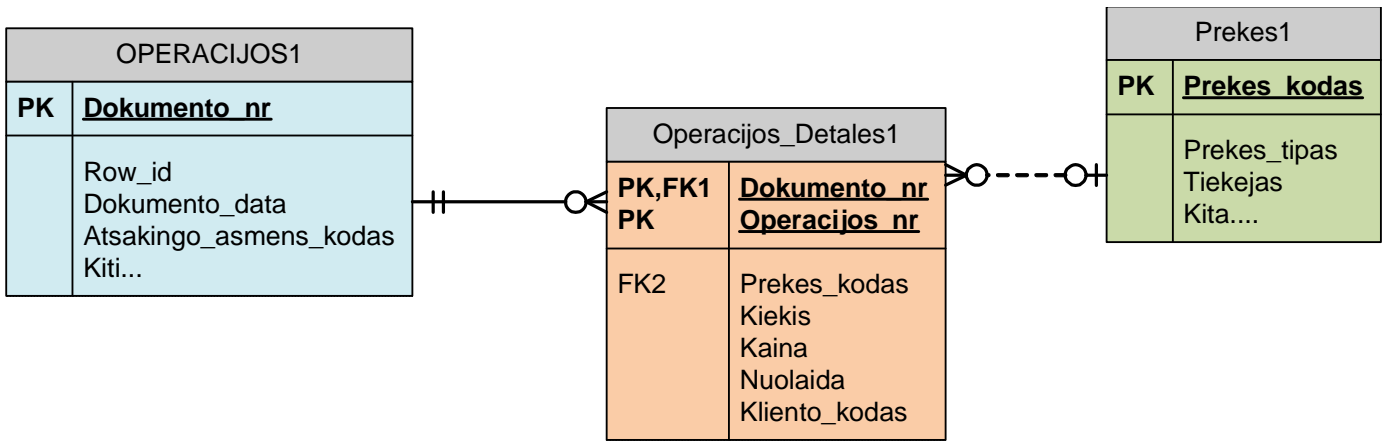
Eksperimento apribojimai:

Kadangi eksperimentuojama su realiais pardavimų duomenimis šiame darbe neskelbiamas įmonės pavadinimas. Siekiant užtikrinti konfidencialumą, eksperimento duomenys buvo koduojami, o kainos buvo pakeistos išlaikant priklausomybę tarp operacijų duomenų (visoms operacijų kainoms buvo parinktas bendras daugiklis)

Dėl didelio duomenų kiekio, atmetos fiskalinių kasų (POS), nutolusių padalinių ir firminių parduotuvių pardavimo operacijos

4.1 Duomenų surinkimas

Visi pardavimo operacijų duomenys randasi duomenų serveryje, apskaitos sistemos bazėje. Kaip galime matyti iš 9 pav., tiksliniai duomenys randasi 3 duomenų bazės lentose. Tai Operacijų lenta, detalių eilučių lenta (detalizuoja operacijas) ir prekes, kurios yra kiekvienos operacijos sudedamoji dalis. Atsižvelgiant į eksperimento apribojimus, buvo numatyti duomenų traukimo parametrai ir tarpinės duomenų keitimo procedūros. Duomenų užklausa suformuota TSQL kalba (PRIEDAS NR 3).



Šaltinis: sukurta autoriaus

10 Pav. Duomenų bazės schema: Duomenų šaltinių lentos

Kaip rezultatą gavome vieną lentą - pardavimų duomenis už numatytą laikotarpį. Jos struktūra ir duomenų pavyzdžiai pateikti lentelėje 2.

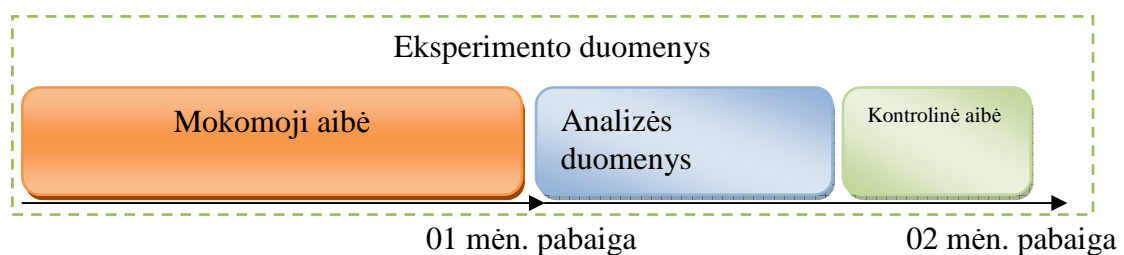
2 lentelė.

Eksperimento duomenys: pardavimų operacijos

Dokumento numeris	Dokumento data	Operacijos numeris	Klientas	Prekes kodas	Kaina	Kiekis	Nuolaida
DK0000000000001	2007.06.09	OP0000000001	KLIJ00001	PREKE1	1.15	9	1
DK0000000000001	2007.06.09	OP0000000002	KLIJ00001	PREKE2	999.99	1	3
DK0000000000001	2007.06.09	OP0000000003	KLIJ00001	PREKE3	12.35	20	0
DK0000000000001	2007.06.09	OP0000000004	KLIJ00001	PREKE4	33.20	1	0
DK0000000000007	2007.06.15	OP0000000009	KLIJ00002	PREKE1	1.15	15	0
***	***	***	***	***	***	***	***

Šaltinis: sukurta autoriaus

Šioje lentelėje yra 2 mėnesių pardavimai prekių lygyje. Ir iš ryšių tarp duomenų bazės lentų ir iš gautos lentos galime pastebėti, jog dokumento numeris (atstovaujantis unikalų apsilankymą) apima kelias operacijas su tam tikru klientu. T.y. per vieną apsilankymą gali būti kelių skirtingų prekių pirkimai..



Šaltinis: sukurta autoriaus

11 Pav. „Eksperimento duomenų sudėtis“

Eksperimentinių duomenų skaidymo dalys yra pavaizduotos 10 pav. Pagal neuroninių tinklų mokymo reikalavimus, iš šios lentos reikia atrinkti (mokymo ir kontrolinės aibės) Tam atrenkame:

- Lojalius pirmo mėnesio klientus ir sudaryti jų pirkimų istoriją dienos lygyje lentą. Formuojant lentą skaičiuojame UP, DW ir CLV dienos pokyčius (pagal formulę 1). Suformuojame lentą „Lojalumo_istorija“.
- Nelojalius pirmo mėnesio klientus ir sudaryti jų pirkimų istoriją dienos lygyje (Priedas NR 3). Formuojant lentą skaičiuojame UP, DW ir CLV dienos pokyčius (pagal formulę 1). Suformuojame lentą „Nelojalumo_istorija“.
- Lentos apjungiamos į bendrą pavadinimu „mokymui“ priskiriant lojalumo/nelojalumo požymius. Lentos dydis sudaro 1011 klientų penkiolikos dienų pirkimo istorijos CLV ir lojalumo požymiai

Šios lentos sudaro mokomąją aibę. Norint gauti analizes duomenis ir kontrolinę aibę, atliekami tie patys žingsniai tik atrenkama antro mėnesio duomenys:

- Atrenkami lojalūs antro mėnesio klientai ir sudaroma jų pirkimų istoriją dienos lygyje (PRIEDAS NR 3) Formuojant lentą skaičiuojame UP, DW ir CLV dienos pokyčiai (pagal formulę 1). Suformuojame lentą „kontrolėi“
- Atrenkami nelojalūs antro mėnesio klientai ir sudaroma jų pirkimų istorija dienos lygyje (Priedas NR 3). Formuojant lentą skaičiuojame UP, DW ir CLV dienos pokyčiai (pagal formulę 1). Suformuojame lentą „Kontrolėi“
- Lentos apjungiamos į bendrą pavadinimu „kontrolėi“, priskiriant lojalumo/nelojalumo požymius. Lentos dydis sudaro 1399 klientų penkiolikos dienų pirkimo istorijos CLV ir lojalumo požymiai periodo pabaigoje

Pabrėžtina, kad lojalumo požymiai atstovaus neuroninių tinklų išeities t.y. tikslų aibę pirmam periodui ir kontrolinius duomenis antram. Jos atskirai bus atrenkamos iš bendros aibės Matlab įrankio pagalba. Suformuoti lentų duomenys eksportuojami į tarpinius failus ir importuojami į Matlab aplinką duomenų apdorojimui ir analizei.

4.2 Duomenų paruošimas eksperimentui, apdorojimas.

Importuoti duomenys į matlab aplinką buvo sukarpomi taip, kad lojalumo požymiai būtų atskirti kaip nepriklausomos aibės (PRIEDAS NR 2). Galiausiai gavome duomenų masyvus:

Treniravimui(15x1010) Mokymo aibės dalis, skirta tinklui apmokyti

Išėjimui(1x1010) Mokymo aibės dalis, skirta tinklo tikslui apibrėžti (nurodoma kurie lojalūs kurie ne)

Analizei(15x1010)Atstovauja kontrolinę aibę, skirta tinklo gebėjimui pritaikyti“ išmoktas žinias“patikrinti.

Kontrolei(1x1010)Atstovauja kontrolinės aibės dalį. Skirta patikrinti kiek tikslios yra tinklo išvados.

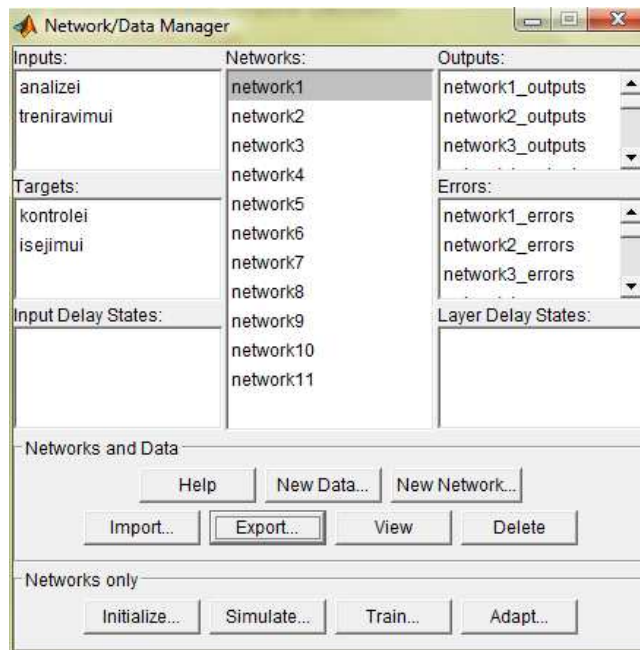
Turėdami omenyje, kad duotuoju laiko momentu (diena d1) turime tik vieną parametą (CLVd- kliento gyvenimo trukmės vertė duotuoju momentu), bet vienam klientui turime 15os dienu parametro kitimo istoriją formuojame neuroninius tinklus kurių įėjimo kintamųjų skaičius yra 15, o išėjimo-1, kadangi tik vienas požymis apsprendžia kliento lojalumą: 0- nelojalus, 1- lojalus.

