

ŠIAULIŲ UNIVERSITETAS  
MATEMATIKOS IR INFORMATIKOS FAKULTETAS  
INFORMATIKOS KATEDRA

**Donatas Tamašauskas**

Informatikos magistro specialybės II kurso dieninio skyriaus studentas

**Duomenų gavybos technologijos nuotolinių studijų  
kursų duomenų bazės analizei**

**Data Mining Technologies for the Analysis of Distance  
Learning Course Database**

Magistro darbas

Darbo vadovas:  
Doc. Dr. S. Turskienė

Recenzentas:  
Prof. Habil. Dr. L. Sakalauskas

2012 m. Šiauliai

*„Tvirtinu, jog darbe pateikta medžiaga nėra plagijuota ir paruošta naudojant literatūros sąrašą pateiktus informacinius šaltinius bei savo tyrimų duomenis“*

Darbo autorius \_\_\_\_\_  
(vardas, pavardė, parašas)

## **Darbo tikslai ir uždaviniai**

### **Tikslas**

Darbo tikslas – nuotolinių studijų kursų duomenų bazių analizė, paremta duomenų gavybos technologijomis.

### **Uždaviniai**

- *Moodle* sistemai statistikos ir ataskaitos įskiepių analizė;
- Institucijos padalinio VMA *Moodle* sistemų duomenų bazių analizė;
- Sukurti programinę įrangą susisteminančią VMA *Moodle* sistemų duomenų bazių duomenis;
- Atlikti gautos duomenų bazės duomenų gavybą, programa STATISTICA 7;
- Parengti rekomendacijas pagal duomenų gavybos proceso metu gautus rezultatus.

Darbo vadovas \_\_\_\_\_  
(vardas, pavardė, parašas)

# Turinys

Įvadas.....	6
1. Teorinė dalis .....	8
1.1. Temos analizė .....	8
1.1.1. Nuotolinio mokymosi apžvalga.....	8
1.1.2. Virtualių mokymosi aplinkų samprata.....	9
1.1.3. Duomenų gavybos sampratos.....	10
1.1.4. Duomenų gavybos metodų analizė.....	11
1.2. Darbinės srities analizė.....	14
1.2.1. Duomenų gavybos samprata nuotoliniame mokyme.....	14
1.2.2. Duomenų gavybos procesas virtualioje mokymosi aplinkoje.....	14
2. Projektinė dalis.....	16
2.1. Įrankių pasirinkimo analizė .....	16
2.2 Pradinis projekto aprašymas.....	19
2.2.1. Moodle sistemos statistikos ir ataskaitos įskiepių analizė .....	19
2.2.2. Institucijos padalinio VMA Moodle sistemų duomenų bazių analizė.....	19
2.2.3. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga .....	20
2.2.4. Gautos duomenų bazės duomenų gavybą programa STATISTICA 7 .....	20
2.2.5. Parengti rekomendacijas pagal duomenų gavybos proceso metu gautus rezultatus ..	20
3. Darbo eigos aprašymas.....	21
3.1. Galutinis projekto būvio aprašymas.....	21
3.1.1. Moodle sistemos statistikos ir ataskaitos įskiepių analizė .....	21
3.1.2. Institucijos padalinio VMA Moodle sistemų duomenų bazių analizė.....	26
3.1.3. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga .....	27
3.1.4. Gautos duomenų bazės ,duomenų gavybą programa STATISTICA 7 .....	33
3.1.4.1. Kiek ir kokius studijų dalykus studijuoja kiekvienas studentas .....	33
3.1.4.2. Studentai ir jų aktyvumas .....	36
3.1.4.3. Vartotojų prisijungimų kiekiai pagal pasirinktą intervalą .....	37
3.1.4.4. Vartotojų aktyvumo kiekiai pagal pasirinktą intervalą .....	40
3.2. Problemos ir jų sprendimo būdai .....	42

3.2.1 VMA <i>Moodle</i> sistemų duomenų bazių pildymo .....	42
3.2.2 VMA <i>Moodle</i> sistemų duomenų bazių susisteminanti programinė įranga .....	44
3.3. Rekomendacijos .....	44
3.3.1 VMA <i>Moodle</i> duomenų bazių pildymo .....	44
3.3.2 VMA <i>Moodle</i> serveriams .....	45
3.3.3 VMA <i>Moodle</i> sistemų vartotojams .....	45
3.4. Darbo rezultatų analizė .....	46
3.4.1. <i>Moodle</i> sistemai statistikos ir ataskaitos įskiepių analizė .....	46
3.4.2. Institucijos padalinio VMA <i>Moodle</i> sistemų duomenų bazių analizė .....	46
3.4.3. VMA <i>Moodle</i> sistemų duomenų bazių susisteminanti programinė įranga .....	47
3.4.4. Gautos duomenų bazės, duomenų gavybą programa STATISTICA 7 .....	47
Išvados .....	48
Informacinių šaltinių ir literatūros sąrašas .....	49
Anotacija .....	51
Priedai .....	52
1. Kompaktinio disko turinys .....	52
2. VMA <i>Moodle</i> sistemų duomenų bazes susisteminančios programinės įrangos diegimo instrukcijos .....	52
3. VMA <i>Moodle</i> sistemų duomenų bazių, susisteminanti programinė įranga .....	52
4. Studentų mokslinių darbų programa .....	53
5. Vartotojų prisijungimų kiekiai kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio .	60
6. Nuspėjimų mėnesių, pagal prisijungimų kiekį, kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio .....	61
7. Sekų paieška duomenyse , pagal prisijungimų kiekį kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio .....	62
8. Vartotojų aktyvumas kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio .....	63
9. Nuspėjimų mėnesių, pagal aktyvumo kiekį kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio .....	64
10. Sekų paieška duomenyse , pagal prisijungimų kiekį kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio .....	65

## Įvadas

Atsiradus pirmiesiems personaliniams kompiuteriams bei internetui ir jiems tapus visuotinai prieinamiems, natūraliai atsirado poreikis didinti duomenų srautus bei saugoti duomenis. Šiuolaikinės technologijos užtikrina didelių duomenų srautų gavimą ir jų saugojimą. Tačiau, vis dar lieka didelė spraga tarp duomenų surinkimo bei saugojimo, ir jų suvokimo. Šią spragą padeda užpildyti duomenų gavyba (angl. *data mining*), kurios duomenų apdorojimo metodai padeda suprasti duomenis, atskirti svarbią informaciją nuo menkavertės, pastebėti susidariusias grupes (klasterius), žymiai išsiskiriančius objektus (taškus-atsiskyrėlius) (angl. *outliers*), atstumus arba panašumus tarp objektų ir pan.[15]

Norint atlikti duomenų analizę, duomenų gavybos apdorojimo metodais, būtina surinkti iš heterogeninių duomenų šaltinių, transformuoti, išfiltruoti ir pritaikyti duomenis, duomenų gavybos apdorojimo metodams, specializuotoms programinėms įrangoms, tokioms kaip *SAS*, *StatSoft STATISTICA 7*, *STATGRAPHICS*, *SPSS Statistica* bei kitoms, kurios atlieka statistinę duomenų analizę, duomenų gavybos metodais prognozuoja ir klasifikuoja duomenis bei sprendžia atskiras problemas.[16]

Šiuo metu yra sukurta daug virtualių mokymosi aplinkų (VMA): *Sakai*, *Moodle*, *WebCT*, *Learning Space*, *First Class*, ir kt., kurios aktyviai naudojamos aukštosiose mokyklose.

Visos VMA kaupia didelį kiekį informacijos, ne tik apie jos vartotojų būseną (kokiam kursui priklauso, kokiam vaidmeniui (angl. *role*) priklauso ir kt.), bet ir apie atliktus veiksmus (darytos užduotys, spęsti testai, skaitytos temos ir kiti veiksmai, bei jų rezultatai) ir jų laiką (kada buvo prisijungęs vartotojas, kada sprendė testus, kada peržiūrėjo testus bei kitų atliktų veiksmų laikus).[5]

Labai svarbi mokymo procesuose yra operuojamų duomenų analizė. Kuo įvairesniais pjūviais atlikta duomenų analizė, galėtų padidinti galimybę planuoti mokymo srities veiklą, besimokančiųjų elgesiui atskleisti ir pan. Tai būtų labai naudinga ne tik nuotolinio mokymosi atžvilgiu, bet ir padėtų prognozuoti naudojamų serverių resursų kiekį, kuriuose būtų patalpintos sistemos.

Lietuvoje gana plačiai paplitusi atvirojo kodo *Moodle* VMA, kurios dėka galima patalpinti įvairius mokymosi kursus internetinėje terpėje, kurie leidžia bet kuriuo metu, iš bet kurios vietos pasiekti ne tik reikiamą informaciją ar atlikti įvairias užduotis, bet ir bendrauti su kitais sistemos vartotojais.

**Tyrimo aktualumas ir naujumas.** Duomenų gavybos technologijos yra labai išpopuliarėjusios, ir dabar yra nuolat taikomos beveik visose korporacijose: telekomunikacijų, finansinių paslaugų, prekybos, medicinos, gamybos, švietimo, ir kitokių tipų kompanijose bei valstybinėse įstaigose.

Tačiau duomenų gavybos technologijų VMA duomenų analizė pradėta taikyti gana neseniai. Šių sričių apjungimas yra vis dar naujas dalykas. Šią sritį tyrinėja G. Dzemyda, P. Gudynas, V. Saltenis, V. Tiesis (2001 m.), atlikdami mokytojų kvalifikacijos, amžiaus kaitos bei jų skaičiaus dinamikos tyrimą taikant dirbtinius neuronų tinklus, panaudojant *Ms Excel* programinį įrankį. R. Burbaitė, G. Kulvietis, S. Turskienė (2003 m.), atliko Jaunųjų kompiuterininkų mokyklos (JKM) vartotojų duomenų analizę, panaudojant įrankį *IBM DB2 Intelligent Miner*. Saulius Preidys, Leonidas Sakalauskas (2011 m.) taikė duomenų gavybos metodus, analizuojant virtualią mokymo aplinką įvairiais pjūviais, panaudojant *Ms Excel*. [12, 13, 14]

**Darbo tikslas ir uždaviniai.** Šiame darbe keliamas tikslas – nuotolinių studijų kursų duomenų bazių analizė, paremta duomenų gavybos technologijomis.

**Tiksli pasiekti formuluojami šie uždaviniai:**

- *Moodle* sistemai statistikos ir ataskaitos įskiepių analizė;
- Institucijos padalinio VMA *Moodle* sistemų duomenų bazių analizė;
- Sukurti programinę įrangą susisteminančią VMA *Moodle* sistemų duomenų bazių duomenis;
- Atlikti gautos duomenų bazės duomenų gavybą, programa STATISTICA 7;
- Parengti rekomendacijas pagal duomenų gavybos proceso metu gautus rezultatus.

**Tyrimo objektas.** Šiaulių universiteto nuotolinių studijų centro *Moodle* sistemų duomenų bazių analizė.

**Tyrimo metodai.** Mokslinės ir techninės literatūros analizė, duomenų analizė.

**Rezultatų skelbimas.** Skaitytas pranešimas studentų mokslinėje konferencijoje „Studentų moksliniai darbai“. [26]

# 1. Teorinė dalis

## 1.1. Temos analizė

### 1.1.1. Nuotolinio mokymosi apžvalga

Nuotolinis mokymas, kuris dar dažnai vadinamas distancinių mokymų, virtualių mokymų ar e-mokymo sąvokomis, apibūdinamas, kaip bet koks mokymas, kai mokantysis (ar mokomąją medžiagą pateikiantysis) yra kitoje vietoje, negu besimokantysis. Šį principą leidžia realizuoti informacinės ir komunikacinės technologijos. Viena seniausių nuotolinio mokymo formų buvo mokymas susirašinėjant. Lietuvoje tokio mokymosi pavyzdys buvo fizikos mokykla „Fotonas“.

[1, 2, 3, 17]

Literatūroje ir informaciniuose šaltiniuose galima rasti įvairių nuotolinio mokymosi sąvokos paaiškinimų:

- Petersas teigia (Peters, 1973), kad distancinis mokymas - žinių, įgūdžių, požiūrių perteikimo mechanizmas, naudojantis daug techninių informacijos priemonių, ypač skirtingose vietose tuo, pačiu metu esantiems studentams mokytis. Tai technizuota mokymo ir švietimo forma; [2]
- Dellingas pažymėjo (Delling, 1986), jog "distancinis švietimas - tai tokia mokymosi forma, kai besimokantieji yra fiziškai atskirti nuo mokytojų. Toks atskyrimas gali būti pritaikytas visam mokymosi procesui, arba tik kai kuriems mokymosi etapams. Galimos tiek akivaizdinės, tiek asmeninės studijos, bet jų funkcijos papildo ar sustiprina vyraujančią distancinę bendravimą"; [19]
- Makholas Tigtas akcentavo (Malcolm Tigt, 1988), kad "distanciniame švietime vyksta bendravimas tarp mokytojo, studento ir techninės informacijos priemonės, reikalingos tokiam bendravimui palaikyti"; [2, 19]
- M. Moore (1990) pateikia bendrą, kiekvienai sistemai, įmonei, metodologijai bei situacijai tinkanti distancinio švietimo apibrėžimą: "Distancinis švietimas yra dialogų struktūra ir savarankiškumu paremtas ryšys, kuriam reikia techninių tarpininkavimo priemonių (media)"; [18]
- Kaip teigia Laužackas (Laužackas, 2005), nuotolinis mokymas(-is) – mokymo(-si) forma, kai besimokantysis nepalaiko betarpiško kontakto su mokytoju. Bendravimas ir



komunikavimas yra užtikrinimas IKT (informacinės ir telekomunikacinės technologijos) priemonėmis. [4]

### 1.1.2. Virtualių mokymosi aplinkų samprata

Šiandien dažnai girdime sakant nuotolinis mokymasis naudojant virtualią mokymosi aplinką, todėl, kad suprasti, kas yra virtuali mokymosi aplinka, pirmiausia reikia išanalizuoti šias tris sąvokas:

1. „Virtuali“ nusako, kad tai susiję su kompiuteriais, tai nėra įprastas mokymo būdas klasėje;
2. „Mokymosi“ – sąvoka yra aiški, tačiau reikia atkreipti dėmesį į „-si“ (ne mokymo, o mokymosi!), kuri parodo besimokančiojo aktyvumą;
3. „Aplinka“ – ją galima suprasti keleriopai, tačiau svarbiausia yra vieta, kurioje susitinka mokiniai bei mokytojas (virtuali vieta), ir susitinka ne šiaip sau, o veikti, mokytis.

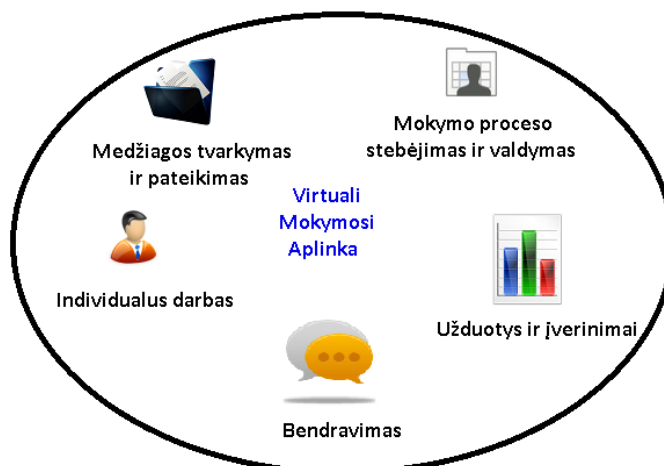
Taigi, išanalizavus šias tris sąvokas galima teigti, kad virtuali mokymosi aplinka (VMA) — tai mokymosi aplinka, realizuota kompiuterinėmis priemonėmis, kurioje pateikiama mokymosi medžiaga, organizuojamas mokymosi procesas bei atliekamas jo valdymas.

Terminas „virtualioji mokymosi aplinka“ bendrąja prasme vartojamas norint įvardinti bet kurią iš toliau išvardintų sistemų, ar jų derinį:

- Mokymosi valdymo sistema (angl., *learning lanagement system*). Tai gali būti ir paprasta besimokančių (ir mokinančių) registravimo sistema, leidžianti prieiti prie mokymosi medžiagos kompiuterių tinkle, bei sudėtinga sistema, stebinti besimokančių mokymosi eigą, ir pagal tai atliekanti kitas ugdymo proceso organizavimo funkcijas.[5]
- Mokymosi turinio tvarkymo sistema (angl., *learning content management system*). Tai individualiems poreikiams pritaikoma mokymosi tvarkymo priemonė. Pavyzdžiui, ji gali atlikti mokymosi medžiagos dalių, kurias besimokantis naudojo, stebėjimą, sekti besimokančiojo atsiskaitymus bei pagal tai parinkti, ar pritaikyti atitinkamą ugdymo turinį.[6]
- Kursų tvarkymo sistema (angl., *course management system*). Ši priemonė leidžia mokinančiam nenaudojant HTML, arba programavimo kalbos parengti mokymo kursą ir kompiuterių tinkle pateikti jo medžiagą bei kitą su mokymusi susijusią informaciją.[23]
- Kolektyvinio kompiuterizuoto mokymosi aplinka (angl., *computer supported collaborative learning environment*). Tai sistema, skirta mokytis kognityviniais metodais bei dirbant grupėmis konstruoti žinias.[6]

- Turinio tvarkymo aplinka (angl., *content management system*). Ši priemonė skirta mokymosi medžiagai parengti. Ji suteikia galimybes mokytojui lanksčiai tvarkyti mokomąją medžiagą: kurti naujus modulius, paimti informaciją iš kitų šaltinių, pertvarkyti, pateikti įvairiais būdais.[6]
- Ugdymo tvarkymo aplinka (angl., *managed learning environment*). Tai platesnė sąvoka nei VMA, ji apima visas, kurios nors institucijos informacines sistemas bei procesus, turinčius tiesiogines ar netiesiogines įtakas, ugdymui bei jo tvarkymui. Ugdymo tvarkymo aplinka apima ir virtualiąją mokymosi aplinką. [6]

VMA vyksta visas ugdymo procesas (Pav. 1): pateikiamas visas kurso ar modulio turinys, bendraujama diskutuojant (diskusijų forumuose, pokalbiuose ar elektroniniu paštu), atliekamos praktinės užduotys, vyksta darbas grupėse, o įgytos žinios bei gebėjimai tikrinami kompiuteriniais testais, vertinama automatinėmis priemonėmis ir pan.[7]



1. VMA ugdymo procesai

VMA pasaulyje parengta ganėtinai daug, praktiškai beveik kiekviena šalis, kasmet jų sukuria po kelias dešimtis. Dažnai girdima minint *WebCT*, *Learning Space*, *First Class*, *Moodle*, *Sakai* ir kitas aplinkas.

### 1.1.3. Duomenų gavybos sampratos

Įvairiausi duomenys (meteorologinės sąlygos ir pan.) buvo kaupiami įvairiais laikotarpiais, o šių duomenų kaupimas jų neapdorojus neturi prasmės, todėl duomenų analizė buvo visuomet atliekama. Visų rūšių duomenų tvarkymas yra pirmas žingsnis norint palengvinti duomenų analizę, apdoroti didelį informacijos srautą. Viena iš informacijos analizės sričių — duomenų gavyba (angl., *data mining*). Tai šiuolaikinė informacijos analizės sritis, atsiradusi

duomenų bazių technologijų, dirbtinio intelekto bei statistinės duomenų analizės sankirtoje. Šiuolaikinės duomenų gavybos technologijos pagrindas yra šablonai (angl., *patterns*). [11]

Keletas duomenų gavybos apibrėžimų:

1. Pasak "Gartner Group", duomenų gavyba yra prasmingų šablonų (angl., *patterns*), dėsningumų, modelių ir tendencijų radimo procesas dideliuose informacijos kiekiuose, pasinaudojant modelių atpažinimo, statistiniais bei matematiniais metodais; [8]
2. Dažniausiai naudojamas G. Piatecki-Shapiro duomenų gavybos apibrėžimas – duomenų gavyba yra neapdorotų duomenų (angl., *raw data*) tyrinėjimo procesas žinioms nustatyti, kurios yra:
  - naujos, prieš tai nežinotos;
  - netrivialios;
  - praktiškai naudingos;
  - interpretuotinos;
  - būtinos sprendimams priimti pasirinktoje veiklos srityje. [8]
3. Antai U.Fayyad, S.Chaudhuri and P.Bradley teigia, kad duomenų gavyba – įdomių struktūrų (modelių, ryšių, statistinių modelių, šablonų) radimas duomenų bazėse. [20]
4. M.J.A.Berry ir G.S.Linoff teigia, kad duomenų gavyba – didelių duomenų kiekių tyrinėjimas bei analizė automatizuotu arba pusiau automatizuotu būdu, siekiant rasti naudingus modelius (angl. *patterns*) ir taisykles. [20]

Duomenų gavyba yra labai plati sritis, apimanti daug metodų, algoritmų bei taikomųjų sistemų. Jei įprasti duomenų analizės metodai padeda atskleisti tiriamų kintamųjų priklausomumą, tai duomenų gavyba unikali tuo, kad analizės rezultatų yra naujų priklausomybių, apie kurių egzistavimą buvo, ar net nebuvo įtariama, radimas.

#### **1.1.4. Duomenų gavybos metodų analizė**

Reikalingų duomenų modelių ar taisyklių radimui galima panaudoti kompiuterį — pagrindinė duomenų gavybos idėja. Duomenų išgavimo technika ir algoritmai priklauso ne tik nuo pačių duomenų, bet ir nuo jų kilmės, užduoties struktūros, ir pan. Vienas iš svarbiausių duomenų išgavimo privalumų yra tai, kad programinė įranga gali rasti tokius modelius, kurie nėra lengvai pastebimi bei akivaizdūs, tačiau labai naudingi tiriamoje srityje. [8]

Turint tam tikrą duomenų rinkinį, galima pradėti analizės procesą. Pirmiausia jam nustatoma metodologija, duomenų struktūros atvaizdavimui. Kai randamos tam tikros žinios ar

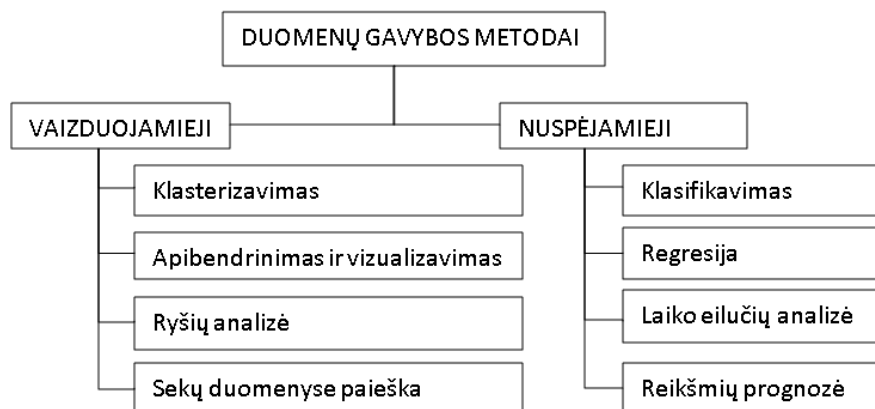
taisyklės, programinei įrangai pateikiami didesni duomenų rinkiniai, kurie turi panašią duomenų struktūrą.

Duomenų gavyboje yra naudojama daugybė algoritmų, vienokiems ar kitokiems uždaviniams atlikti. Visi šie algoritmai, analizuoja duomenis bei mėgina pritaikyti modelį labiausiai atitinkanti duomenų charakteristikas.

Duomenų gavybos metodus galima skaidyti į du tipus:

1. Vaizduojamieji duomenų gavybos metodai — skirsto duomenis į grupes, atlieka klasifikavimą. Vaizduojami duomenų gavybos metodai priešingai, nei nuspėjamieji arba prognozės duomenų gavybos metodai, neranda naujų duomenų, o analizuoja esamus duomenis ir priklausomai nuo to kokie metodo algoritmai yra naudojami, vaizduojamieji skirstomi į klasterizavimo, asociacijų bei priežastinių ryšių metodus.
2. Nuspėjamieji arba prognozės duomenų gavybos metodai — atlieka duomenų prognozavimą, kintamųjų priklausomybių tyrimą, vertina kintamųjų faktines vertes, kurie gali remtis kitų duomenų rezultatais.

Pateiktoje schemeje (Pav. 2) pavaizduoti fundamentalūs duomenų gavybos metodai. Šių metodų kompleksiškas taikymas veda į naujų jungtinių metodų apibrėžtis [11].



2. Duomenų gavybos metodai

Keletas duomenų gavyboje naudojamų metodų:

- Klasterizavimas (Grupavimas) — dažniausiai būna vienas pirmųjų duomenų gavybos žingsniu. Tai visos duomenų aibės suskaidymas į poaibius, pagal skiriamuosius bruožus, pavyzdžiui, studentų grupavimas pagal bruožus į grupes.[8] Šį metodą naudojo Turskienė S., Kulvietis G., Burbaitė R. [13] tyrinėjant lyties demografinę klasterizaciją.
- Apibendrinimas ir vizualizavimas — norint suvokti tiriamus duomenis: įvertinti statistikas, tokias kaip vidurkis, dispersija, standartinis nuokrypis, histogramos ir pan.,

nustatyti duomenų pasiskirstymą, iškelti hipotezes, kurias duomenų gavybos proceso metu teks patvirtinti ar paneigti. [15]

- Ryšių (Asociacijų) analizė – apima uždavinius, kuriuose reikia nustatyti ryšius, koreliacijas tarp duomenų bei jų parametrų. Gerai žinomas ryšių analizės pavyzdys yra laikai, kuomet studentai prisijungia prie *Moodle* virtualios mokymosi aplinkos, ši analizė parodo ryšius tarp laiko ir studentų prisijungimų prie *Moodle* virtualios mokymosi aplinkos. [15]
- Sekų paieška duomenyse — apima duomenų gavybos metodikas, nustatančias tam tikrų įvykių dažnumus, kurie yra pritaikomi duomenų rinkinių analizei. Šiuo metu labai aktualu rasti, tarp didelių duomenų masių mums svarbią informaciją, kurią būtų galima panaudoti ateityje. Vienas iš svarbiausių jos tikslų yra dažnų pasikartojimų radimas. [11]
- Klasifikavimas — dažniausiai atliekamas po grupavimo. Kai nagrinėjama aibė, jau padalinta į pogrupius, dažnai kyla klausimas, kam priskirti naujus elementus. Grupavimu anksčiau neįvardinti poaibiai išskiriami iš duomenų visumos, o klasifikuojant sprendžiama, kaip sudėti elementus į žinomas grupes. Klasifikavimo tikslas – identifikuoti parametrus, kurie nusakytų grupę (klasę), kuriai priklauso objektas. [8]
- Regresija — prognozuojant naujų duomenų reikšmes, naudojasi žinomais, jau turimais duomenimis. Ji naudoja standartinius statistinius metodus, pavyzdžiui, tiesinę regresiją. [15]
- Laiko eilučių analizė — naudojant laiko eilučių analizės metodus yra tiriamas parametro reikšmių kitimas laike. Parametro reikšmės yra gaunamos fiksuotais laiko momentais (kas dieną, kas savaitę, kas valandą ir pan.). Prognozuojant laiko eilutės būsimas reikšmes remiamasi anksčiau gautais duomenimis, panašiai kaip ir regresijoje. Sukurtas laiko eilutės modelis atspindi laiko ypatumus, sezoniškumą, tokius kalendoriaus ypatumus kaip atostogos bei švenčių dienos ir pan. [22]
- Reikšmių prognozė — apibrėžiama kaip tam tikra klasifikavimo atmaina. Skirtumas tas, kad prognozė nusako būsimą duomenų reikšmę, o ne klasifikuoja jau esamą. Dažnai duomenų gavybos uždaviniui išspręsti taikomi keli metodai iš eilės, ar net sudėtingi jų deriniai. [15]
- Ir kiti.