4.2.1 Neuroninių tinklų formavimas:

Neuroniniuose tinkluose „Reikalingas sluoksnių skaičius ir elementų skaičius juose nėra nusakomas vienareikšmiškai. Didinant sluoksnių skaičių ir neuronų skaičių juose, didėja tinklo sudėtingumas. Kaip jau minėta, reikalingų neuronų skaičių nusako ne įėjimo erdvės matavimų skaičius, tačiau problemos sudėtingumas. <...> Jau dviejų sluoksnių neuroninis tinklas gali aproksimuoti bet kokią tolydinę funkciją. Tačiau dažnai, didinant sluoksnių skaičių, tinklo apimtys (neuronų skaičius, svorių skaičius) mažėja.“¹⁴

Matlab įrankio NNTOOL pagalba formuojami 11 Neuroninių tinklų su skirtingais neuronų skaičiais ir numatomos skirtingi mokymo algoritmai Priklausomai nuo testavimo rezultatų bus parinkta optimaliausia struktūra.. Šių neuroninių tinklų principinė schema pateikta (PRIEDE 1). Importuoti duomenys į Matlab aplinką ir sukurti neuronai tinklai pavaizduoti 11 pav.

¹⁴ Habil. dr. Rimvydas Simutis, Neuroninių tinklų paskaitų medžiaga, 2006.



Šaltinis: sukurta autoriaus

12 Pav. Neuroniniai tinklai ir importuoti duomenys matlab aplinkoje

Sekantis eksperimento žingsnis- suformuotų neuroninių tinklų apmokymas.

4.2.2 Neuroninių tinklų mokymas, kontrolinės aibės analizė

„Neuroninio tinklo apmokymo tikslas nėra tiksliai atstovauti mokymo duomenų imtį. <..> Neuroninio tinklo apibendrinimo savybės labai priklauso nuo modelio sudėtingumo ir duomenų kiekio mokymo imtyje. Modelis turintis mažai parametrų gali būti per mažai lankstus, tuo tarpu modelis su pernelyg dideliu parametrų kiekiu gali labai prisitaikyti prie duomenų imties ir joje esančių triukšmų. Todėl siekiama parinkti reikiamą neuroninio tinklo sudėtingumą“¹⁵

Todėl mokymo epochų skaičius (kaip ir neuronų/sluoksnių skaičius) parenkamas bandymo būdu. Optimaliausią epochų skaičių nulems testavimo paklaida.

Mokymas

Visi neuroniniai tinklai buvo apmokyti naudojant įėjimo duomenis iš kintamojo „treniravimui“ ir išėjimo tikslus „isejimui“. Siekiant patikrinti ar tarp duomenų vyrauja tiesinė ar netiesinė priklausomybė, visiems tinklams priskirta Tansig (netiesinę) perdavimo funkcija, išskyrus tinklą „Network8“. Jam pritaikyta Purelin (tiesinė) perdavimo funkcija.

Kiekvienam tinklui perdavus mokymo duomenis, buvo fiksuojama paklaida mokymo metu. Lentelėje 5 ji yra stulpelyje „paklaida treniruojant“. Kaip galima matyti, mokymo metu didžiausią paklaidą darė tinklas „Network2“ (1.544634), o mažiausią tinklas „Network1“ (0.027424)

¹⁵ Habil. dr. Rimvydas Simutis, Neuroninių tinklų paskaitų medžiaga, 2006.

Apmokymo ir testavimo rezultatai

Tinklo pavadinimas	Epochų skaičius	Neuronų skaičius	Paklaida treniruojant	Std. treniravime	Std. testavime
Network1	100	20 (tansig)	0.027424	0.1658	0.1795
Network2	100	25 (tansig)	1.544634	0.1688	0.2221
Network3	100	31 (tansig)	0.108021	0.8362	0.6225
Network4	100	25 (tangsig)	0.0679332	0.2571	0.4227
Network5	100	18 (tansig)	0.027425	0.1657	0.1967
Network6	150	20(tansig)	0.027611	0.1658	0.1760
Network7	150	26(tansig)	0.027429	0.1657	0.1880
Network8	100	21(purelin)	0.0274288	0.1657	0.1750
Network9	100	20(logsig)	0.027754	0.1658	0.1773
Network10	200	40(tansig)	0.0277471	0.1660	0.1799
Network11	250	20(tansig)	0.0274289	0.1657	0.1873

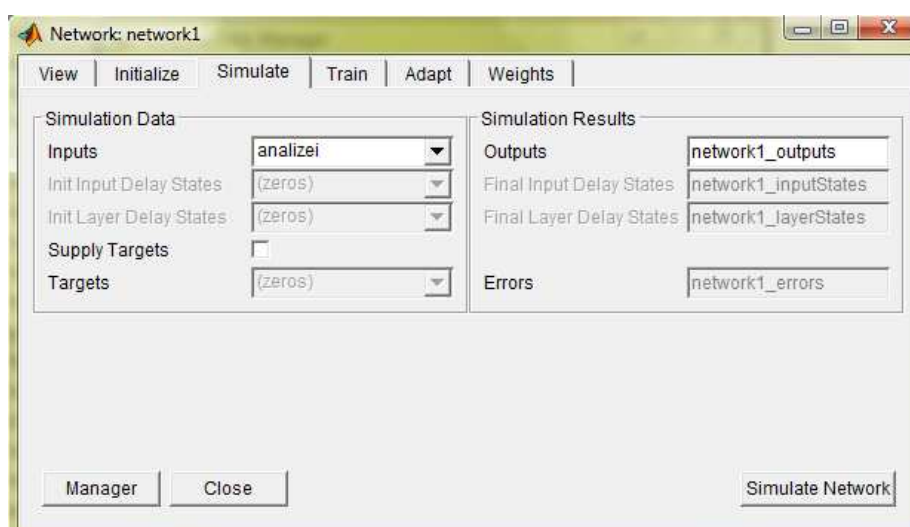
Šaltinis: sukurta autoriaus

Paskaičiavus standartinį nuokrypį tarp duomenų „įsejimui“ ir neuroninio tinklo išėigos (network_output) paaiškėjo, kad mažiausias nuokrypis yra tinklo „Network 5“. Jis sudarė 0.1657, o didžiausias- tinklo „Network 3“, kuris siekė 0.8362.

Tačiau nei paklaida, gauta mokymo metu, nei standartinis nuokrypis tarp išėigos duomenų ir tinklo sugeneruotų išėigos duomenų negali vienareikšmiškai nusakyti koks tinklas labiausiai tinka klientų lojalumo vertinimui. Tam neuronai tinklai turi būti patikrinti kontrolinių duomenų pagalba.

Kontrolinės aibės analizė

Norint išbandyti tinklus lojalių klientų atpažinime, kiekvienam iš jų buvo pateikta įeiga: antro (t.y. kontrolinio) mėnesio 15 os dienų pardavimų duomenys. (pav. 12.)

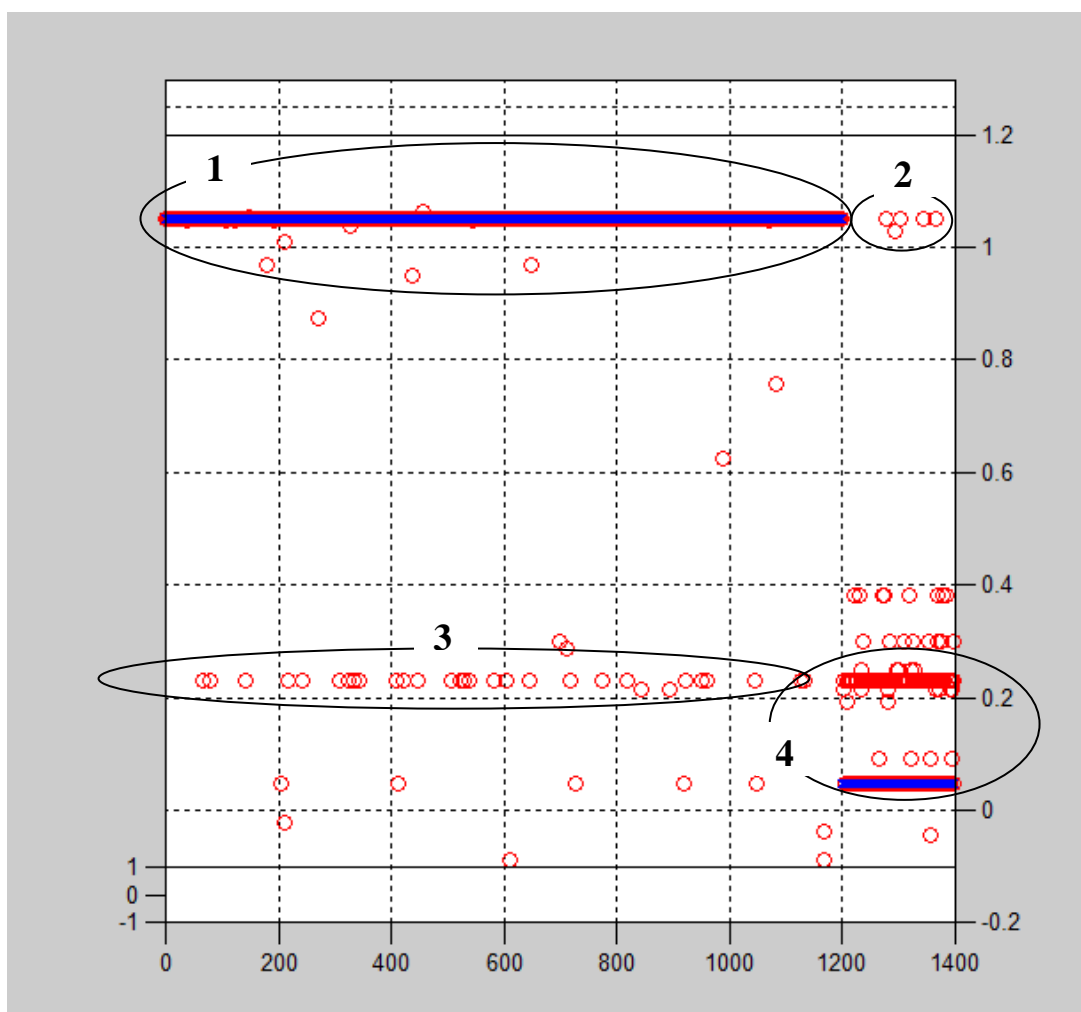


13 Pav. „Kontrolinės aibės pateikimas tinklui“

Šiuo atveju tinklai simuliuo išeią pavadinimu „Network_Output“, kuri yra lojalumo/nelojalumo požymių aibė pagal padarytas „išvadas“ mokymo metu. Ši aibė buvo lyginama su kontroline aibe: paskaičiuotas standartinis nuokrypis (kontrolinė aibė- tinklo išeią). Gauti rezultatai yra pateikti lentelėje 11, stulpelyje „Std. testavime“