## **1.2. Darbinės srities analizė**

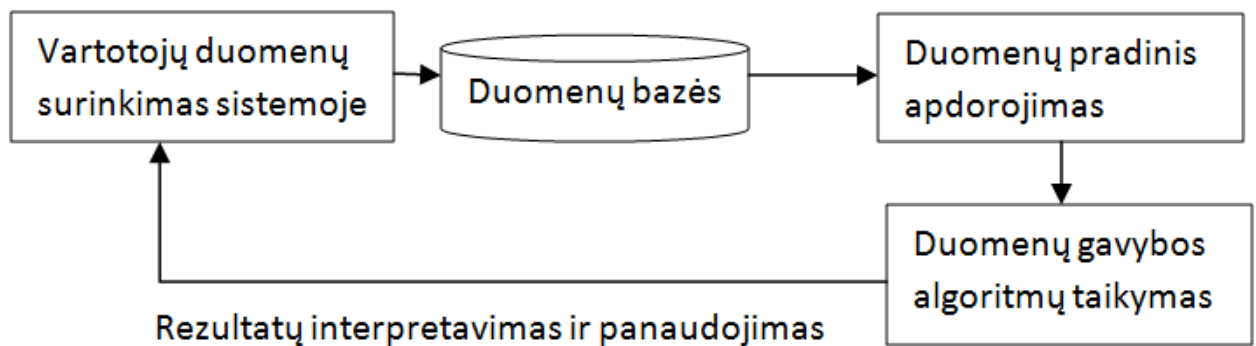
### **1.2.1. Duomenų gavybos samprata nuotoliniame mokyme**

Nuotolinio mokymosi sritis yra pakankamai nauja disciplina Lietuvoje ir duomenų gavybos taikymas būtų vienas iš įdomesnių bei naudingų faktorių jos tyrime. Pasaulyje nuotolinis mokymas yra gerai žinomas ir tiriamas. Nepaisant to, jis lieka vienas iš įdomesnių duomenų gavybos taikymo sričių. Pagrindinė priežastis, kodėl yra taikomos duomenų gavybos technologijos yra tai, kad visose populiariausiose VMA (*WebCT, BlackBoard, Moodle* ir kt.), visų vartotojų veiksmai yra saugomi duomenų bazėse. Ten yra fiksuojami tokie veiksmai, kaip: prisijungimas prie atitinkamo įrankio (paštas, forumai, mokymo medžiaga), dalyvavimas virtualiuose pokalbiuose ir pan. Todėl pradėta kurti įrankius, skirtus vartotojams stebėti, jų veiksams analizuoti. Jų padedamas kurso kuratorius gali gauti informaciją apie vartotojų naudotus VMA įrankius, prisijungimo trukmę, testavimo rezultatus ir pan. Kuratorius, žinodamas šią informaciją, gali valdyti kuruojamą vartotojų grupę, aktyvinti pasyvesnius vartotojus, suteikti daugiau veiklos aktyviesiems ir t.t. , kadangi studijų kokybė priklauso nuo visų studijų procese dalyvaujančių vartotojų: kokybiškai parengtas nuotolinis modulis bus patogus besimokančiajam, palengvins darbą kuratoriui, nuo kurio veiksmų, taip pat labai priklauso nuotolinių studijų proceso kokybė. Nuo administratoriaus, atsakingo už VMA serverio nepertraukiamą darbą veiklos, taip pat priklauso visas studijų procesas – nustojus veikti serveriui, bent po vieną valandą per parą, studijos taps neprognozuojamos ir nepatrauklios vartotojams bei kuratoriams. Kad šis procesas vyktų sklandžiai ir koordinuotai, būtinas visų dalyvių grupių nenutrūkstamas darbas: besimokančiųjų veiklos stebėjimas bei skatinimas, grįžtamojo ryšio užtikrinimas, galimų techninių problemų prognozavimas, nepavėluotas jų šalinimas ir pan. Tačiau, didėjant nuotolinių studijų kursų pasiūlai ir tokias studijas besirenkančių vartotojų skaičiui, darosi vis sunkiau prižiūrėti anksčiau minėtus procesus. Dėl šių priežasčių nukenčia nuotolinių studijų kokybė, vartotojai praranda norą studijuoti šiuo būdu, prastėja ir paties kuratoriaus darbo kokybė. [9] [10]

### **1.2.2. Duomenų gavybos procesas virtualioje mokymosi aplinkoje**

Duomenų gavybos taikymas yra pasikartojantis ciklas, todėl VMA duomenų gavybos taikymas yra pasikartojantis ciklas, kadangi duomenų gavybos pagalba išgaunama informacija turėtų būti panaudota VMA.

VMA duomenų gavybos procesas susideda iš keturių pagrindinių duomenų gavybos žingsnių (Pav. 3):



3. Duomenų gavybos procesai virtualioje mokymosi aplinkoje

- Duomenų rinkimas: VMA kaupia visą informaciją apie vartotoją (jo atributus, veiksmus ir t.t.).
- Duomenų pradinis apdorojimas: duomenys yra paimami iš duomenų bazės bei paruošiami duomenų gavybos algoritmų taikymui.
- Duomenų gavyba: duomenų gavybos algoritmai pritaikomi, siekiant sukurti ir vykdyti modelius atrandančius, ir apibendrinančius vartotojus, kurso kuratorius bei administratorius dominančią informaciją. Šiam tikslui gali būti panaudoti bendro pobūdžio ar specifiniai duomenų gavybos įrankiai.
- Rezultatų interpretavimas ir panaudojimas: modelis ar rezultatai yra interpretuojami ir panaudojami tolesniam kurso tobulinimui. Kurso kuratoriai gali pasinaudoti atrastą informaciją priimdami sprendimus, dėl besimokančių mokymosį procesų ir kurso veiklų tobulinimo.

## 2. Projektinė dalis

### 2.1. Įrankių pasirinkimo analizė

Institucijos padalinyje yra naudojama keturiolika *Moodle* sistemų, kurios yra įdiegtos į *MySQL* (*Moodle* sistemos, taip pat gali būti suderinamos su *PostgreSQL* ir kitomis duomenų bazėmis) duomenų bazes. Pagrindinis kuriamos institucijos padalinio susisteminančios VMA *Moodle* sistemų duomenų bazių duomenis programinės įrangos tikslas – kad būtų galima ja naudotis per interneto naršyklės, todėl buvo nuspręsta įrankį kurti PHP programavimo kalba.

Kuriamo įrankio viena iš funkcijų yra duomenų atvaizdavimas, norint jam suteikti galimybę dinamiškai filtruoti ar rūšiuoti duomenis, pagal pasirinktus kriterijus, buvo panaudota *HTML Table Filter Generator* biblioteka. Taip pat buvo integruotos jos papildomos bibliotekos, tai *Sortable table* — skirta HTML lentelės dinaminiam rūšiavimui ir *Columns resizer* — skirta *HTML* lentelės dinaminiam stulpelių pločio redagavimui. Ši biblioteka lyginant su kitomis labai panaši į *DataTables* biblioteką, tačiau nuspręsta pasirinkti pastarąją, dėl geros dokumentacijos, bei pavyzdžių gausos (1 lentelė).

Pavadinimas	Nemokama	Įskiepių įvairovė	Dinaminis filtrų veikimas	Visų duomenų tipų palaikymas	Lietuvių kalbos palaikymas	Dokumentacija
HTML Table Filter Generator	+	+	+	+	+	+
Sort and filter HTML table	+	-	+	+	+	+
zFilter	+	-	+	+	-	-
FilterTable	+	-	+	+	-	-
DataTables	+	+	+	+	+	-

1 lentelė. HTML lentelių dinaminiam rūšiavimui bibliotekų analizė

Norint suteikti kuriamam įrankiui galimybę atvaizduoti duomenis diagramose, buvo nuspręsta panaudoti *JpGraph*, kuris yra vienas iš populiariausių šiuo metu naudojamų braižymo bibliotekų. Pagrindiniai pasirinkimo kriterijai, kad bibliotekos turėtų savybes braižyti skritulines, stulpelines ir linijines diagramas. Kadangi, visos analizuojamos bibliotekos (2 lentelė) turėjo šias savybes, tai pirmenybė buvo teikiama išsamiai dokumentacijai. Iš visų įrankių labiausiai išsiskyrė savo savybėmis *Google Chart API* įrankis, kuris suteikia pasinaudojant *Google Api* braižyti



diagramas iš *Google* serverių, tačiau šio įrankio atsisakyta, dėl nepriklausomybės nuo *Google* serverių.

	Programavimo technologija	Stulpelinės diagramos	Skritulinės diagramos	Linijinės diagramos	Papildoma informacija
<i>JpGraph</i>	OOP	+	+	+	Gerai dokumentuota, didelė diagramų įvairovė, patogios klasių sąsajos, daug pavyzdžių
<i>pChart</i>	OOP	+	+	+	Duomenis galima atvaizduoti iš <i>SQL</i> užklausos. Didelė diagramų įvairovė.
<i>Google Chart API</i>	API	+	+	+	Visos funkcijos yra saugomos <i>google</i> serveriuose, todėl nereikia savo puslapio serveryje talpinti bibliotekų ar papildomų modulių.
<i>ezComponents</i>	OOP	+	+	+	Neturi galimybės kurti kitokias diagramas

2 lentelė. Braižymo bibliotekų įrankių analizė

Kiti pasirinkti įrankiai programinio įrankio kūrimui yra:

1. *NetBeand IDE (6.9.1)* — Šis *IDE*, kuris skirtas puslapių kūrimui pastaruoju metu sparčiai populiarėja tarp *PHP* programuotojų. Greičiausiai todėl kad, aktyviai vystomas projektas ir periodiškai pridedamos naujos funkcijos, buvo galima rinktis iš tokių įrankių, kaip *dreamviewer* ir pan.
2. *SQL Manager Lite* — tai nemokamas įrankis, kuris buvo pasirinktas, dėl jo gebėjimo vizualiai kurti *SQL* užklausas.
3. *MySQL* — duomenų bazių valdymo sistema (DBVS).

Duomenų analizei pasirinkta specializuota duomenų gavybos programinė įranga *STATISTICA 7*, kuri lyginant su kitais duomenų gavybos įrankiais (3 lentelė) savo galimybėmis yra panašus, šis įrankis buvo pasirinktas kadangi, testavimo metu dažnai pristatomas kaip aukšto lygio programinis produktas. Reikia išskirti, kelis pagrindinius *STATISTICA 7* privalumus – aukštą skaičiavimų tikslumą, bei patogią interaktyvaus dialogo sistemą.

	Privalumai	Kaina	OS	Dokumentacija	Populiarumas Lietuvoje
SPSS Clementine	Patogi vartotojo sąsaja, Algoritmų įvairovė	Kaina prasideda 26 800 \$	Windows	Labai gerai dokumentuotas	Nėra labai populiarus, lyginant su kitomis programomis.
DB2 Intelligent Miner	Algoritmų įvairovė, patogi vartotojo sąsaja, daugybė grafikų braižymo funkcijų	Kaina prasideda 15 000 \$	Windows, Unix	Gerai dokumentuota	Šio produkto populiarumas krenta, vis mažiau atliekama duomenų analizė šia programine įranga.
Statistica 7	Algoritmų įvairovė, įvairių būdų duomenų vizualizacija, patogi vartotojo sąsaja	Kaina prasideda nuo 15 000 \$ vienai licenzijai	Windows	Labai gerai dokumentuota, yra daug video medžiagos	Gana plačiai paplitęs aukštosiose mokyklose.
WEKA	Patogi vartotojo sąsaja, Algoritmų įvairovė.	Nemokama	Daugiaplatformė	Pakankamai gerai dokumentuota	Gana aktyviai naudojamas, kaip atvirojo kodo produktas.

3 lentelė. Duomenų gavybos įrankių analizė

## 2.2 Pradinis projekto aprašymas

### 2.2.1. Moodle sistemai statistikos ir ataskaitos įskiepių analizė

Yra daugybė produktų skirtų *Moodle* sistemų duomenų bazių analizei, taip pat nemažai informatikos studentų rašo darbus susijusius su duomenų gavybos metodų panaudojimu, analizuojant savo universiteto *Moodle* sistemų duomenų bases.

Tam, kad susipažinti su sukurtais produktais, kurie yra skirti *Moodle* sistemai duomenų bazių analizei, reikia atlikti oficialiame *Moodle* tinklalapyje siūlomus duomenų analizės įskiepių skirtų *Moodle* sistemos 1.9.x versijai ir 2.x versijai analizę.

Prieš pradėdant analizuoti pastebėta, kad oficialiame *Moodle* tinklalapyje duomenų gavybos įskiepių nepateikiama, yra tik statistikos ir ataskaitos įskiepių, todėl nuspręsta atlikti statistikos ir ataskaitos įskiepių analizę siekiant įvertinti, kokie dabar yra sukurti statistikos ir ataskaitos įskiepiai.

### 2.2.2. Institucijos padalinio VMA Moodle sistemų duomenų bazių analizė

Norint pritaikyti duomenų gavybos metodus, analizuojant institucijos padalinio VMA *Moodle* sistemų duomenų bazių duomenims, reikia atlikti visų analizuojamų duomenų bazių analizę. Ši analizė, taip pat bus naudojama kuriant programinę įrangą, susisteminančią VMA *Moodle* sistemų duomenų bases.

Konsultuojantis su institucijos padalinio vienu iš *Moodle* sistemų administratoriumi, bus siekiama išsamiai išanalizuoti duomenų bases bei vartotojų pasiskirstymą jose, šiam tikslui keliami šie reikalavimai:

1. Kiek yra *Moodle* duomenų bazių;
2. Kokios yra naudojamos *Moodle* duomenų bazės versijos;
3. Kokie yra vartotojai;
4. Kaip grupuojami vartotojai;
5. Kaip vartotojai pasiskirstė duomenų bazėse;
6. Kaip administruojami vartotojai;
7. Kiti pastebėjimai.

### **2.2.3. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga**

Reikės sukurti universalią programinę įrangą, susisteminančią institucijos padalinio VMA Moodle sistemų duomenis. Šiai programinei įrangai keliami šie reikalavimai:

1. Duomenų gavimas iš neriboto kiekio Moodle sistemų duomenų bazių nepriklausomai nuo jų versijos.
2. Analizuotų visų duomenų bazių duomenis iš karto;
3. Duomenų peržiūra įvairiais pjūviais.
4. Statistinė duomenų analizė įvairiais pjūviais.
5. Duomenų filtravimas ir rūšiavimas įvairiais kriterijais;
6. Atvaizduotus duomenis išsaugoti duomenų bazėje.
7. Išsaugoti atvaizduoti duomenis į CSV failo formatą.
8. Vartotojų autentifikavimas.

### **2.2.4. Gautos duomenų bazės duomenų gavybą programa STATISTICA 7**

Duomenis gauti iš sukurtos programinės įrangos sisteminančios VMA Moodle sistemų duomenų bazių duomenis, bus analizuojami duomenų gavybos metodais. Duomenų analizę atliekant, kaip įmanoma įvairesniais pjūviais naudojantis STATISTICA 7 programinės įrangos duomenų gavybos paketu.

### **2.2.5. Parengti rekomendacijas pagal duomenų gavybos proceso metu gautus rezultatus**

Pagal gautus duomenų gavybos proceso metu rezultatus parengti rekomendacijas, kurios sietųsi ne tik su programinės įrangos, sisteminančios VMA Moodle sistemų duomenų bazių duomenų gavybos rezultatais, bet ir su institucijos padalinio VMA Moodle duomenų bazių pildymo rekomendacijomis.

## 3. Darbo eigos aprašymas

### 3.1. Galutinis projekto būvio aprašymas

#### 3.1.1. Moodle sistemai statistikos ir ataskaitos įskiepių analizė

*Moodle* sistemai yra sukurta daugybė įvairių modulių ir įskiepių, kurie yra skirstomi į kategorijas: *Activity Module*, *Admin Report*, *Assignment Type*, *Authentication Method*, *Block*, *Course format*, *Course report*, *Database field*, *Enrolment Method*, *Filter*, *Grade Export*, *Grade Report*, *Integration*, *Message Processor*, *Package*, *Question Import/Export Format*, *Question Type*, *Quiz Report*, *Resource Type*, *Repository*, *SSO Plugin*, *Major Patch*, *Small Hack*, *User Profile Field* ir kiti. Tačiau iš tokios gausos modulių bei įskiepių, labai mažą dalį sudaro statistikos ir ataskaitos įskiepiai, o duomenų gavybos įskiepių išvis nėra.

Standartinė *Moodle* sistema turi bazinius statistikos ir ataskaitų modulius, tačiau jų nepakanka, trūksta tokių modulių, kuriais pasinaudojus būtų galima gauti norimą informaciją apie kiekvieno vartotojo ar grupės veiksmus *VMA* (kiek laiko vartotojas ar grupė praleido tinklalapyje, kiek laiko sugaišo vartotojas ar grupė kiekvienoje veikloje ir pan.). Tai būtų labai naudinga informacija, kadangi dabar *Moodle* galima gauti informaciją, tik apie pažymius ir testus. Todėl norint išplėsti *Moodle* galimybes reikia įdiegti papildomus įskiepius.

*Moodle* tinklalapyje (kaip jau minėta anksčiau) galima rasti daugybę įskiepių, tarp jų galima rasti tik statistikos ir ataskaitos įskiepių. Statistikos ir ataskaitos įskiepius galima suskirstyti į grupes:

1. Administratoriaus ataskaitos (angl., *Admin reports*) — tai įvairiausios ataskaitos, kurios būtų įdomios administratoriui. Pvz. (Sąrašas kiekvieno kurso ir jo užimama vieta, ir pan.);
2. Kursų ataskaitos (angl., *Course reports*) — tai įvairiausios ataskaitos susijusios su kursais. Pvz. (Bendras visų kursų vidurkis ir pan. ).
3. Pažymių ataskaitos (angl., *Grade reports*) — tai įvairiausios ataskaitos susijusios su vartotojų įvertinimu. Pvz. (Galimybė peržiūrėti keleto kursų vartotojų pažymius ir pan.).
4. Testų ataskaitos (angl., *Quiz reports*) — tai įvairiausios ataskaitos susijusios su testais. Pvz. (Parodyti visus testų atsakymus lentelėje).
5. Kiti — tai visos kitos ataskaitos nesusijusios, su nei viena aukščiau paminėta grupe. Pavyzdžiui, (*Certificate* modulio sukurtas ataskaitos įskiepis parodantis vartotojus, kuriems priskirti sertifikatai ir pan.)

### 3.1.1.1. Administratoriaus ataskaitos

Įskiepis *User's roles report* (Pav. 4), gana paprastas, tačiau naudingas, jis leidžia patikrinti kokius vaidmenis priskirti vartotojui, tai galima pamatyti įvedus vartotojo vardą. Jo didžiausias trūkumas, kad negalima pasirinkti vartotojo iš sąrašo ir gauti informacijos apie jo priskirtus vaidmenis.

**User's role assignments report**

**Report settings**

Username

#### 4. User's roles report

Įskiepis *Module use* (Pav. 5) parodo kiek kursų naudoja kiekvieną modulį ir kiek modulių pavyzdžių yra naudojama kiekviename kurse.

Activity module	Activities
Assignment	0
Chat	0
Choice	0
Database	0
Forum	5 instances in 4 courses
Glossary	3 instances in 2 courses
Hot Potatoes Quiz	0
Journal	0
Label	17 instances in 2 courses
LAMS	0
Lesson	1 instance in 1 course

#### 5. Module use

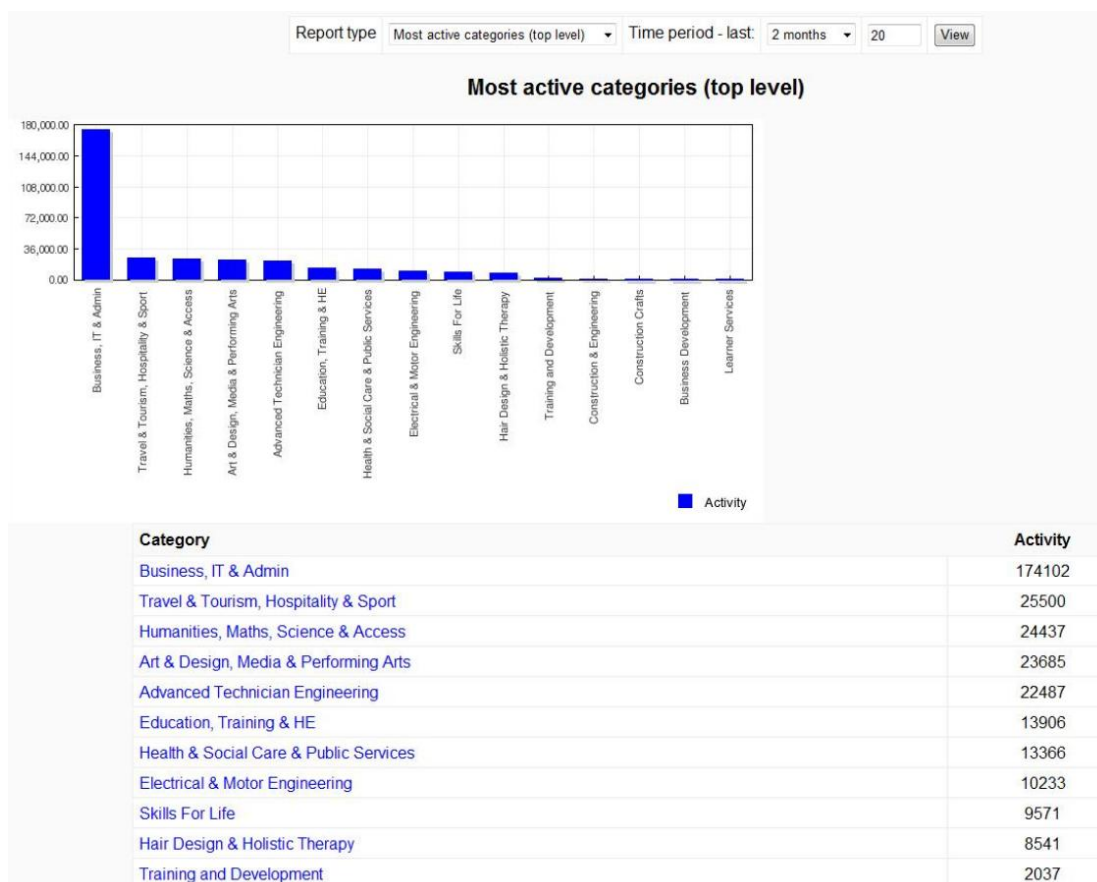
Įskiepis *Course size report* (Pav. 6) parodo sąrašą visų kursų ir kiek kiekvienas kursas užima atminties iš *moodle data* aplankalo. Jame galima pasirinkti kokiais dydžiais rodyti užimamą vietą ir/ar rodyti prie kursų, kokiai kategorijai jie priklauso.

Display size in  Display parent category name

Course ID	Course Name	Disk Size
1	Moodle (Site)	0 bytes
2	Informatikos inžinerija (Informatika)	44KB
3	Informatikos pagrindai (Informatika)	0 bytes
4	Mikro ekonomika (Ekonomika)	0 bytes
	Total	44KB

#### 6. Course size report

Įskiepis *Category activity report* (Pav. 7) parodo aktyvumų skaičių kiekvienoje kategorijoje. Patogus įrankis tuo, kad galima pasirinkti kaip norėtume gauti ataskaitą, bei pasirinkti nuo kada norėtume gauti ataskaitą.



## 7. Category activity report

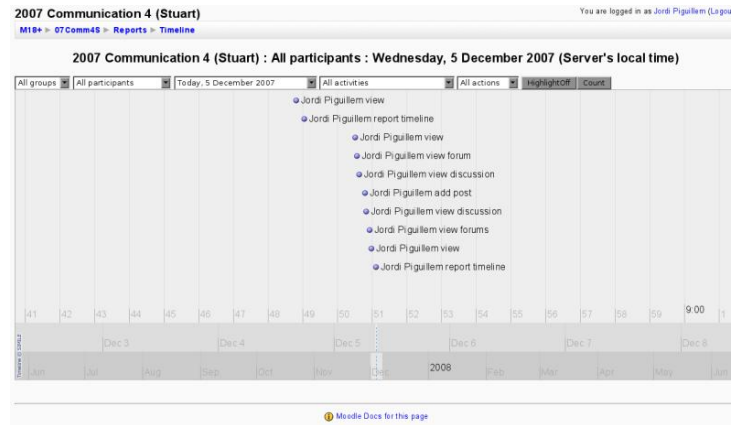
Įskiepis *Comparative reports* pateikia net penkias ataskaitas:

1. Parodo aktyvumo palyginimą kiekvieno mėnesio, lyginant su praeitų metų tuo pačiu mėnesiu;
2. Parodo pasirinktos dienos aktyvumą. Diena suskaidoma į valandas, tam kad būtų galima pamatyti aktyvius ir ramius periodus. Ši ataskaita puikiai tinka planuoti dienos resursų pasiskirstymą;
3. Parodo pasirinktinai aktyviausius vartotojus per visus kursus;
4. Parodo pasirinkto vartotojo aktyvumą per visus kursus;
5. Parodo kiek, kiekvienas kursas užima serverio atminties.

Šis įskiepis turi didžiausią funkcionalumą palyginus su kitais įskiepais iš administratoriaus ataskaitų grupės įskiepių. Lyginant su anksčiau paminėtais įskiepiams, tai šis turi visų jų savybes, išskyrus įskiepio *Module use*.

### 3.1.1.2. Nuotolinio mokymosi kursų ataskaitos

Įskiepis *Course reports plugin timeline* (Pav. 8) pateikia kurso ataskaitą laiko linijoje. Taip pat suteikia galimybę pasirinkti norimą kursą, datą ar norimus vartotojus, bei nustatyti kokius veiksmus ir kokias veiklas pavaizduoti laiko linijoje.



#### 8. Course reports plugin timeline

Įskiepis *Raw record count* (Pav. 9) pateikia vartotojus bei jų veiksmų skaičių. Taip pat suteikia galimybę pateiktą lentelę parsisiųsti *TXT*, *Excel* ar *CSV* formatu.