Mažiausią nuokrypį padarė tinklas „Network8“ jis siekė 0.175, o didžiausią- tinklas „Network3“. Jo standartinis nuokrypis siekė 0.6225

Siekiant vaizdžiai pamatyti kontrolinę aibę ir tinklo sugeneruotą išeią aibę, nubraižome grafiką. Tinklo su mažiausia paklaida grafikas yra pateiktas pav. 10, tinklo su didžiausia- pav. 13



Šaltinis: sukurta autoriaus

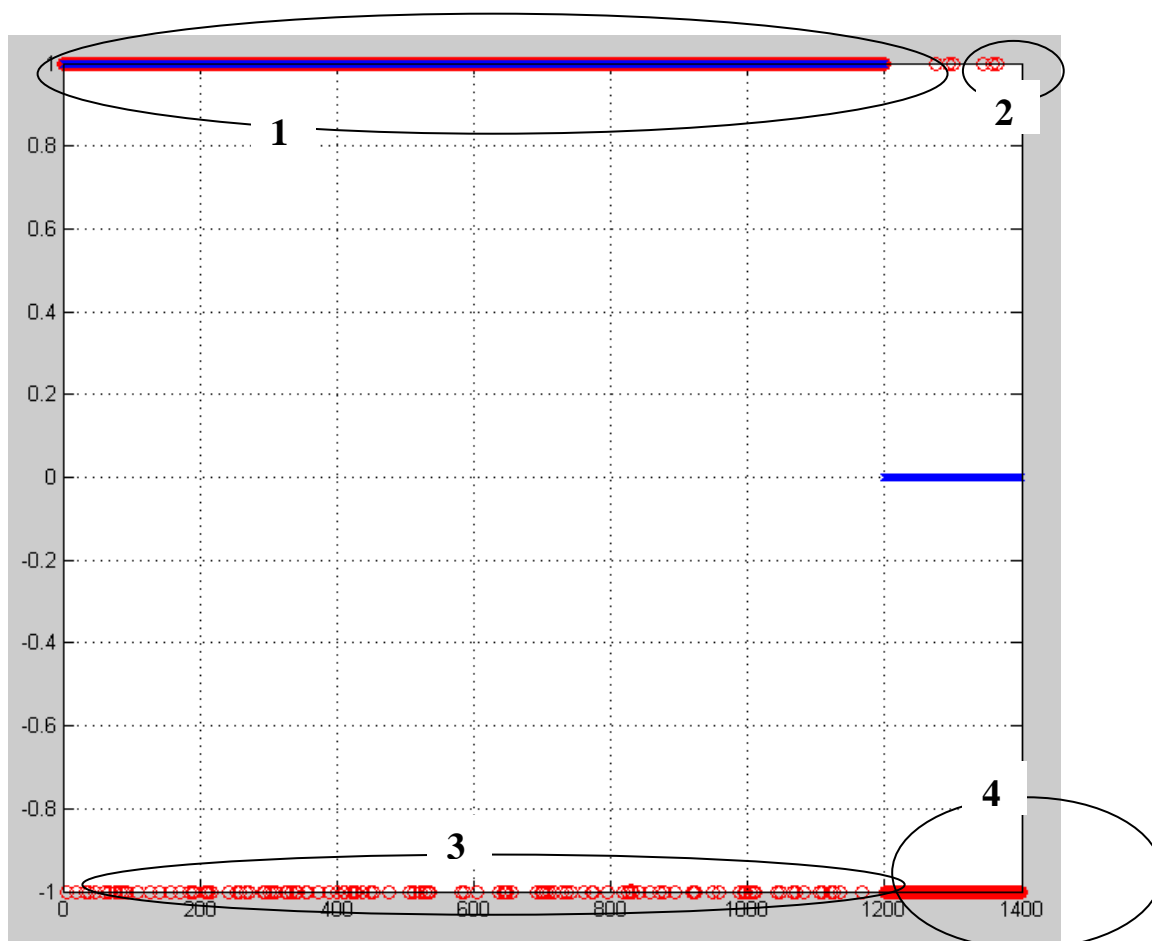
14 Pav Grafikas: „Network 8“ išeią duomenys ir kontrolinė aibė.

Raudoni apskritimai- neuroninio tinklo išeią taškai. Mėlyni- kontrolinės aibės taškai. 1- tiksliai atpažinti lojalūs klientai, 2- klaidos, atpažįstant nelojalius klientus, 3- klaidos atpažįstant lojalius klientus, 4- tiksliai ir su didele tikimybe atpažinti lojalūs klientai.

Šiame grafike (pav. 13) x ašyje yra pateiktas kliento x atvejis, o y- lojalumo požymis. 0- klientas nelojalus, 1- klientas lojalus. Mėlyna tiesė, einanti ties $Y=1$ o x (1..1200) parodo kaip buvo sugrupuoti klientai: pirmiausiai buvo renkama lojalių klientų aibė, o paskui – nelojalių.

Neuroninis tinklas „Network8“ tiksliausiai (standartinis nuokrypis mažiausias- 0.175) sugebėjo atrinkti lojalius klientus nuo nelojalių. Ties lojalių klientų kontroline aibe, atrinktos neuroninio tinklo aibės taškai suformuoja kontūrą, kuris glaudžiai atkartoja kontrolinės aibės tiesės kontūrus.

Prie nelojalių klientų kontrolinės aibės tiesės ($y=0$) neuroninis tinklo generuotos išeigos taškai suformuoja tiesę, kuri per 0.2 yra aukščiau. Tai reiškia kad tinklas su aukšta tikimybe nustatė kad šie klientai nelojalus. Klientų, kuriuos jis vienareikšmiškai apibūdino kaip nelojalius, aibės taškai suformuoja kontūrą aplink nelojalių klientų kontrolinės aibės tiesės.



Šaltinis: sukurta autoriaus

15 Pav. Grafikas: „Network3“ išeigos duomenys ir kontrolinė aibė.

Raudoni apskritimai- neuroninio tinklo išeigos taškai. Melyni- kontrolinės aibės taškai. 1- tiksliai atpažinti lojalūs klientai; 2 ir 4 –klaidos atpažįstant nelojalius klientus, 3- klaidos atpažįstant lojalius klientus.

Kaip matome iš pav. 14, didžiausia tinklo sugeneruotų išeigos tašku nesutampa su kontroline aibe. Taip jis neatpažino nei vieno „nelojalaus“ kliento (4 ir 2 grafike). O bandant

atpažinti lojalūs klientus, jis sugeneravo taškų aibę 3, kuri nesutapo su lojalių klientų aibe daugiau nei per $Y=2$ plotą. Tai parodo kad tinklas visai nesugeba atpažinti nelojalius, ir daro santykinai didelę klaidą atpažįstant lojalūs.

Tačiau net ir su tiksliausiu šiame eksperimente gautame neuroniniu tinklu, praktinis tokio modelio pritaikomumas yra komplikuotas. Pagal formulę 1, CLV vertė yra suminis rodiklis. T.y. jo skaičiavimas kiekvienai pasirinktai laikotarpio atkarpai sudaro priklausomybę tarp pačių atkarpų. Kitaip tariant pirmos dienos paskaičiuota CLV vertė tiesiogiai įtakoja sekančios dienos CLV vertę, kadangi pirmoji diena įeina į antrosios paskaičiavimą. Labai tikėtina, jog toks sudaryto neuroninio tinklo tikslumas yra sąlygotas būtent šios priklausomybės. Todėl kintamųjų/ kliento lojalumą apibūdinantys kiekybiniai rodikliai turi būti nepriklausomi.

Kitas neobjektyvų neuroninio tinklo vertinimą galėjo sąlygoti neproporcingas mokymo duomenų skaidymas. Yra pastebėta, jog neuroninis tinklas iš esmės yra linkęs daryti paklaidą tos klasės naudai, kurios egzempliorių skaičius yra didesnis. (pvz. jei turime 500 lojalių ir 1000 nelojalių, tinklas darys paklaidą atpažindamas daugiau nelojalių nei lojalių)

Atsižvelgiant į šiuos atlikto eksperimento trūkumus, atliksime sekantį eksperimentą.

5. Lojalių klientų atpažinimas remiantis kliento potraukį pirkti apibūdinančiais rodikliais

Kliento potraukio pirkti apibūdinančių rodiklių analizei, skirtingai nei CLV analizei, naudosime statistica neural networks programinį paketą. Esantis jame Intelligent problem solver vedlys palengvina mums darbą, sutrumpinant eksperimento eigą. Jis sugeba atrinkti svarbius/nesvarbius įeigos kintamuosius, atskirti duotus duomenis į testinę ir kontrolinę aibes, automatiškai sudaro kelis neuroninių tinklų variantus, visus apmoko ir parenka turintį mažiausią paklaidą.

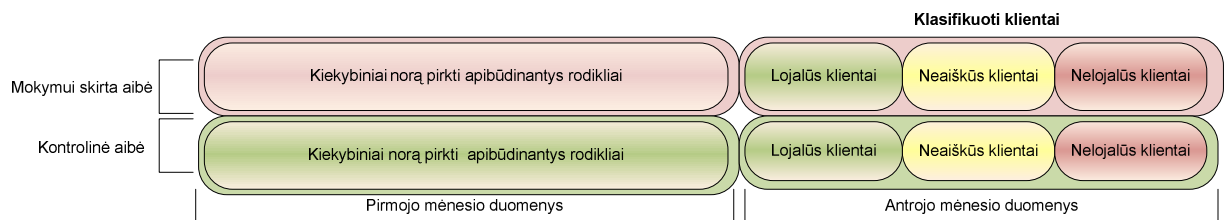
Surinkti duomenys tinka ir šiam eksperimentui atlikti, todėl atlikimo etapai sekantys:

- Duomenų paruošimas eksperimentui
- Automatinis neuroninių tinklų formavimas, mokymas ir testavimas.