Users	Total
 sss2 sss2	14

#### 9. Raw record count

Įskiepis *Summary Report of Log/Activities between sequence of Dates* (Pav. 10) pateikia ataskaitą pagal nustatymus. Galima nustatyti nuo kada iki kada analizuoti įrašus (angl., *Log*), failus, kokias veiklas, ir veiksmus įtraukti analizavimui bei kokį kursą, ir vartotoją analizuoti. Taip gautą ataskaitą leidžia parsisiųsti *TXT*, *Excel* ar *ODS* failo formatu.

Action	Count	Information
Admin User	4	4 March 2009, 10:30 PM to 8 June 2009, 12:00 AM
course report log	323	Reporting Course copy 1
course user report	2	Admin User
blog view	37	view blog entry
joel henson	5	5 March 2009, 10:00 PM to 8 June 2009, 12:00 AM
quiz continue attempt	1	Pre-CSW course test
quiz report	1	CSW Post Course Quiz
course report log	2	Reporting Course copy 1
course report participation	2	Reporting Course copy 1
course update	2	
course view	32	Reporting Course copy 1
Jay Beavan	6	6 March 2009, 09:30 PM to 8 June 2009, 12:00 AM
quiz report	5	CSW Post Course Quiz
course report log	118	Reporting Course copy 1
course report outline	6	Reporting Course copy 1
course report participation	34	Reporting Course copy 1
course report stats	7	Reporting Course copy 1
course update	1	

#### 10. Summary Report of Log/Activities between sequence of Dates



### 3.1.1.3. Pažymių ir testų ataskaitos

Įskiepis *Multi Course Grader Report* (Pav. 11) parodo visų kursų (prie kurių yra suteiktos teisės prieiti) vartotojų *moodle* sistemoje pažymius. Taip pat yra galimybė pamatyti pasirinkto kurso vartotojų pažymius bei galimybė pažymių lentelę parsisiųsti *Excel*, *XML*, *ODS* ir *TXT* failo formatu.

Annotations in the screenshot:

- Link to individual Grader Report
- Icon links to Excel, XML, ODS & TXT Export pages

11. Multi Course Grader Report

Įskiepis *MCQ breakdown* (Pav. 12) pateikia informaciją, apie tai kiek vartotojų dalyvavo teste, kiek nedalyvavo teste, koks aukščiausias pažymys, koks vidurkis ir mediana. Taip pat pateikia išsamią kiekvieno vartotojo informaciją, apie teste atsakytus klausimus. Teisingai, iš dalies teisingai ar neteisingai atsakytus klausimus išskiria skirtingomis spalvomis.

Quiz statistics:

- Users completing quiz: 2 (20%)
- Users yet to complete quiz: 8 (80%)
- Highest grade achieved (max possible): 8 (80%)
- Median grade: 6 (60%)
- Mean grade: 5 (50%)

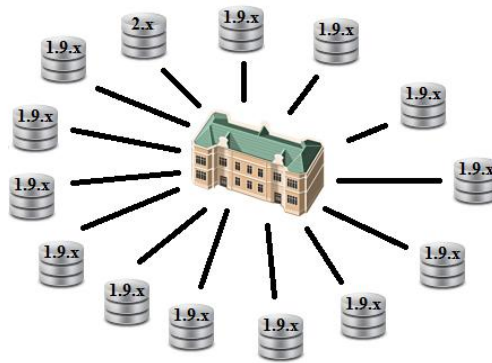
Legend: Correct option (Green), Partially correct option (Yellow), Unanswered (Grey)

Question attempted (%)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8
7 (80%)	5	2	0	0	0	0	0	0
5 (62%)	5	5	2	0	0	0	0	0
5 (62%)	5	1	0	0	0	0	0	0
3 (38%)	2	4	5	4	5	3	1	4
3 (38%)	5	1	0	4	0	0	4	1
2 (25%)	5	2	5	2	5	4	5	4
2 (25%)	4	6	6	6	6	6	6	3

12. MCQ breakdown

### 3.1.2. Institucijos padalinio VMA Moodle sistemų duomenų bazių analizė

Analizuojamos institucijos padalinio Moodle sistemos, kurių iš viso yra naudojama keturiolika. Daugiausia šių sistemų yra 1.9.x versijų ir viena 2.x versijos (Pav. 13). Visas šias sistemas prižiūri tik vienas institucijos padalinys.



13. Institucijos padalinio Moodle sistemos ir jų versijos

Moodle sistemų vartotojai suskirstyti į tris grupes: dėstytojai, studentai ir administratoriai. Kiekvienos grupės vartotojas turi institucijos vidinį kodą, kuris yra vienodas visose Moodle sistemose.

Studentų grupės vartotojų paskyros yra kuriamos, studijų pradžioje ir laikomos tris metus po universiteto studijų baigimo, pavyzdžiui, studentų grupės vartotojas baigęs bakalauro studijas ir įstojęs į magistrantūros studijas gali naudotis ta pačia studento paskyra, su tuo mačiu vidiniu kodu. Dėstytojų grupės vartotojo paskyra yra sukuriama, kai jie įsidarbina, ir laikomos tol, kol dėstytojų grupės vartotojas išsina iš darbo. Vienu ir kitu atvejų vartotojų sąrašai, kuriems reikia sukurti paskyras yra gaunami iš kito Institucijos padalinio.

Studentų grupės vartotojų fakultetai yra nurodami skaičiais, kurie atitinkamai yra: 01 — EF, 02 — SGNFS, 03 — MIF, 04 — MF, 06 — SMF, 07 — GMF, 13 — TF, 50 — TSI. Fakulteto numeris ir du papildomi skaičiai nurodo fakulteto doktorantus.

Savarankiškai prisiregistruoti prie sistemų nėra leidžiama. Ši tvarka atsirado gana neseniai, todėl visose Moodle sistemose dar galima rasti vartotojų, kurių paskyros buvo sukurtos kitais būdais (elektroniniu paštu, be patvirtinimo).

Visiems aktyviems vartotojams yra priskirtas vaidmuo (angl. *role*). Nors jų nustatymai visose duomenų bazėse yra vienodi, t.y. jų identifikavimai su vaidmens pavadinimu sutampa, tačiau gali būti, kad ateityje gali ir nesutapti, į tai reikės atsižvelgti kuriant programinę įrangą sisteminančia VMA Moodle sistemų duomenų bazių duomenis. Pagal vaidmenį bus galima filtruoti kokius vartotojus, kokiais vartotojų grupei priklauso.

Administratorių grupės vartotojai naudoja savo paskyras visose Moodle sistemose, tik administravimo tikslams. Šie vartotojai norėdami kurti dalykus, turi susikurti dėstytojų grupės vartotojus.

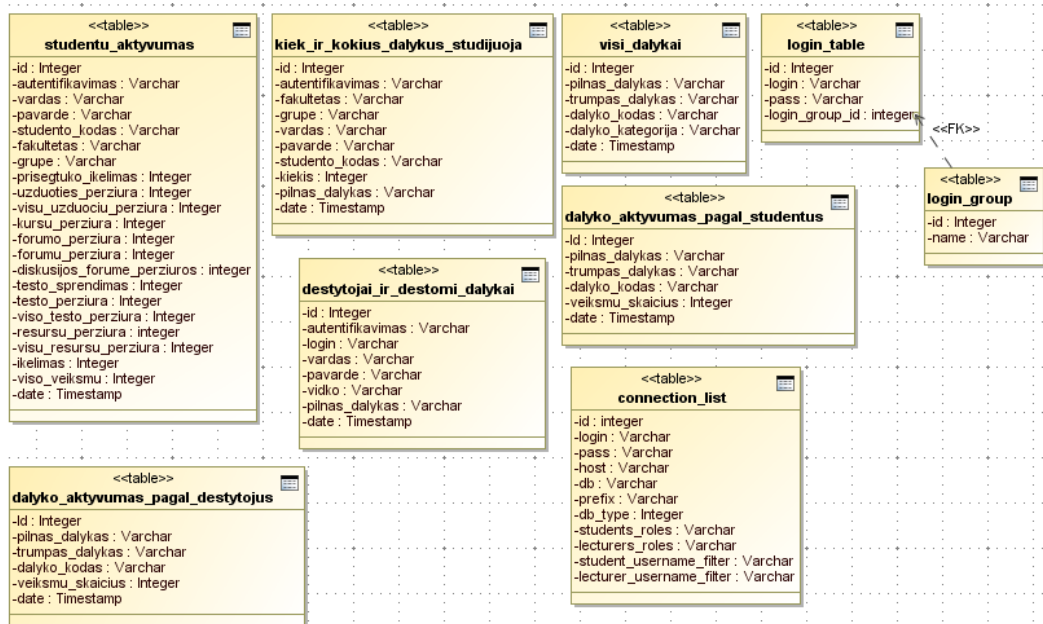
### 3.1.3. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga

Kiek ir kokius studijų dalykus studijuoja kiekvienas studentas  
Visi dalykai  
Dėstytojai ir jų dėstomi dalykai  
Dalykų aktyvumas pagal studentus  
Dalykų aktyvumas pagal dėstytojus  
Studentai ir jų aktyvumas  
Vartotojų prisijungimų kiekiai pagal pasirinktą intervalą  
Vartotojų veiksmų kiekiai pagal pasirinktą intervalą

Duomenų bazių sąrašo redagavimas  
vartotojų redagavimas  
atsijungti

## 14. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga

Sukurta VMA Moodle sistemų duomenų bazių, susisteminanti programinė įranga (Pav. 14) (Priedas nr. 3). Šios programinės įrangos duomenų bazė pavaizduota diagramoje (Diagrama 1). Šios programinės įrangos testavimas buvo atliktas su vienu iš institucijos padalinio Moodle sistemų administratoriumi.



1. Duomenų bazės diagrama

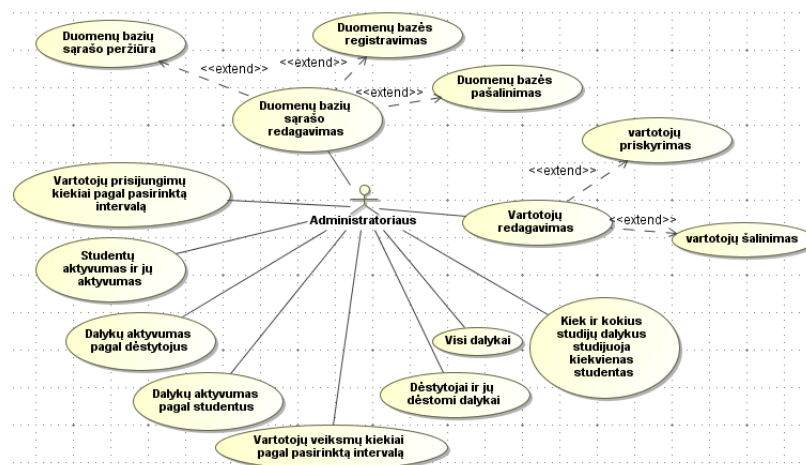
Ši programinė įranga turi vartotojų autentifikavimą — galimybę suteikti prisijungimą prie sistemos keliems vartotojams. Šiuo metu vartotojai gali būti dviejų grupių:

- Administratorius — suteiktos visos teisės naudotis programine įranga;
- Peržiūros teisės turintys — šiai vartotojų grupei suteiktos teisės peržiūrėti tik ataskaitas ir naudotis šių ataskaitų visomis funkcijomis, apribotos teisės redaguoti duomenų bazių nustatymus, iš kurių yra imami duomenys bei vartotojų redagavimas.

Taip pat šioje programinėje įrangoje numatyta galimybė redaguoti *Moodle* duomenų bazių sąrašą. Kiekvienai *Moodle* duomenų bazei galima nustatyti šiuos nustatymus:

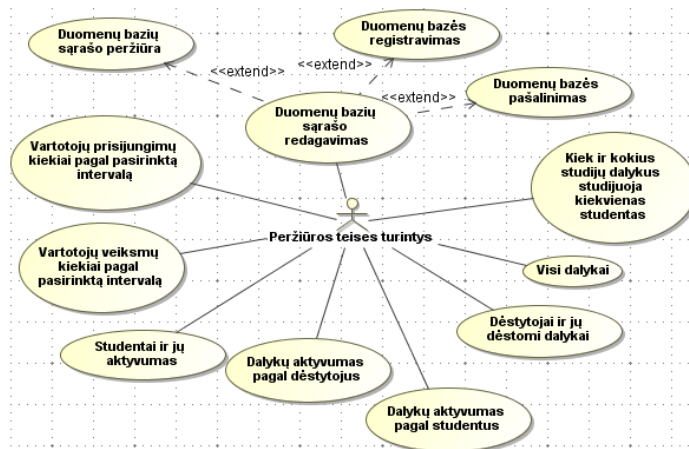
- Prisijungimo vardas;
- Slaptažodis;
- Ip adresas;
- Duomenų bazės pavadinimas;
- Lentelių prefiksas;
- *Moodle* duomenų bazės versija;
- Studentų rolių identifikatorius;
- Dėstytojų rolių identifikatorius;
- Studentų filtravimą pagal prisijungimo vardą;
- Dėstytojų filtravimą pagal prisijungimo vardą.

Šios programinės įrangos administratoriaus teises turintys vartotojų galimybes nusako poreikių diagrama (Diagrama 2).



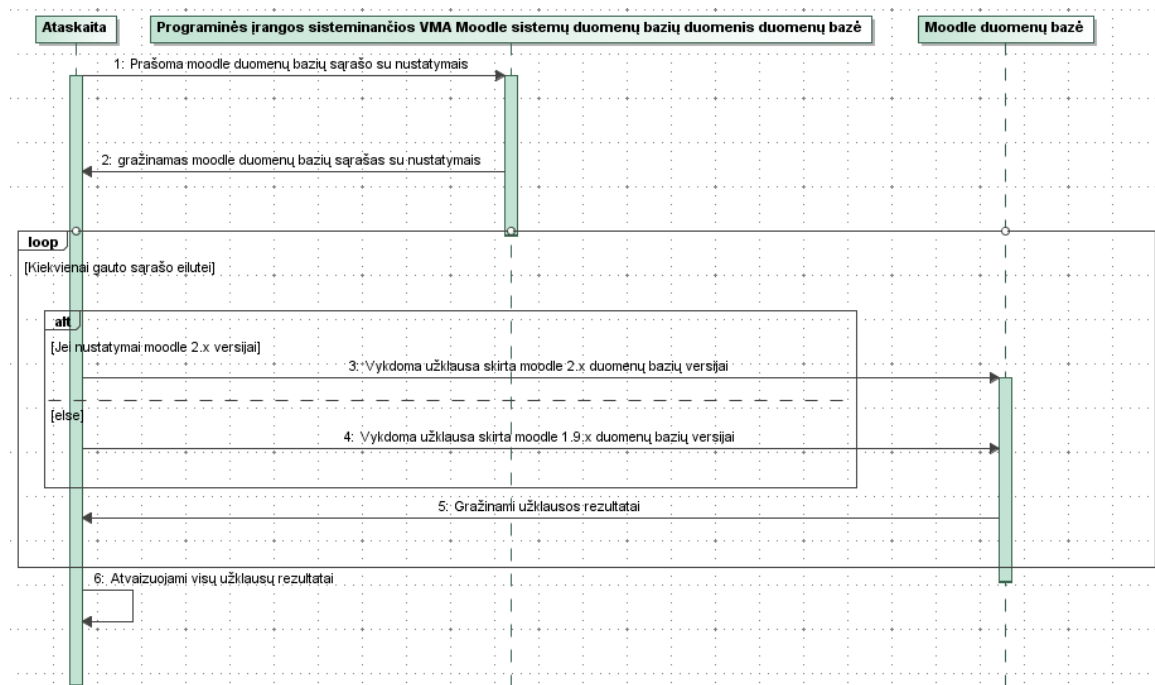
2. Administratoriaus teises turinčių vartotojų poreikių diagrama

Šios programinės įrangos peržiūros teises turinčių vartotojų galimybes nusako panaudos diagrama (Diagrama 3).



### 3. Peržiūros teises turinčių vartotojų panaudos diagrama

Visose ataskaitose duomenų surinkimas, nepriklausomai nuo užklauso yra vienodas, tai pavaizduota sekų diagramoje (Diagrama 4).



### 4. Duomenų surinkimo sekos diagrama

Taip pat visas, šios programinės įrangos ataskaitas galima išsaugoti duomenų bazėje, o vėliau prireikus jas užkrauti. Taip pat yra galimybė visas ataskaitas ir statistinius duomenų analizės rezultatus išsaugoti CSV failo formatu.

Norint geriau įvertinti statistinės analizės rezultatus buvo prijungta diagramų braižymo biblioteka, suteikianti galimybę braižyti trijų rūšių diagramas:

1. Stulpelinė diagrama (duomenų vaizdavimas skaičiais);
2. Stulpelinė diagrama (duomenų vaizdavimas procentais);

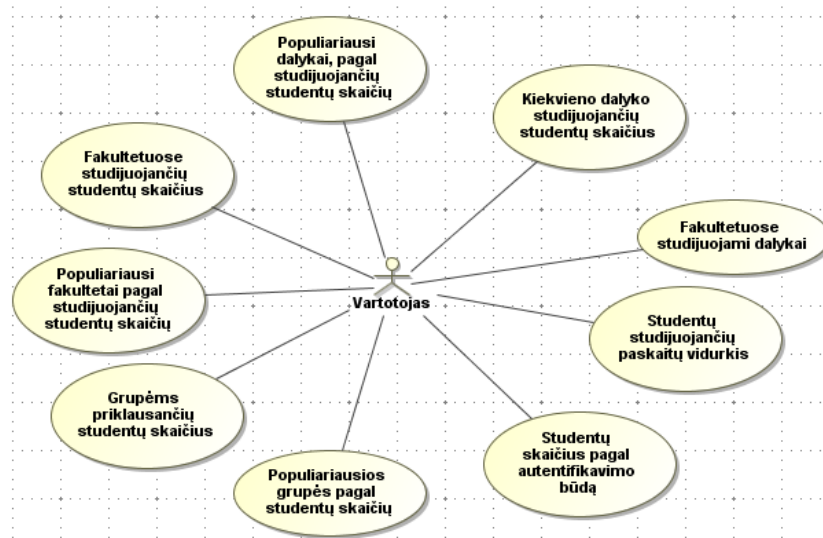
### 3. Skritulinė diagrama.

VMA Moodle sistemų duomenų bazių sisteminančios programinės įrangos ataskaitos yra:

#### 1. Kiek ir kokius studijų dalykus studijuoja kiekvienas studentas — šioje ataskaitoje atvaizduojami duomenys yra:

- Autentifikavimo būdas;
- Fakultetas;
- Grupė;
- Studento vardas;
- Studento pavardė;
- Studento vidinis kodas;
- Studijuojamų kursų skaičius;
- Studijuojami dalykai.

Šios ataskaitos statistinės analizės galimybės pavaizduotos panaudos diagramoje (Diagrama 5).



5. Statistinės analizės galimybių panaudos diagrama

#### 2. Visi dalykai — šioje ataskaitoje atvaizduojami duomenys yra:

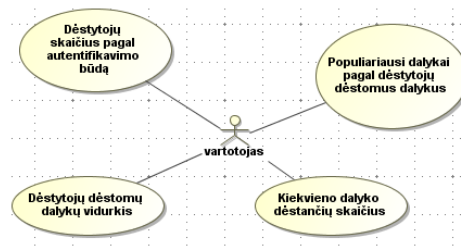
- Dalyko pavadinimas;
- Dalyko sutrumpinimas;
- Dalyko kodas;
- Dalyko kategorija.

#### 3. Dėstytojai ir jų dėstomi dalykai — šioje ataskaitoje atvaizduojami duomenys yra:

- Autentifikavimo būdas;

- Prisijungimo vardas;
- Dėstytojo vardas;
- Dėstytojo pavardė;
- Vidinis kodas;
- Dalyko pavadinimas.

Šios ataskaitos statistinės analizės galimybės pavaizduotos panaudos diagramoje (Diagrama 6).



#### 6. Statistinės analizės galimybių panaudos diagrama

4. Dalykų aktyvumas pagal studentus — šioje ataskaitoje pavaizduoti duomenys yra:

- Dalyko pavadinimas;
- Dalyko sutrumpinimas;
- Dalyko kodas;
- Veiksmų skaičius.

Šioje ataskaitoje numatytas filtras, filtruoti duomenys pagal laiką.

5. Dalykų aktyvumas pagal dėstytojus — šioje ataskaitoje pavaizduoti duomenys yra:

- Dalyko pavadinimas;
- Dalyko sutrumpinimas;
- Dalyko kodas;
- Veiksmų skaičius.

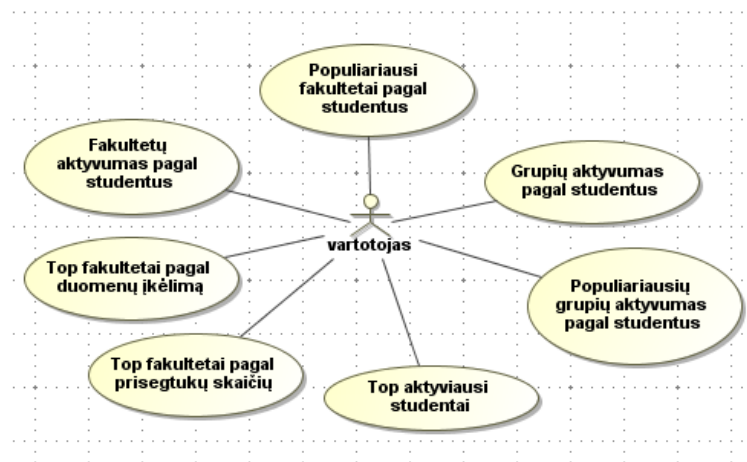
Šioje ataskaitoje, kaip ir „Dalykų aktyvumas pagal studentus“ ataskaitoje numatytas filtras, filtruoti duomenys pagal laiką.

6. Studentai ir jų aktyvumas — šioje ataskaitoje duomenys buvo pasirinkti pavaizduoti tie, kurie atspindėtų bendrą studentų aktyvumą mokymosi tikslams, šie duomenys yra:

- Autentifikavimo būdas;
- Studento vardas;
- Studento pavardė;
- Studento vidinis kodas;

- Fakultetas;
- Grupė;
- Prisegtuko įkėlimo skaičius;
- Užduoties peržiūros skaičius;
- Visų užduočių peržiūros skaičius;
- Kursų peržiūros skaičius;
- Forumo peržiūros skaičius;
- Forumų peržiūros skaičius;
- Diskusijų forume peržiūros skaičius;
- Testo sprendimų skaičius;
- Testo sprendimų peržiūros skaičius;
- Viso testo peržiūros skaičius;
- Resursų peržiūros skaičius;
- Visų resursų peržiūros skaičius;
- Įkėlimo skaičius;
- Viso veiksmų skaičius.

Šioje ataskaitoje taip pat numatytas filtras, skirtas duomenų filtravimui, pagal laiką. Šios ataskaitos statistinės analizės galimybės pavaizduotos vartotojų diagramoje (Diagrama 7).



7. Statistinės analizės galimybių panaudos diagrama

7. Vartotojų prisijungimo kiekiai pagal pasirinktą intervalą — šiai ataskaitai formuoti, visi duomenis yra imami iš įrašų lentelės (angl. *logs*). Ataskaitoje numatytas filtras, kurio



pagal pasirinktus kriterijus duomenys galima atvaizduoti grafike arba juos išsaugoti CSV failo formatu.

8. Vartotojų aktyvumo kiekiai pagal pasirinktą intervalą — šiai ataskaitai formuoti, kaip ir vartotojų prisijungimų kiekiai, pagal pasirinktus intervalus duomenys yra imami iš įrašų lentelės (angl. *logs*). Aktyvumo kiekius nusako visi veiksmai, kuriuos atliko vartotojai. Ataskaitoje taip pat numatytas filtras, kurio pagalba duomenis pagal pasirinktus kriterijus galima atvaizduoti grafike arba juos išsaugoti CSV failo formatu.

#### **3.1.4. Gautos duomenų bazės, duomenų gavybą programa STATISTICA 7**

Duomenų gavybai duomenys buvo imami iš visų keturiolikos Institucijos padalinio VMA Moodle sistemų duomenų bazių. Duomenų gavimui pasinaudota VMA Moodle sistemų duomenų bazių, susisteminanti programine įranga. Jos išsaugoti duomenys CSV failo formatu buvo eksportuojami į STATISTICA 7 programinę įrangą, duomenų gavybos taikymui.

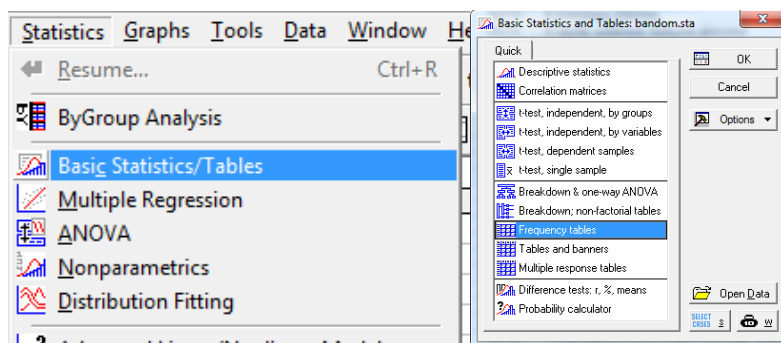
Duomenų analizavimui duomenys iš VMA Moodle sistemų duomenų bazių susisteminančios programinės įrangos paimti iš kelių ataskaitų:

1. Kiek ir kokius studijų dalykus studijuoja kiekvienas studentas;
2. Studentai ir jų aktyvumas;
3. Vartotojų prisijungimų kiekiai pagal pasirinktą intervalą;
4. Vartotojų aktyvumo kiekiai pagal pasirinktą intervalą.

##### **3.1.4.1. Kiek ir kokius studijų dalykus studijuoja kiekvienas studentas**

Analizuojant duomenys apie studentų grupės vartotojus ir jų studijuojamų dalykų skaičių buvo panaudotas *Association rules* metodas ieškant asociacijų tarp fakultetų ir studijuojamų dalykų skaičiaus, fakultetų ir studentų grupės vartotojų vardų, fakultetų ir studentų grupės vartotojų pavardžių. Taip pat buvo bandoma rasti asociacijų, tarp studento grupės vartotojo vardų ir studijuojamų dalykų skaičiaus, bei studentų grupės vartotojo pavardžių ir studijuojamų dalykų skaičiaus. Tačiau jokių asociacijų tendencijų nepavyko rasti. Taip pat nepavyko rasti, kokių nors tendencijų analizuojant ir kitais būdais (Boosted Tree Classifiers and Regression, General Classification and Regression Tree ir kt.).