Daroma prielaida, jog bendradarbiavimo su klientu trukmė- 1as mėnuo. Užlaikymo periodas- 2as mėnuo.

5.1 Duomenų paruošimas eksperimentui, apdorojimas.

Turimus pirmojo ir antrojo mėnesio pardavimų duomenis pertvarkome (pav. 15):



16 Pav. Duomenų paruošimas eksperimentui

Remiantis antrojo mėnesio duomenimis, klasifikuojame klientus į lojalius, nelojalius ir neaiškius:

Lojalių klientų klasifikavimui atlikti pasinaudojome Pareto 20/80 metodu. Surūšiuojame antrojo mėnesio klientai pagal perkamumą mažėjimo tvarka. Iš jų atrinkti klientai, kurie neša įmonei 80% antrojo mėnesio pelno. Tai sudarė 11% visų tą mėnesį pirkusių klientų (589 klientų).

Neaiškių klientų aibeį suformuoti buvo paimti sekantys 10% pagal perkamumą klientai. Daroma prielaida, jog tai yra neapsisprendę klientai, kurie su tikimybe ~50% gali pavirsti lojaliais arba nelojaliais. Šią aibę sudarė 500 klientų.

Nelojalių klientų atrinkimas: remiantis prieš tai atlikto eksperimento rezultatais, nelojalių klientų atrinkimui buvo paimti 500 klientai (kad išlaikytume klasių egzempliorių pusiausvyrą)

Kiekvienos klasės egzemplioriams skaičiuojami potraukį įmonei apibūdinantys rodikliai¹⁶ už pirmą mėnesį:

- Kliento pirkimai per pastarąsias 5 kalendorines mėnesio dienas
- Kliento pirkimai per pastarąsias 10 kalendorinių mėnesio dienų
- Suminiai kliento pirkimai per nagrinėjamą mėnesį
- A kategorijos prekių pirkimai per nagrinėjamą mėnesį
- B kategorijos prekių pirkimai per nagrinėjamą mėnesį
- C kategorijos prekių pirkimai per nagrinėjamą mėnesį
- D kategorijos prekių pirkimai per nagrinėjamą mėnesį
- Suminis nupirktų prekių per nagrinėjamą mėnesį skaičius
- Nupirktų A kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų B kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų C kategorijos prekių skaičius per nagrinėjamą mėnesį
- Nupirktų D kategorijos prekių skaičius per nagrinėjamą mėnesį
- Paskutinio apsilankymo naujumas (dienų skirtumas po paskutinio apsilankymo)
- Apsilankymo naujumo standartinis nuokrypis (standartinis nuokrypis dienų skirtumo tarp apsilankymų)
- Vidutinis apsilankymo naujumas (vidutinis dienų skirtumas tarp apsilankymų)

Čia:

A kategorijos prekė tai prekė, kurios kaina yra tarp 2000Lt ir 12 000 Lt

¹⁶ Wouter BUCKINX, Geert VERSTRAETEN, Dirk Van den POEL, (2007) „Predicting customer loyalty using internal transactional database“

B kategorijos prekė tai prekė, kurios kaina yra tarp 1000 Lt ir 1999 Lt

C kategorijos prekė tai prekė, kurios kaina yra tarp 100 Lt ir 999 Lt

D kategorijos prekė tai prekė, kurios kaina yra tarp 0Lt ir 99 Lt

Atlikus tokį duomenų pertvarkymą gavome duomenų masyvą, kuriame suskaičiuoti visi aukščiau išvardinti klientų rodikliai ir lojalumo požymis, nusakantis kokiai klasei yra priskirtas klientas.

Papildomų duomenų pertvarkymų (skaidymas į testinę ir kontrolinę aibes) nebūtinai, kadangi statistica neural network intelligent problem solver tai atlieka savarankiškai

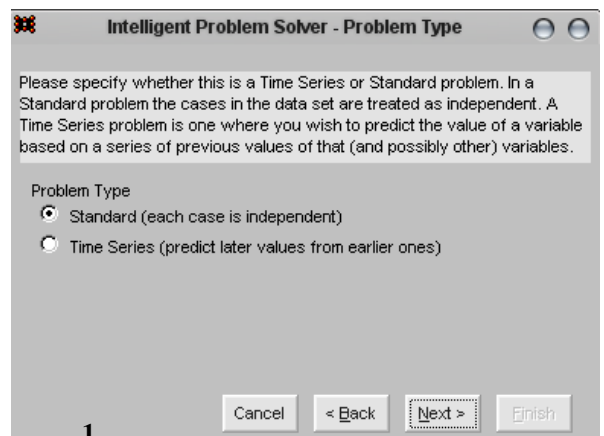
5.2 Neuroninių tinklų formavimas, mokymas ir rezultatų analizė

Sudarytas duomenų masyvas įkeliamas į statistica aplinką. Kintamasis lojalumas automatiškai perkoduojamas, nurodant nominalias reikšmes „Lojalus“, „Neaiškus“ ir „Nelojalus“.

Paleidžiamas intelligent problem solver: (17 pav.) Nurodomas Standart tipas (kiekvienas klientas) yra vertinamas kaip atskiras egzempliorius.

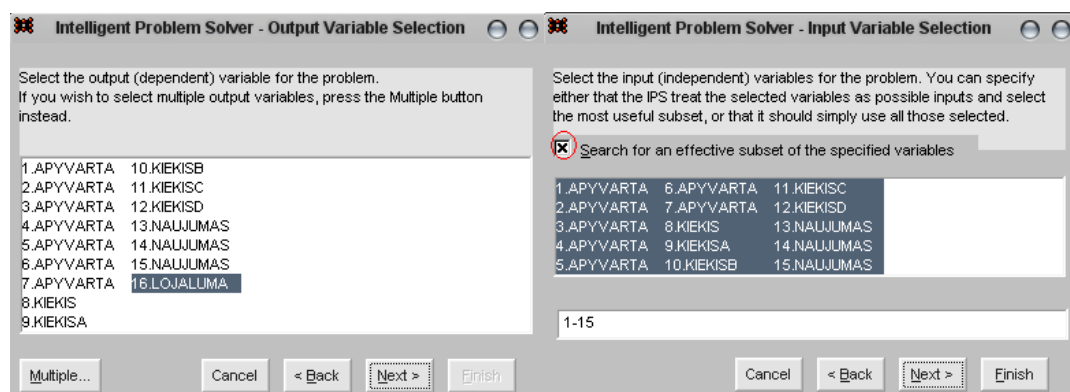
„Time series“ skirtas laiko įtakojamų egzemplioreių prognozavimui (pvz. akcijų vertės pokytis)

Toliau pasirenkame įeigos ir išeigos kintamuosius



7 Pav. IPS vedlio paleidimas

Mūsų atveju potraukį apibūdinantys rodikliai yra įeigos, o lojalumas- išeigos kintamieji. (18 pav.)

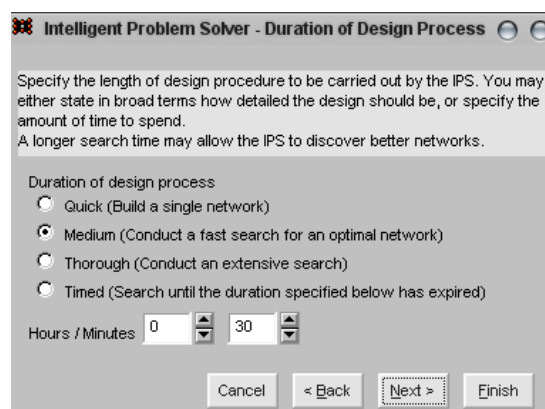


18 Pav. Įeigos ir išeigos kintamųjų pasirinkimas

Pasirinkę įreigos kintamuosius, pažymime kad Intelligent Problem solver atliktų kintamųjų svarbos analizę (Search for an effective subest of the specified variables) kurios metu bus parinkti tik išreigos kintamąjį įtakojantys įreigos kintamieji.

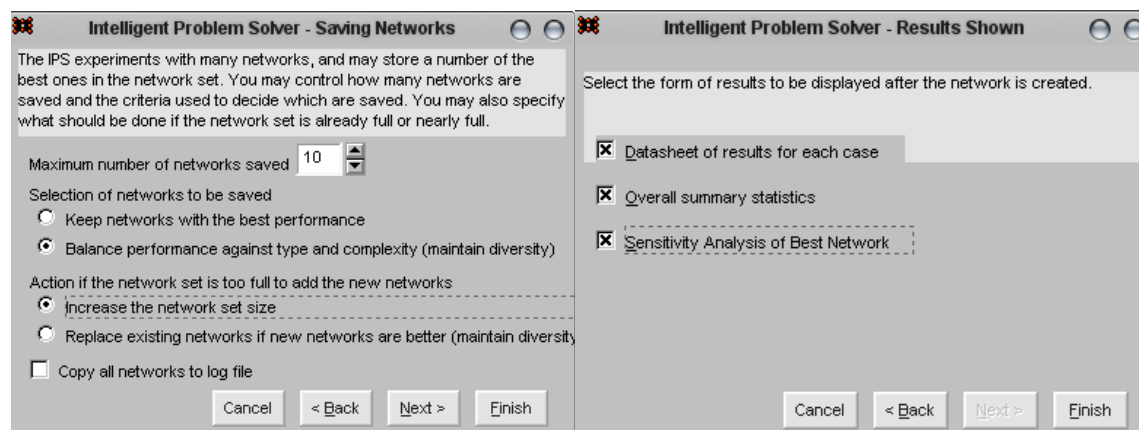
Sekantys pasirinkimai apibrėžia tyrimo trukmes. Pabrėžtina kad trukmė tiesiogiai įtakoja kuriamų tinklų skaičių ir apmokymo/testavimo kruopštumą.

Iš duotų pasirinkimų, paliekame pagal nutylėjimą siūloma Medium laiko sąnaudas tinklų sukūrimui ir apmokymams. (19 pav.)



19 Pav. Tyrimo trukmės pasirinkimas

Toliau nurodome kiek ir kokių tinklų norime saugoti, ir nusimatome veiksmus, jeigu neuronų tinklas persipildo.(20 Pav.)



20 Pav. Kuriamų tinklų nustatymai

Šiuo atveju tinklo dydis yra didinamas. Sekančiame ir paskutiniame vedlio žingsnyje nurodome kokią statistiką norėtume gauti tyrimo gale. Visapusiškai analizei pažymime visus siūlomus variantus. Mokymo/testavimo/ rezultatas kiekvienam egzemplioriui, bendra suminė statistika ir jautrumo analizė geriausiam atrinktam tinklui.