Nuspręsta duomenis analizuoti, panaudojant statistinius duomenų gavybos metodus. Pasirinkus dažnių lentelės metodą (Pav. 14), atliktą studentų studijuojančių fakultetuose

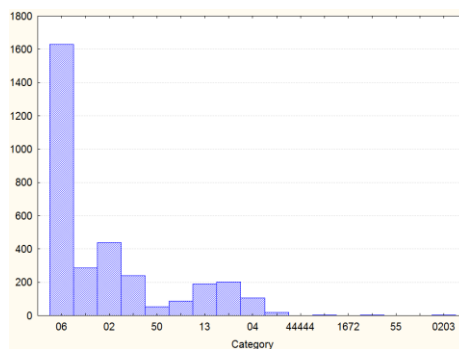


#### 14. Dažnių lentelės metodo kelias

dažnių analizė (Pav. 15). Pastebėta, kad fakultetas „06“ turi daugiausia studentų grupės vartotojų, dažnių lentelėje pastebima, kad studijuojančių studentų grupės vartotojų šiame fakultete sudaro, net ~49,9 % visų studentų grupės vartotojų, studijuojančiu nuotoliniu būdu, tai puikiai atsispindi stulpelinėje diagramoje (Pav. 16)

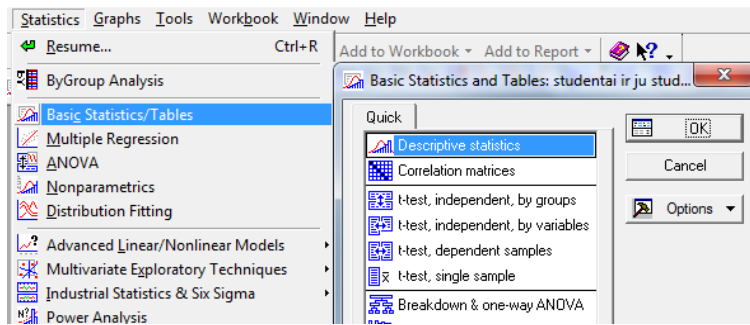
Category	Count	Cumulative Count	Percent	Cumulative Percent
06	1629	1629	49,90809	49,9081
01	287	1916	8,79289	58,7010
02	439	2355	13,44975	72,1507
05	241	2596	7,38358	79,5343
50	52	2648	1,59314	81,1275
07	87	2735	2,66544	83,7929
13	190	2925	5,82108	89,6140
03	203	3128	6,21936	95,8333
04	107	3235	3,27819	99,1115
60	19	3254	0,58211	99,6936
44444	1	3255	0,03064	99,7243
NENURODYTA	2	3257	0,06127	99,7855
1672	1	3258	0,03064	99,8162
0108	2	3260	0,06127	99,8775
55	1	3261	0,03064	99,9081
0182	1	3262	0,03064	99,9387
0203	2	3264	0,06127	100,0000
Missing	0	3264	0,00000	100,0000

15. Fakultetų dažnių lentelė



16. Fakultetų dažnių stulpelinė diagrama

Norint analizuoti studentų grupės vartotojų studijuojančių, dalykų kiekius buvo panaudotas aprašomosios statistikos metodas (Pav. 17), jo pagalba buvo gauta studentų grupės vartotojų studijuojamų dalykų kiekio dažnių lentelė (Pav. 18), kurioje pastebima, kad daugiausia studentų grupės vartotojai, studijuoja iki dvejų dalykų, net ~74,57 %, tai puikiai atsispindi

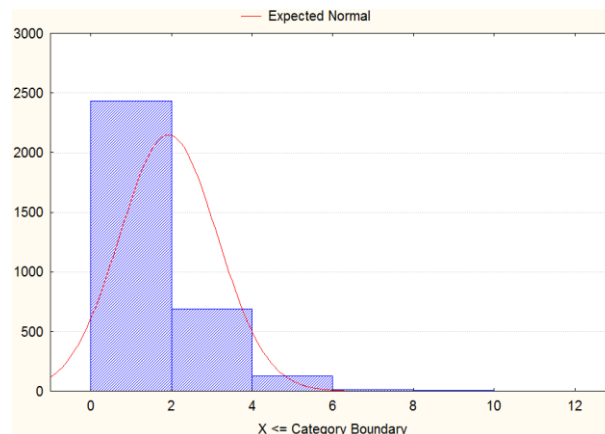


### 17. Aprašomosios statistikos metodo kelias

studentų vartotojų grupės, studijuojamų dalykų kiekio dažnių lentelės stulpelinėje diagramoje (Pav. 19).

Category	Count	Cumulative Count	Percent of Valid	Cumul % of Valid	% of all Cases	Cumulative % of All
0,000000<x<=2,000000	2434	2434	74,57108	74,5711	74,57108	74,5711
2,000000<x<=4,000000	688	3122	21,07843	95,6495	21,07843	95,6495
4,000000<x<=6,000000	127	3249	3,89093	99,5404	3,89093	99,5404
6,000000<x<=8,000000	10	3259	0,30637	99,8468	0,30637	99,8468
8,000000<x<=10,000000	4	3263	0,12255	99,9694	0,12255	99,9694
10,000000<x<=12,000000	1	3264	0,03064	100,0000	0,03064	100,0000
Missing	0	3264	0,00000		0,00000	100,0000

### 18. Studentų grupės vartotojų studijuojamų dalykų kiekio dažnių lentelė



### 19. Studentų grupės vartotojų studijuojamų dalykų kiekio dažnių lentelės stulpelinė diagrama

Aprašomosios analizės būdu gauti rezultatai (Pav. 20) parodo, kad studijuojamų dalykų vidurkis yra ~1,93, taip pat parodo, kad daugiausia studentų grupės vartotojų studijuoja po vieną dalyką, net 1614 iš 3264 analizuojamų studentų grupės vartotojų, o daugiausia studentų grupės vartotojai studijuoja po 11 dalykų. Bendras studijuojamų, studentų grupės vartotojų skaičius yra 3264, o bendras studijuojamų dalykų kiekis yra 6293.

Variable	Valid N	Mean	Median	Mode	Frequency of Mode	Sum	Minimum	Maximum
Studijuojamų dalykų kiekis	3264	1,928002	2,000000	1,000000	1614	6293,000	1,000000	11,00000

### 20. Aprašomoji analizė studentų grupės vartotojų studijuojamų dalykų kiekio

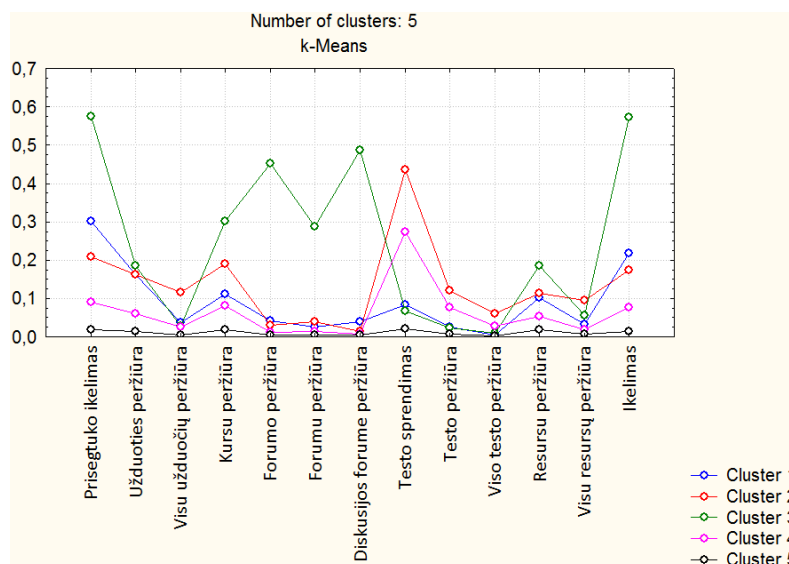
### 3.1.4.2. Studentai ir jų aktyvumas

Analizuojant šių metų duomenys, apie studentų grupės vartotojus ir jų aktyvumą, buvo naudojamas klasterizavimo k-vidurkių metodas. Analizuojant duomenis, pagal 5 klasterius, pastebėta, kad daugiausia studentų grupės vartotojų pateko į penktą klasterį, kuris pagal pateiktos lentelės (21 pav.) aktyvumo kategorijų vidurkius, lyginant su kitų klasterių aktyvumo kategorijų

Number of clusters: 5 Total number of training cases: 3255															
Cluster	Prisegtuko i kelimas	Užduoties peržiūra	Visu užduočių peržiūra	Kursu peržiūra	Forumo peržiūra	Forumu peržiūra	Diskusijos forume peržiūra	Testo sprendimas	Testo peržiūra	Viso testo peržiūra	Resursu peržiūra	Visu resursų peržiūra	Ikelimas	Number of cases	Percentage(%)
1	13,91518	91,7321	9,65179	255,9643	30,8482	3,33482	31,6786	5,35714	18,51339	1,69196	207,7054	3,94643	14,34821	224	6,88172
2	9,63636	91,4949	29,19697	436,9091	22,1667	5,19192	11,5152	27,91414	87,78283	17,30808	230,5202	11,61111	11,46970	198	6,08295
3	26,40741	104,7778	7,29630	691,5185	341,0000	37,59259	387,7037	4,29630	17,51852	2,59259	375,2222	6,92593	37,81481	27	0,82949
4	4,20659	34,3275	6,27253	186,4110	9,1648	1,78901	5,1275	17,55604	55,36044	7,87692	110,3714	2,45055	5,05055	455	13,97849
5	0,86984	7,3696	1,48660	43,4883	4,0927	0,72905	3,5151	1,36963	4,87325	0,76946	37,8392	0,89621	0,96725	2351	72,22734

#### 21. Klasterių vartotojų aktyvumo kategorijų vidurkiai

vidurkiais visais aspektais yra žemiausi, todėl galima teigti, kad studentų grupės vartotojai yra pasyvus, jie sudaro ~72% visų studijuojančių studentų grupės vartotojų, tai labai gerai matosi klasterių pasiskirstymo grafike (22 pav.)



#### 22. Klasterių pasiskirstymo grafikas

Analizuojant MARSplines [28] metodu studentų grupės vartotojų aktyvumo priklausomybę nuo studentų grupės vartotojų aktyvumo kategorijų apibendrinimo lange (Pav. 23). Išanalizuoti 13 nepriklausomų kintamųjų, analizes metu buvo naudojama 18 bazinių

Model specifications	Value
<b>Independents</b>	<b>13</b>
Dependents	1
Number of terms	19
Number of basis functions	18
Order of interactions	1
Penalty	2,000000
Threshold	0,000500
GCV error	50,66974
Prune	Yes

#### 23. MARSplines apibendrinimo langas

funkcijų. Bazinės funkcijos yra sudaromos remiantis tik eksperimentiniais duomenimis, bet ne ryšiais tarp jų. Modelio apmokymas šiam tyrimui realizuotas per 18 atvejų, kurių kiekvieno analizės metu nustatomi reikšminiai koeficientai iš tų kintamųjų grupių, kurios turi didžiausią įtaką prognozuojamam kintamajam. Paveiksle pateikti modelio apmokymo rezultatai (Pav. 24).

Coefficients, knots and basis functions	Model coefficients (Studentai ir jų aktyvumai sta)													
	NOTE: Highlighted cells indicate basis functions of type max(0, independent-knot), otherwise max(0, knot-independent)													
	Coefficients	Knots	Knots	Knots	Knots	Knots	Knots	Knots	Knots	Knots	Knots	Knots	Knots	Knots
	Viso veiksmu	Prisegtuko ikelimas	Užduoties peržiūra	Visu užduočių peržiūra	Kursu peržiūra	Forumu peržiūra	Forumu peržiūra	Diskusijos forume peržiūra	Testo sprendimas	Testo peržiūra	Viso testo peržiūra	Resursu peržiūra	Visu resursų peržiūra	Knots ikelimas
Intercept	1216,950													
Term.1	1,012				164,00									
Term.2	-1,011				164,00									
Term.3	0,996												440,00	
Term.4	-0,997												440,00	
Term.5	1,158						33,00							
Term.6	-0,998						33,00							
Term.7	1,016													
Term.8	-1,372									52,00				
Term.9	1,079		76,00							52,00				
Term.10	-1,270		76,00											
Term.11	0,977				69,00									
Term.12	-0,996				69,00									
Term.13	0,960							270,00						
Term.14	-1,074							270,00						
Term.15	0,986													
Term.16	-1,203										10,00			
Term.17	1,077										10,00			
Term.18	-1,210													34,00

## 24. MARSplines modelio koeficientai

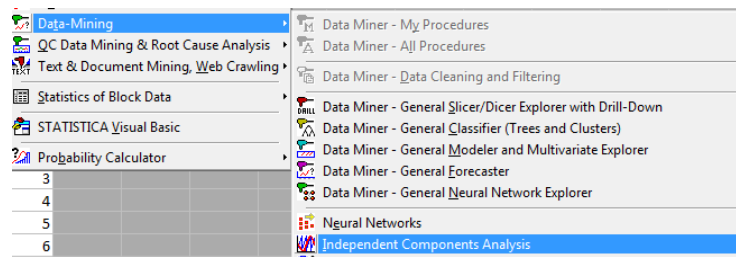
MARSplines modelis konstruoja regresines funkcijas, remiantis svertiniais koeficientais kartu su bazinių funkcijų rezultatais. Raudonai pažymėtos tos nepriklausomų kintamųjų reikšmės, kurios dalyvaus rezultatų modelio formavime. Pateiktoje lentelėje matome, kurie kintamieji turi didžiausios įtakos studentų grupės vartotojų aktyvumo kiekio prognozei bei kurias aktyvumo kategorijas galima eliminuoti (Pav. 25). Pastebima, kad visų aktyvumo kategorijų įtaka maždaug vienoda, mažiausiai įtakos turi prisegtuko įkėlimas, forumų peržiūrų, testo sprendimai ir įkėlimo aktyvumo kategorijos.