Po intelligent problem solver paskaičiavimų gauname klientų tyrimo rezultatus. Jie pateikti sekančiame skyriuje.

5.3 Gautų klientų lojalumo tyrimo rezultatų analizė

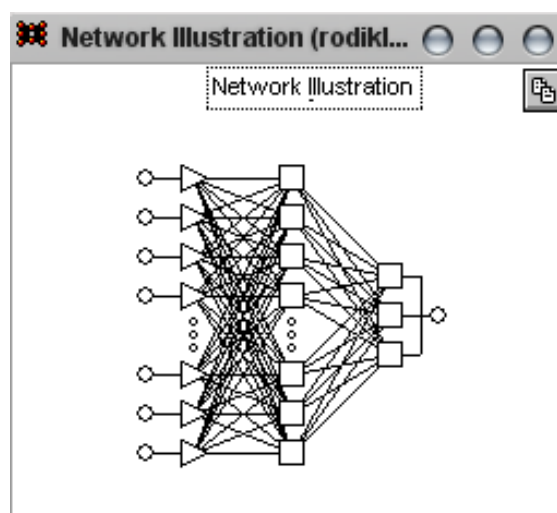
Vedlio nustatymų metu pasirinkome palikti 10 neuroninių tinklų ir parodyti suminę jų statistiką

Type	Error	Inputs	Hidden	Performance	
01	Linear	25,98775	9	-	0,4128878
02	RBF	0,4504832	3	1	0,4892601
03	RBF	0,4482466	3	1	0,4892601
04	Linear	0,4477435	13	-	0,5250597
05	RBF	0,4325021	3	2	0,5155131
06	Linear	0,4207579	15	-	0,5799523
07	MLP	0,3782272	6	5	0,7016706
08	MLP	0,3677575	7	9	0,7326969
09	MLP	0,3553016	15	30	0,7470167
10 *	MLP	0,3535663	13	23	0,7541766

21 Pav. Sukurtų neuroninių tinklų struktūros abibendrinimas

Šiame (pav. 19) lange yra pateikti 10 neuroninių tinklų. Turintis mažiausią paklaidą ir didžiausią našumą yra pažymėtas * žyme. Tai Multi Linear Perceptron tipo tinklas. Turėdamas 13a įėjimo neuronų (Apsilankymo naujumo standartinis nuokrypis ir Vidutinis apsilankymo naujumas interpretuoti kaip nereikšmingi) ir 23 paslėptų, jis gali prognozuoti su 0,3535663 paklaida ir turi didžiausią našumą 0,7541766 (našumas- kaip greitai apdorojami duomenys (connection per second))

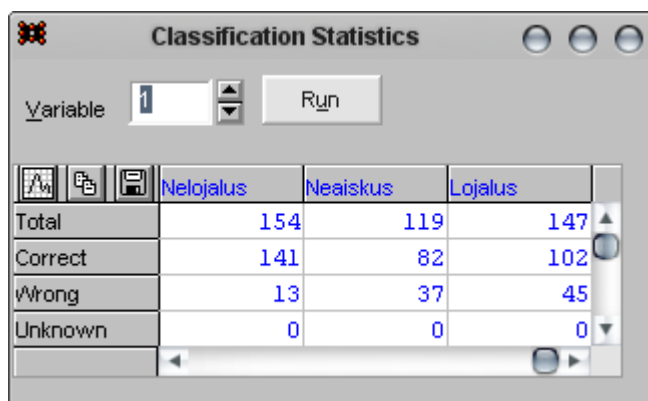
Radialinės funkcijos pagrindu(RBF) sudaryti tinklai taip pat kaip ir tiesinės (Linear) turi didžiausią paklaidą, kas patvirtina, jog tarp klientų lojalumo ir pasirinktų kintamųjų vyrauja ne tiesinė priklausomybė. Principinę tiksliausio sudaryto neuroninio tinklo struktūrą galime matyti iš žemiau pateikto paveikslo (Pav.20).



22 Pav. Principinė neuroninio tinklo struktūra

Kaip matome jis turi 3 paskutinio paslėptojo sluoksnio neuronus (kiekvienai klasei po neuroną)

Klasifikacijos statistika šiam neuroniniam tinklu pateikia tokią suvestinę:



	Neįėjus	Neaiškus	Lojalus
Total	154	119	147
Correct	141	82	102
Wrong	13	37	45
Unknown	0	0	0

23 Pav. Klasifikavimo statistika

Atliekant kontrolinį duomenų apie klientus tyrimą, neuroninis tinklas iš 154 lojalių klientų atpažino 141, suklydo atitinkamai dėl 13 klientų. Atpažįstant neaiškius, iš 119 pateiktų jis nustatė 82 o suklydo dėl 37 klientų. Atrenkant lojalius, iš 147 klientų jis atrinko 102 teisingai, o dėl 45 suklydo. Nežinomų variantų nepasitaikė.

Šie tyrimo rezultatai iš dalies paneigia vadybinę taisyklę 55-15 kuri teigia, jog įmonės bandant klasifikuoti geriausius (20 % pelningiausių) klientus, maždaug 55 % yra klasifikuojami klaidingai. Šio tyrimo metu tinklas suklydo tik dėl 13 klientų, kas sudaro 8%.

Lyginant su CLV analizės metu gautais rezultatais, šie rezultatai yra artimesni realybei, kadangi tarp pačių įeigos kintamųjų nėra priklausomybės. Be to intelligent solver pagalba iš visų pateiktų potraukį apibūdinančių rodiklių nereikšmingi atmesti (Apisilankymo naujumo standartinis nuokrypis ir Vidutinis apsilankymo naujumas interpretuoti kaip nereikšmingi) ir palikti tik didžiausią įtaką išeigos kintamajam turintys rodikliai.

Pastarojo neuroninio tinklo tikslumas gali būti padidintas, naudojant iteracinį grįžtamąjį ryšį ir pastoviai jį bemokant.

IŠVADOS

1) Išnagrinėjus pagrindinius lojalumo apibrėžimus, lojalių klientų atskyrimo bei lojalumo įvertinimo metodus paaiškėjo, kad jie tik iš dalies apibūdina lojalius klientus, ir neleidžia įmonei vienareikšmiškai įvertinti lojalumą.

Apžvelgus pagrindinius modelius, o taip pat (pareto 20/80, RFM ir Banginio kreivės) lojaliems klientams atskirti bei lojalumui įvertinti paaiškėjo, kad plačiausiai paplitęs metodas yra Pareto 20/80, kuris teigia jog 20 procentų geriausių (pelningiausių) įmonės klientų duoda 80 procentų įmonės pelno. Remiantis šiais rodikliais, įmonė gali atskirti savo gyvybiškai svarbius klientus ir klientus, kurie įmonei yra nuostolingi

Pasak Wouter Buckinx, Geert Verstraeten, Dirk Van den Poel, kliento „potraukį pirkti“ galima aprašyti tokiais rodikliais kaip: Kliento pirkimai per pastarąsias 5, 10 kalendorinių mėnesio dienų, suminiai kliento pirkimai; A,B,C,D kategorijos prekių pirkimai bei suminis nupirktų prekių per nagrinėjamą mėnesį skaičius

Nagrinėjant E. Ksevelonakiso pasiūlytą CLV etapų modelį pastebėtas pagrindinis trūkumas: klausimas „kurie klientai yra linkę atsitraukimui“ gimsta tik užlaikymo stadijos pabaigoje. Nustačius kurie klientai yra linkę atsitraukti įmonė pereina prie susigrąžinimo stadijos. Tuo tarpu Reinartz and Kumar (2002) teigia “let butterflies fly.” (leiskime drugeliams išskristi) ir koncentruotis ties potencialiai pelningais klientais.

Paaiškėjo, kad lojalumo tyrimas susiveda į sekančius uždavinius: klasifikavimas remiantis kliento norą pirkti apibūdinančiais rodikliais ir klasifikavimas remiantis CLV vertės pokyčio analize.

2) Apžvelgus pagrindinius statistinius metodus (sprendimo medžiai, daugiaregresinė analizė ir kt.), ir išnagrinėjus jų tinkamumą šiems uždaviniams spręsti paaiškėjo kad minėtiems uždaviniams realizuoti labiausiai tinka neuroninis tinklas kadangi:

- Nėra žinių apie priklausomybę tarp kliento elgesio bendradarbiavimo stadijoje ir lojalumo, nelojalumo statuso užlaikymo stadijoje. (Neuroninis tinklas nereikalauja a priori žinių apie duomenis.)
- Mokosi iš savo klaidų: Susigrąžinimo etapo rezultatas, kuomet mes žinome post-factum kuris klientas ištiktųjų lojalus o kuris ne- gali būti pateiktas tinklui kaip papildoma mokymosi imtis. Taip kiekvieną kartą tinklo klasifikavimo tikslumas didės. (Klaidas toleruojantis mechanizmas.)
- Užpildo duomenų „spragas“: Jei apie klientą surinkta nevisapusiška informacija: nepavyko identifikuoti atliekant pirkimo operacijas ir pan. (Duomenų trūkumas sintetiškai pašalinamas tinklu imituojant reikšmes.)

- Iš esmės atitinka darbo uždavinį. Tai klasifikavimas ateičiai, kitaip tariant prognozavimas.

Vienas algoritmas gali būti panaudotas kaip prognozavimui taip ir klasifikavimui

„Neuroniniai tinklai negali atlikti nieko tokio, ko negali kiti metodai“¹⁷, tačiau jie gali sutaupyti laiką ir pinigus ieškant priklausomybių tarp Klientų lojalumo ir kiekybinių kliento vertinimo rodiklių. Klientų lojalumo tyrimui pasirinktas neuroninis tinklas.

3) Išnagrinėjus neuroninių tinklų struktūrą bei taikymo galimybę duomenų apie klientus analizei bei ateities elgesio prognozei paaiškėjo, kad“:

Formuojant duomenų analizės modelį neuroninių tinklų pagalba, naudojama 7ių žingsnių metodologija: 1. Kintamųjų parinkimas; 2. Duomenų apdorojimas; 3. Mokymasis ir testavimas; 4. Neuroninių tinklų paradigmos; 5. Perdavimo funkcijos; 6. Įvertinimo (pripažinimo, angl. evaluation) kriterijai; 7. Neuroninių tinklų apmokymas

Remiantis šiais metodologiniais etapais, realizuojamas sprendimas pagal pasiūlyta metodą.

4) Remiantis teorinėje dalyje išskirtais esamų klientų gyvenimo ciklo vertės modelio trūkumais, pasiūlyti 2 sprendimo metodai:

Lojalių klientų atrinkimas, jų istorijos tyrimas naudojant CLV vertės prognozavimą:

Modelio esmė – lojalių klientų atpažinimas remiantis CLV rodiklio pokyčio istorija. CLV tai rodiklis, kuris skaičiuoja kliento (brangių/ pigių prekių) pirkimus, tų pirkimų pokyčius duotai bendravimo su klientu dienai. Skaičiavimus galima atlikti pagal Bolton, Lemon pasiūlytą formulę (formulė 1).

Klasifikavimas remiantis kliento „norą pirkti apibūdinančiais rodikliais“: Kliento lojalumas prognozuojamas pagal Wouter BUCKINX, Geert VERSTRAETEN, Dirk Van den POEL pasiūlytų kliento „norą apibūdinančių rodiklius.

Kadangi CLV analizės pagrįstas klientų lojalumo vertinimas ir vertinimas remiantis kliento patrauklumo įmonei analizės nėra palygintos mokslinėje literatūroje, eksperimentinėje dalyje atlikti 2 eksperimentai kurių metu įvertintas jų tinkamumas klientų lojalumo vertinimo uždaviniui spręsti

5) **CLV pokyčio analizės eksperimento** metu gauta santykinai maža prognozavimo paklaida, tačiau išskirtas pagrindinis trūkumas ir neuroninio tinklo vertinimo neobjektyvumas dėl skaičiavimo kiekvienai pasirinktai laikotarpio atkarpai(tai sudaro priklausomybę tarp pačių atkarpų) Taip pat neobjektyvumas sąlygotas neproporcingu mokymo duomenų skaidymu. Eksperimentui atlikti buvo naudojamas Matlab NNtool programinis paketas

Kliento potraukio pirkti analizės eksperimento dalyje skirtingai nei CLV analizei, panaudotas statistica neural networks programinį paketą:

17 (Habil. dr. Rimvydas Simutis, Neuroninių tinklų paskaitos, 2006)

Šio eksperimento rezultatai iš dalies paneigė vadybinę taisyklę 80-15 kuri teigia, jog įmonės bandant klasifikuoti geriausius (20 % pelningiausių) klientus, maždaug 55 % yra klasifikuojami klaidingai. Šio tyrimo metu tinklas suklydo tik dėl 13 klientų, kas sudaro 8%.

Lyginant su CLV analizės metu gautais rezultatais, šie rezultatai yra artimesni realybei, kadangi tarp pačių įeigos kintamųjų nėra priklausomybės. Pastarojo neuroninio tinklo tikslumas gali būti padidintas, naudojant iteracinį grįžtamąjį ryšį ir pastoviai jį bemokant.

Skirtingai nei kitos šio metu siūlomos priemonės, šiame darbe pasiūlytas metodas išsiskiria savo lankstumu. Jis nereikalauja griežto lojalių klientų apibrėžimo, kadangi gali mokytis iš pateiktų pavyzdžių. Kadangi tai klasifikavimas remiantis istoriniu potraukį apibūdinančių rodiklių pokyčiu (prognozavimas), įmonė gali laiku (užlaikymo stadijos pradžioje) išgauti informaciją klientus, kurie yra linkę atsitraukimai, o kurie potencialiai Lojalūs. Įmonė gali nedelsiant taikyti tikslines strategijos priemones ir laiku pateikti klientui atitinkamą pasiūlymą, sumažinant jo reakcijos laiką (lag)

LITERATŪROS SĄRAŠAS

1. ANTON, J. (1996) Customer Relationship Management. Making Hard Decisions with Soft Numbers. Upper Saddle River, NJ: Prentice Hall.
2. MALTHOUSE E. C. , BLATTBERG R. C. (2005m.), "Can We Predict Customer Lifetime Value?", Kellogg School of Management, Northwestern University."
3. PROCTOR, R.A. (1992m.), "Marketing Decision Support Systems: A Role of Neural Networking", Marketing Intelligence&Planning, Vol.10, Nr.1
4. THALL, N (1992m.), "Neural Forecasts: A Retail Sales Booster", Discount Merchandiser, Vol. 32, Nr 10.
5. VAVRA, T.G. (1997) "Improving Your Measurement of Customer Satisfaction: A Guide to Creating, Conducting, Analyzing and Reporting" Customer Satisfaction Measurement Program. Milwaukee, Wis.: ASQ Quality Press.
6. VENUGOPAL V., BAETS W.,(2005) "Neural Networks and Statistical Techniques in Marketing Research", Marketing Intelligence & Planning Vol.12, Nr7.
7. WHITE, H. (1998), "Neural Networks Learning and Statistics", AI expert, Vol4, Nr12.
8. Habil. dr. Rimvydas Simutis, Neuroninių tinklų paskaitų medžiaga, 2006.

SANTRAUKA

BREDICHIN, Michail. (2008) *Evaluation of Customer Loyalty by Applying Neural Networks*. Master Graduation Paper. Kaunas: Vilnius University, Kaunas Faculty of Humanities, Department of Informatics. 65 pages.

SUMMARY

The client loyalty is one of the most vital things to any company. During last two decades the attention to customer relationship management systems has increased significantly. There are many criterias and calculation methods described in customer relationship theories.

The existing management rule 80-15 states that when trying to classify best customers, almost 55% are misclassified and therefore organization loses its investments (there's not enough attention being paid to the best customers). The main reason of such misclassification is continuous change in business environment and subjective customer loyalty definition. Loyal customers do not stay loyal all the time.

Using scientific literature analysis there were selected most widely used methods, which help the company to identify it's the most important customers. Additionally "wish to buy" indicators were selected, which describe the customer buying potential.

Using selected methods and indexes, there is a proposition of a new classifying model, which helps to predict customer class using collected data during contribution period. Model was implemented by using neural network and tested using experiment. It predicted customer class using sales data and the results were as follow: it classified 141 customer from 154 customers in total, made a mistake in 13 cases while classifying loyal customers. When classifying disloyal customers, it made a mistake in 45 cases and classified 102 correctly of 147 in total.

Differently from the other tools that are offered at present, the proposed method distinguishes by its flexibility. It does not require strict customer loyalty definition, because it can teach itself by using subjective examples given by any company so they can get information on a customer class in time and implement marketing strategy.

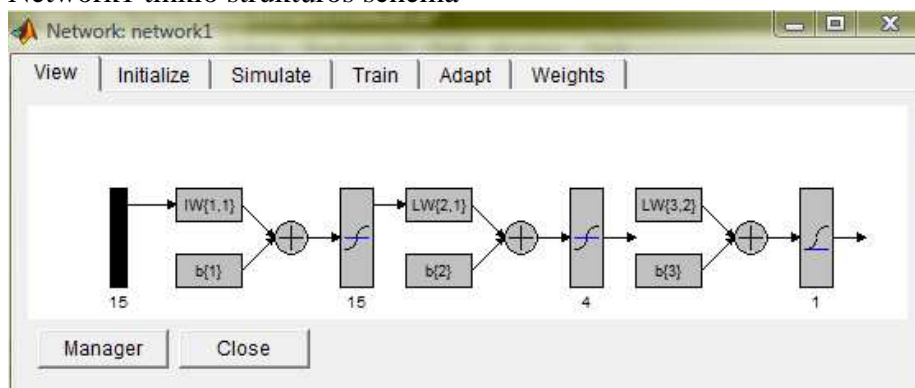
Projects consist of 66 pages, including 3 tables, 21 figures and 13 pages of appendixes.

PRIEDAI

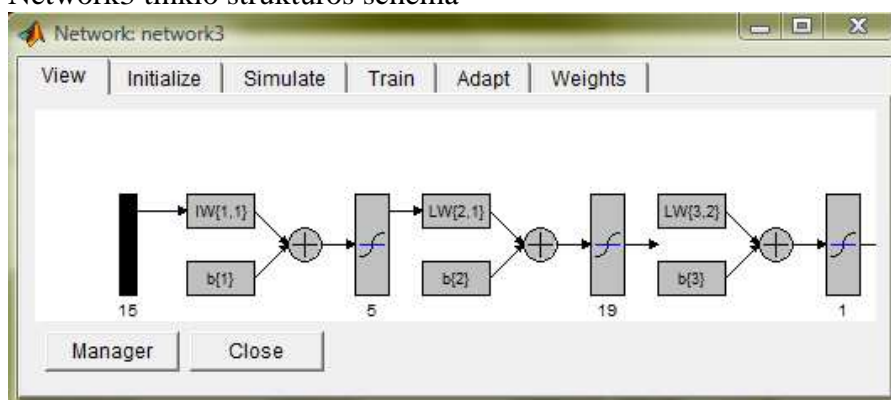
PRIEDAS 1 „Neuroninių tinklų struktūros schemas“	55
PRIEDAS 2 „MATLAB programinis tekstas duomenų apdorojimui“	58
PRIEDAS 3 „TSQL programinis tekstas eksperimento duomenų formavimui“	59

PRIEDAS 1 „Neuroninių tinklų struktūros schemas“

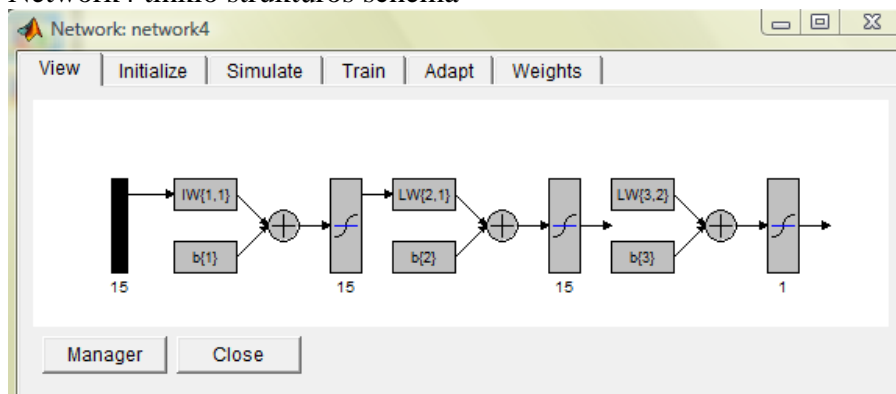
Network1 tinklo struktūros schema



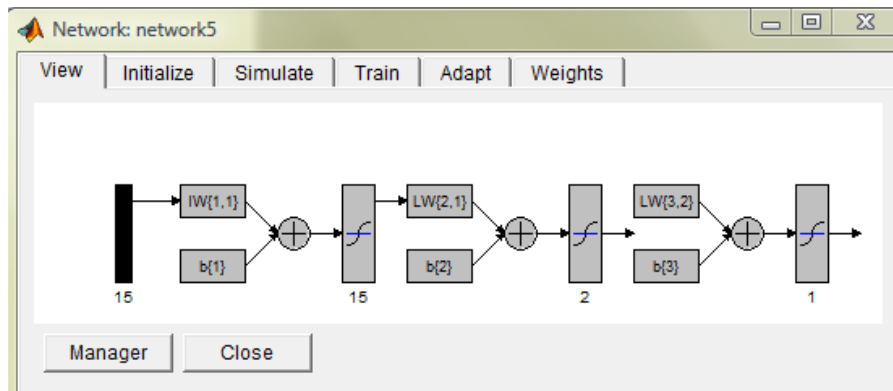
Network3 tinklo struktūros schema



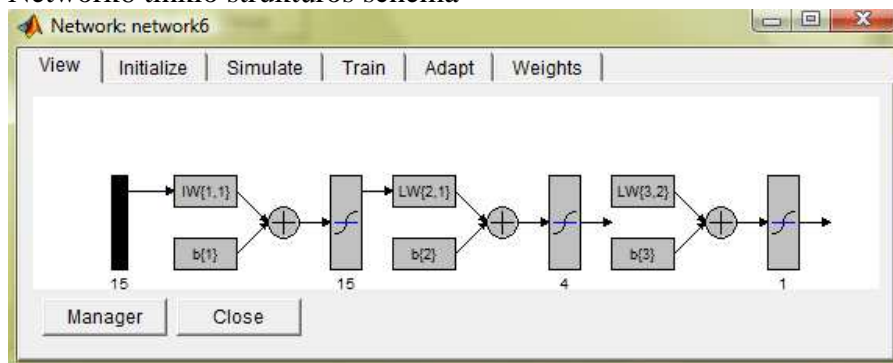
Network4 tinklo struktūros schema



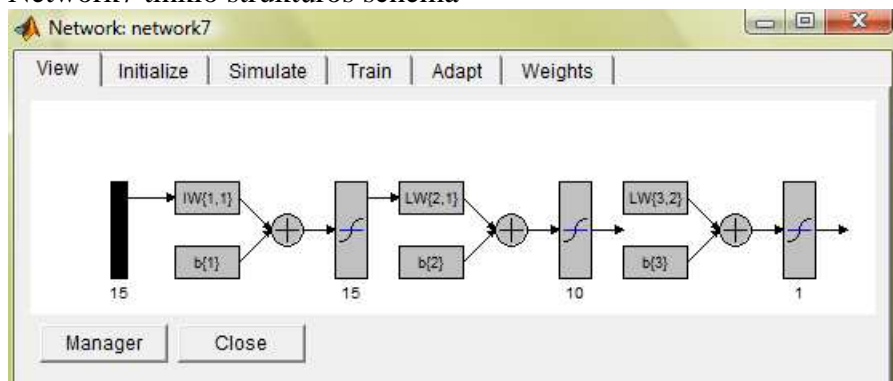
Network5 tinklo struktūras schema



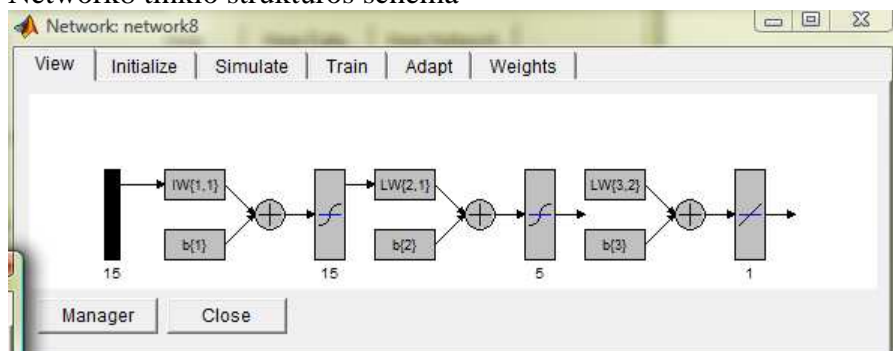
Network6 tinklo struktūras schema



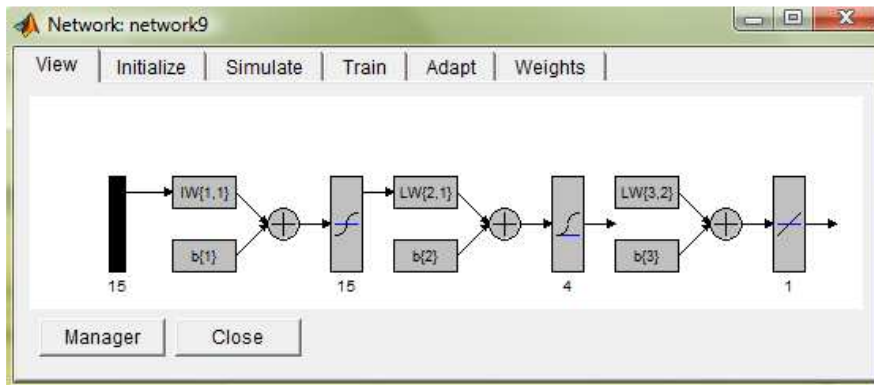
Network7 tinklo struktūras schema



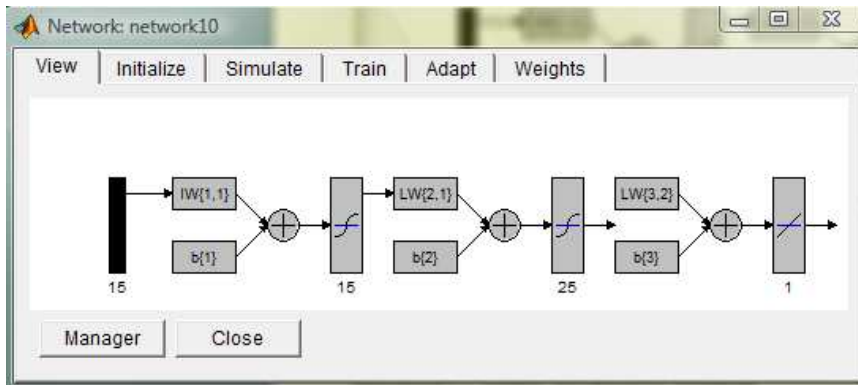
Network8 tinklo struktūras schema



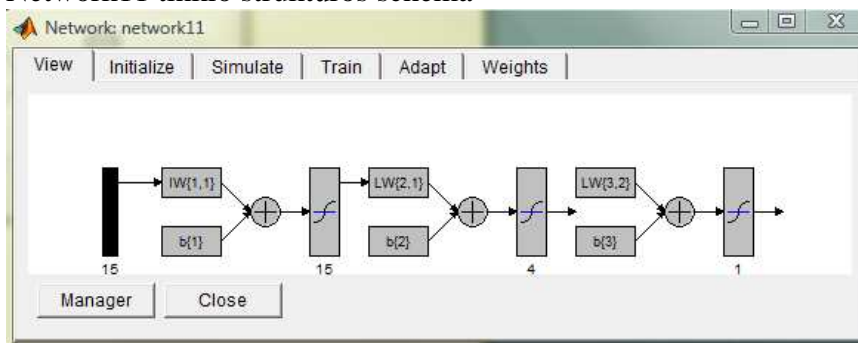
Network9 tinklo struktūras schema



Network10 tinklo struktūras schema



Network11 tinklo struktūras schema



PRIEDAS 2 „MATLAB programinis tekstas duomenų apdorojimui“

```
% [i,j]=size(transpuonuoti_duomenys) % matricos apimtis (stulpeliai ir eilutes)
% d1(1,:)=data(1:i,2); % Iejimas - praleidziamas pirma stulpelis nes tai auto identity
% d2(:,1)=data(1:i,3)
% d3(:,1)=data(1:i,4)
% d4(:,1)=data(1:i,5)
% d5(:,1)=data(1:i,6)
% d6(:,1)=data(1:i,7)
% d7(:,1)=data(1:i,8)
% d8(:,1)=data(1:i,9)
% d9(:,1)=data(1:i,10)
% d10(:,1)=data(1:i,11)
% d11(:,1)=data(1:i,12)
% d12(:,1)=data(1:i,13)
% d13(:,1)=data(1:i,14)
% d14(:,1)=data(1:i,15)
% d15(:,1)=data(1:i,16)
% treniravimui(:,15)=data(1:i,15)
% isejimui(1:i)=d15(1:i,1)
% treniravimui(1:15,1:j)=transpuonuoti_duomenys(1:15,1:j);
% isejimui(1:j)=transpuonuoti_duomenys(16,1:j);
% [k,v]=size(isejimui)
% [n,m]=size(treniravimui)
[i,j]=size(kontra);
analizei(1:15,1:j)=kontra(1:15,1:j);
kontrolei(1:j)=kontra(16,1:j);
% t=(1:1398);plot(t,network11_outputs,'ro-',t,kontrolei,'bx-'), grid paklaida vazidziai
% t=(1:1398);plot(t,network2_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network3_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network4_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network5_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network6_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network7_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network8_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network9_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network10_outputs,'ro',t,kontrolei,'bx'), grid
% t=(1:1398);plot(t,network11_outputs,'ro',t,kontrolei,'bx'), grid
% for v=0:i
% v=v+1;
% isejimas(v)=d15(1v,1);
% end
```

PRIEDAS 3 „TSQL programinis tekstas eksperimento duomenų formavimui“

```
--select count (distinct klientas) from Mishai..misai
--in to #prekes
--select top 1 * from misai

--suformuojame kiek kas pirko
-- select distinct klientas klientas, count(dok_nr) skaicius
-- into #klientai_pirkimu_sk
-- from misai
-- group by klientas

--select top 10 * from #klientai_pirkimu_sk

--select * from #klientai_pirkimu_sk
--order by skaicius desc
-----

-- declare @Tipas char (12)
-- --atrenkame kokios prekes buvo pirktos ir kokios ju kainos
-- select distinct preke preke, max(kaina) kaina
-- into #prekes_kainos_tipai
-- from misai
-- group by preke

--select top 10 * from #prekes_kainos_tipai
--order by kaina desc
-----
--
-- declare curs1 cursor for select preke, kaina from #prekes_kainos_tipai (nolock)
-- declare @prek char (12), @kain float, @tipas char (12)
-- open curs1
-- fetch next from curs1 into @prek, @kain
-- while @@fetch_status = 0
-- begin
-- if @kain >=3000 set @tipas = 'Dideles_vertes' else set @tipas='Bazine'
-- select @prek preke, @kain kaina, @tipas tipas into #prekes_tipai
-- --select * from #prekes_tipai
-- fetch next from curs1 into @prek, @kain
-- end
-- close curs1
-- deallocate curs1
-- drop table #prekes_tipai
--
-- select * from #prekes_tipai

--atrenkam perkamiausiu
select top 400 klientas, count (dok_nr) pirko_kartu, sum(kaina*kiekis) pelnas
into #top400_pelnas
```

```

from misai
where month(data)=6
group by klientas
order by pelnas desc

--atrenkam lankomiausiu
select top 800 klientas, count (dok_nr) pirko_kartu, sum(kaina*kiekis) pelnas
into #top400_apsilankymas
from misai
where month(data)=6
and klientas not in (select klientas from #top400_pelnas)
group by klientas
order by pirko_kartu desc

---drop table #top400_apsilankymas

---select top 8 * from #top400_apsilankymas

--sudarome prikimu istorija dienomis perkamiausiems (15 dienu istorija)

--select top 1 * from Misai

declare @i int
set @i=1
while (@i< 30)
begin
declare cursorius cursor for
select klientas from #top400_pelnas (nolock)
open cursorius
declare
@klientas char (12),
@CLV numeric, --clv
@DW numeric, -- Baziniu transakciju per laika i skaicius
@UP numeric, -- dideles vertes transakciju vykdytu per laika i skaicius
@j int -- laikinas skaitliukas
fetch next from cursorius into @klientas
while @@fetch_status = 0 begin

--skaiciuojame up
set @UP = (
select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina > 3000 )
and (kaina < 12000) )
--skaiciujame DW
set @dw = (
select count(dok_nr)from misai
where

```

```

misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina < 3000 )
--skaiciuojame CLV
set @clv= @up+@dw
set @j=1 --(skaiciuojame laipsni)
while (@j< @i)
begin
set @clv=@clv/(1.07)
set @j=@j+1
end
--irasome paskaiciuotus parametrus i lenta lojalumo_istorija
INSERT INTO lojalumo_istorija
(klientas, diena, up, dw ,clv, menuo )
VALUES ( @klientas, @i, @up, @dw, @clv, 6)
--skaiciuojame dw
--select sum(dok_nr) @DW from misai
--where klientas=@klientas
--and month (data)=5
--and day (data) is between 1 and @i
--and preke in (select prek from #prekes_tipai where tip = 'Bazine')
--select @klientas klientas,@diena @up+@dw CLW
--into #TOP_PELNAS_CLW
set @up=0
fetch next from cursorius into @klientas
end
close cursorius
deallocate cursorius
set @i=@i+1
end

```

--Skaiciuojame lojalumo istorija lankomiausiems klientams

```

declare @i int
set @i=1
while (@i< 30)
begin
declare cursorius cursor for
select klientas from #top400_apsilankymas (nolock)
open cursorius
declare
@klientas char (12),
@CLV numeric, --clv
@DW numeric, -- Baziniu transakciju per laika i skaicius
@UP numeric, -- dideles vertes transakciju vykdytu per laika i skaicius
@j int -- laikinas skaitliukas
fetch next from cursorius into @klientas
while @@fetch_status = 0 begin
--skaiciuojame up
set @UP = (

```

```

select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina > 3000 )
and (kaina < 12000) )
--skaiciujame DW
set @dw = (
select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina < 3000 ))
--skaiciuojame CLV
set @clv= @up+@dw
set @j=1 --(skaiciuojame laipsni)
while (@j< @i)
begin
set @clv=@clv/(1.07)
set @j=@j+1
end
--irasome paskaiciuotus parametrus i lenta lojalumo_istoirija
INSERT INTO lojalumo_istoirija
(klientas, diena, up, dw ,clv, menuo )
VALUES ( @klientas, @i, @up, @dw, @clv, 6)
--skaiciuojame dw
--select sum(dok_nr) @DW from misai
--where klientas=@klientas
--and month (data)=5
--and day (data) is between 1 and @i
--and preke in (select prek from #prekes_tipai where tip = 'Bazine')
--select @klientas klientas,@diena @up+@dw CLW
--into #TOP_PELNAS_CLW
set @up=0
fetch next from cursorius into @klientas
end
close cursorius
deallocate cursorius
set @i=@i+1
end

--delete
--from lojalumo_istoirija
--Parenkame nelojaliu klientu aibe:

--Atsitiktiniu atranka
select top 400 klientas, count (dok_nr) pirko_kartu, sum(kaina*kiekis) pelnas

```

```

into #top400_pelnas_atstiktiniai
from misai
where month(data)=6
group by klientas
having sum(kaina*kiekis)between 0 and 5 and count (dok_nr)=1
order by pirko_kartu desc, pelnas asc

--select * from #top400_pelnas_atstiktiniai

-- Daug lankiusiubet nieko nepirkusiu
select top 400 klientas, count (dok_nr) pirko_kartu, sum(kaina*kiekis) pelnas
into #top400_pelnas_nepirke
from misai
where month(data)=6
and klientas not in (select klientas from #top400_pelnas_atstiktiniai)
group by klientas
having sum(kaina*kiekis)= 0
order by pirko_kartu desc
-- select * from #top400_pelnas_nepirke
-- drop table #top400_pelnas_nepirke

--linke atsitraukimui
select top 400 klientas, count (dok_nr) pirko_kartu, sum(kaina*kiekis) pelnas
into #top400_pelnas_atstiraukimui
from misai
where month(data)=6
and klientas not in (select klientas from #top400_pelnas_atstiktiniai)
and klientas not in (select klientas from #top400_pelnas_nepirke)
group by klientas
having sum(kaina*kiekis)< 20
and count (dok_nr)< 3
order by pirko_kartu desc

--sudarome nelojaliu klientu pirkimu istorija
-- select top 10 * from neloyalumo_istorija
--atsitiktiniu
declare @i int
set @i=1
while (@i< 30)
begin
declare cursorius cursor for
select klientas from #top400_pelnas_atstiktiniai (nolock)
open cursorius
declare
@klientas char (12),
@CLV numeric, --clv
@DW numeric, -- Baziniu transakciju per laika i skaicius
@UP numeric, -- dideles vertes transakciju vykdytu per laika i skaicius
@j int -- laikinas skaitliukas
fetch next from cursorius into @klientas

```

```

while @@fetch_status = 0 begin
--skaiciuojame up
set @UP = (
select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina > 3000 )
and (kaina < 12000) )
--skaiciujame DW
set @dw = (
select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina < 3000 ))
--skaiciuojame CLV
set @clv= @up+@dw
set @j=1 --(skaiciuojame laipsni)
while (@j< @i)
begin
set @clv=@clv/(1.07)
set @j=@j+1
end
--irasome paskaiciuotus parametrus i lenta nelojalumo_istorija
INSERT INTO nelojalumo_istorija
(klientas, diena, up, dw ,clv, menuo )
VALUES ( @klientas, @i, @up, @dw, @clv, 6)
--skaiciuojame dw
--select sum(dok_nr) @DW from misai
--where klientas=@klientas
--and month (data)=5
--and day (data) is between 1 and @i
--and preke in (select prek from #prekes_tipai where tip = 'Bazine')
--select @klientas klientas,@diena @up+@dw CLW
--into #TOP_PELNAS_CLW
set @up=0
fetch next from cursorius into @klientas
end
close cursorius
deallocate cursorius
set @i=@i+1
end

---skaiciuojame prikimu istorija linkusiems pasitraukimui klientams ir irasome i lenta

declare @i int
set @i=1
while (@i< 30)

```



```

begin
declare cursorius cursor for
select klientas from #top400_pelnas_nepirke (nolock)
open cursorius
declare
@klientas char (12),
@CLV numeric, --clv
@DW numeric, -- Baziniu transakciju per laika i skaicius
@UP numeric, -- dideles vertes transakciju vykdytu per laika i skaicius
@j int -- laikinas skaitliukas
fetch next from cursorius into @klientas
while @@fetch_status = 0 begin
--skaiciuojame up
set @UP = (
select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina > 3000 )
and (kaina < 12000) )
--skaiciujame DW
set @dw = (
select count(dok_nr)from misai
where
misai.klientas= @klientas
and month (data)=6
and (day (data) between 1 and @i)
and (kaina < 3000 ))
--skaiciuojame CLV
set @clv= @up+@dw
set @j=1 --(skaiciuojame laipsni)
while (@j< @i)
begin
set @clv=@clv/(1.07)
set @j=@j+1
end
--irasome paskaiciuotus parametrus i lenta neloyalumo_istoirija
INSERT INTO neloyalumo_istoirija
(klientas, diena, up, dw ,clv, menuo )
VALUES ( @klientas, @i, @up, @dw, @clv, 6)
--skaiciuojame dw
--select sum(dok_nr) @DW from misai
--where klientas=@klientas
--and month (data)=5
--and day (data) is between 1 and @i
--and preke in (select prek from #prekes_tipai where tip = 'Bazine')
--select @klientas klientas,@diena @up+@dw CLW
--into #TOP_PELNAS_CLW
set @up=0
fetch next from cursorius into @klientas
end

```

```
close cursorius
deallocate cursorius
set @i=@i+1
end
```

```
drop table #top400_pelnas_atsitraukimui
drop table #top400_pelnas_atsitiktiniai
--delete
--from lojalumo_istoirija
select count (row_id) from lojalumo_istoirija where menuo = 6
```

```
-- declare @i int
-- set @i=30
-- declare @up float, @klij char (12)
-- set @klij='KL0000004281'
-- set @UP = (
-- select count(dok_nr)from misai
-- where
-- misai.klientas= @klij
-- and month (data)=5
-- and (day (data) between 1 and @i)
-- and (kaina < 3000 )
-- --and (kaina < 12000) )
-- select @up
--
-- select top 100 * from lojalumo_istoirija
-- where clv = 0
-- order by klientas, diena
--
-- select top 400 klientas, count (dok_nr) pirko_kartu, sum(kaina*kiekis) pelnas
-- into #laikinas
-- from misai
-- where month(data)=5
-- group by klientas
-- order by pelnas asc
--
-- select * from #laikinas
-- where pirko_kartu>0 and pelnas=0
```