Dependents	References (to Basis Functions)
Prisegtuko ikelimas	0
Užduoties peržiūra	2
Visu užduočių peržiūra	2
Kursu peržiūra	2
Forumu peržiūra	2
Forumu peržiūra	0
Diskusijos forume peržiūra	2
Testo sprendimas	0
Testo peržiūra	2
Viso testo peržiūra	2
Resursu peržiūra	2
Visu resursų peržiūra	2
ikelimas	0

## 25. MARSplines aktyvumo kategorijų įtakų pasiskirstymas

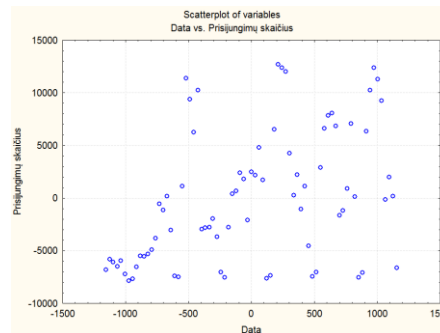
### 3.1.4.3. Vartotojų prisijungimų kiekiai pagal pasirinktą intervalą

Analizuojant nuo 2006 metų iki 2012 metų gegužės mėnesio visų Moodle duomenų bazių vartotojų prisijungimo kiekius kiekvieną mėnesį, panaudotas nepriklausomų komponentų analizės (angl. *independent component analysis*) metodas (Pav. 26).

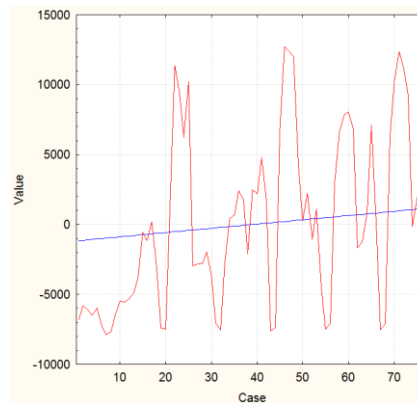


## 26. Nepriklausomų komponentų analizės metodo kelias

Vartotojų prisijungimo kiekio pagal laiko intervalą taškų sklaida pavaizduota grafike (Pav. 27). Pagal vartotojų prisijungimų grafiką (Pav. 28) pastebėta, kad vartotojų prisijungimų

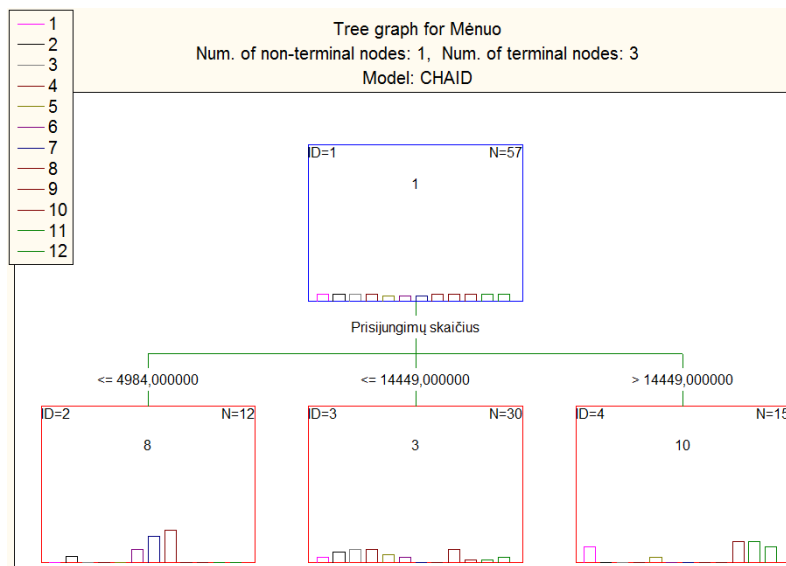


27. Vartotojų prisijungimų kiekio pagal laiko intervalą taškų sklaida kiekis nuo 2006 metų nuolat augo.

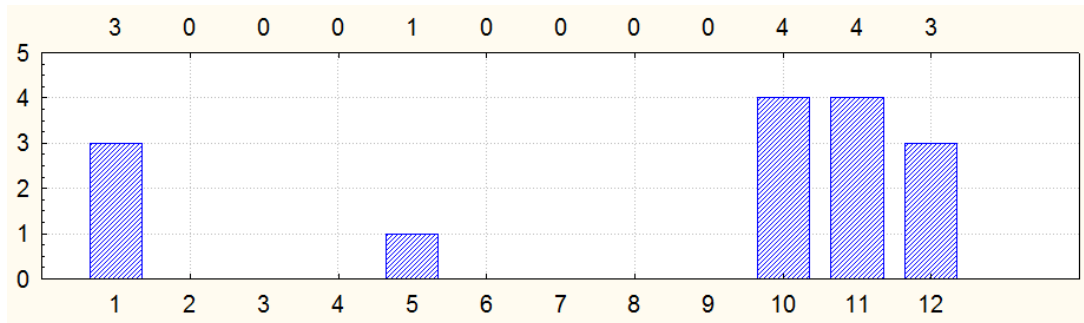


## 28. Vartotojų prisijungimų kiekio pagal laiko intervalą grafikas

Analizuojant duomenis nuo 2006 metų iki 2012 metų gegužės mėnesio, visų *Moodle* sistemų vartotojų prisijungimų kiekius. Grafike (Priedas nr. 5) pastebima, kad 2006 m. išsiskiria iš visų kitų mokslo metų, savo pasyvumu pagal prisijungimų skaičių, taip gali būti todėl, kad tai vieni iš pirmųjų metų, kai imta taikyti *Moodle* sistemas nuotoliniame mokyme. Nuspręstą analizuoti duomenys nuo 2007 metų rugpjūčio 8 d., t.y. nuo tada kai prasideda nauji mokslo metai. Duomenys klasifikuojant, pagal kiekvienų metų mėnesį *CHAID* metodu [27] (Pav. 29)



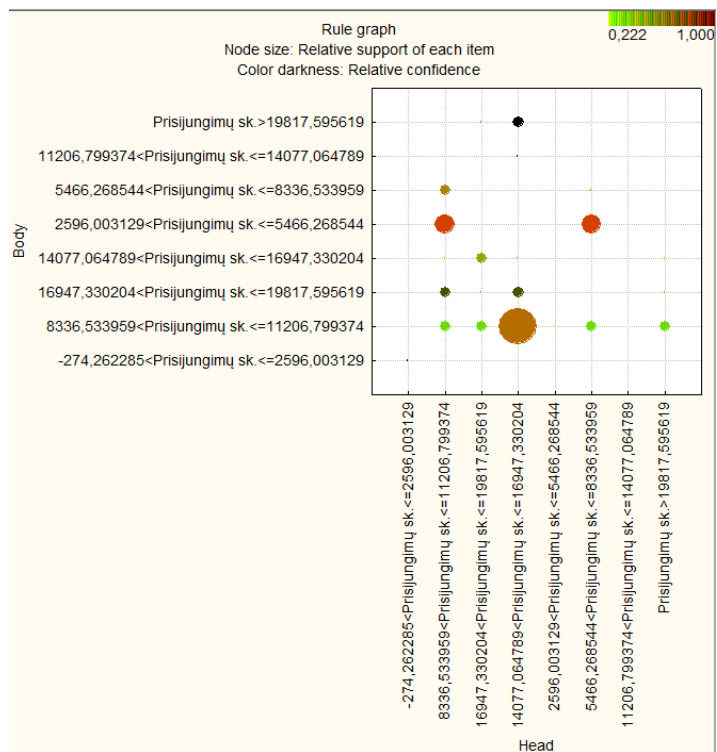
29. Vartotojų prisijungimų kiekio klasifikavimas, pagal kiekvienų metų mėnesį *CHAID* metodu pastebima, kad pagal prisijungimų skaičių duomenys galima išskaidyti į tris pagrindinės dalis. Vartotojų prisijungimo skaičių iki 4964, vartotojų prisijungimo skaičių iki 14449, ir kai prisijungimų skaičius yra virš 14449. Pagal analizuojamų prisijungimų skaičių pastebima, kad daugiausia virš 14449 prisijungimų skaičiaus sudaro prisijungimai pirmąjį mokslo metų pusmetį. Tai puikiai parodo dažnių stulpelinė diagrama (Pav. 30).



### 30. Vartotojų prisijungimų kiekio tendencijos pagal kiekvienų metų mėnesį

Analizuojant kiekvieno mėnesio nuspėjimo reikšmių duomenys (Priedas nr. 6), pastebimos tam tikros tendencijos, kad kiekvienų metų 7 ir 8 mėnesį vartotojų prisijungimų skaičiai yra stabilūs, taip pat 10 mėnesį prisijungimų skaičius apie 75% yra stabilus.

Analizuojant prisijungimų skaičius Sequence, Association, and Link Analysis metodu pastebėta įdomi tendencija (Priedas nr. 7). Staigus prisijungimų šuolis į ~14077 – 16947 režį, kuris puikiai pavaizduotas grafike (Pav. 31), toks staigus prisijungimų šuolis yra siejamas su kiekvienų mokslo metų pradžia.

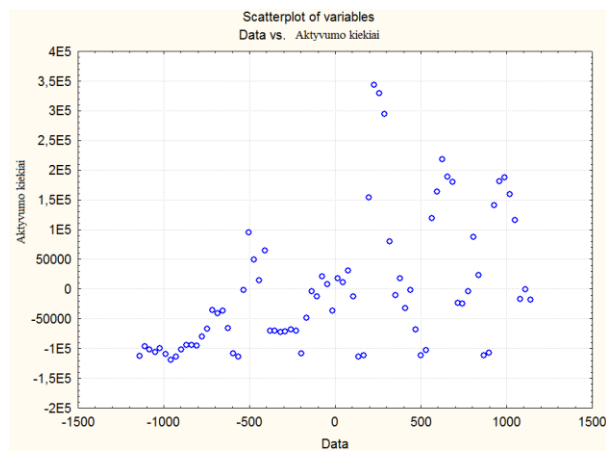


31. Vartotojų prisijungimų kiekio sekų analizė

### 3.1.4.4. Vartotojų aktyvumo kiekiai pagal pasirinktą intervalą

Analizuojant nuo 2006 metų iki 2012 metų gegužės mėnesio visų *Moodle* duomenų bazių vartotojų aktyvumo kiekius, kiekvieną mėnesį buvo naudojamas nepriklausomų komponentų analizės metodas (Pav. 26).

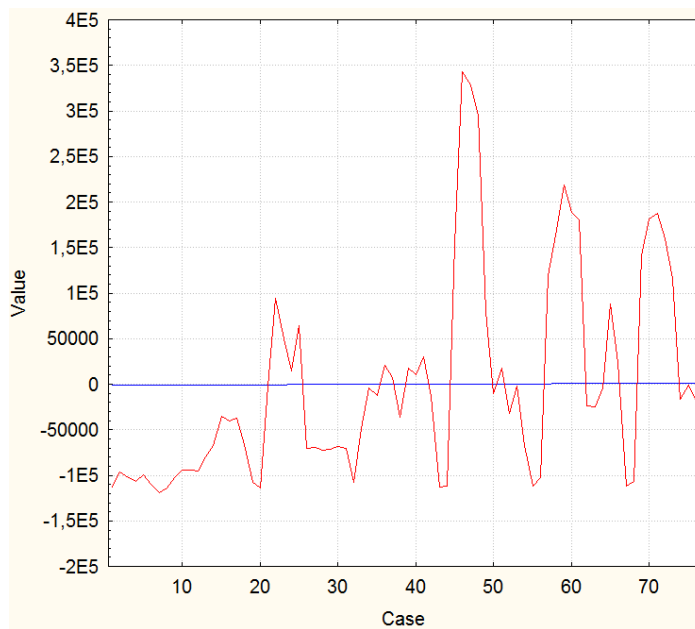
Vartotojų aktyvumo kiekio, pagal laiko intervalą taškų sklaida pavaizduota grafike (Pav. 32). Pagal vartotojų aktyvumo grafiką (Pav. 33) pastebėta, kad vartotojų aktyvumo kiekis nuo



32. Vartotojų aktyvumo kiekio pagal laiko intervalą taškų sklaida

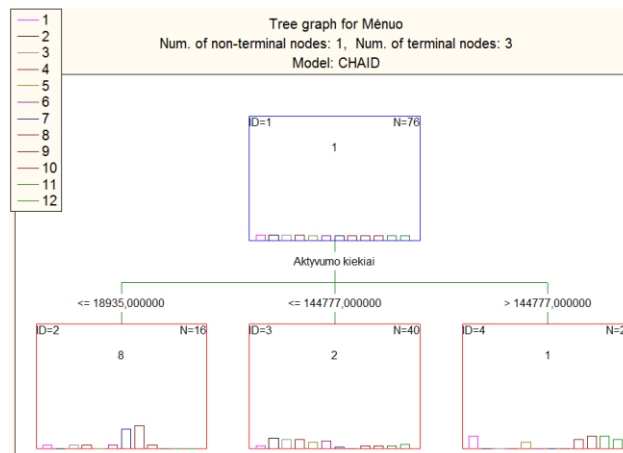
2006 metų nežymiai pakito.





### 33. Vartotojų aktyvumo kiekio priklausomybės nuo laiko grafikas

Analizuojant duomenys nuo 2006 metų iki 2012 metų gegužės mėnesio visų *Moodle* sistemų vartotojų aktyvumo kiekius. Analizuodami grafiką (Priedas nr. 8) ir atsižvelgiant į tai, kad 2006 metai, yra vieni iš pirmųjų metų, kai imta taikyti *Moodle* sistemas nuotoliniame mokyme. Nuspręstą analizuoti duomenis nuo 2007 metų rugpjūčio 8 d., t.y. nuo tada kai prasideda nauji mokslo metai. Duomenys klasifikuojant, pagal kiekvienų metų mėnesį *CHAID*

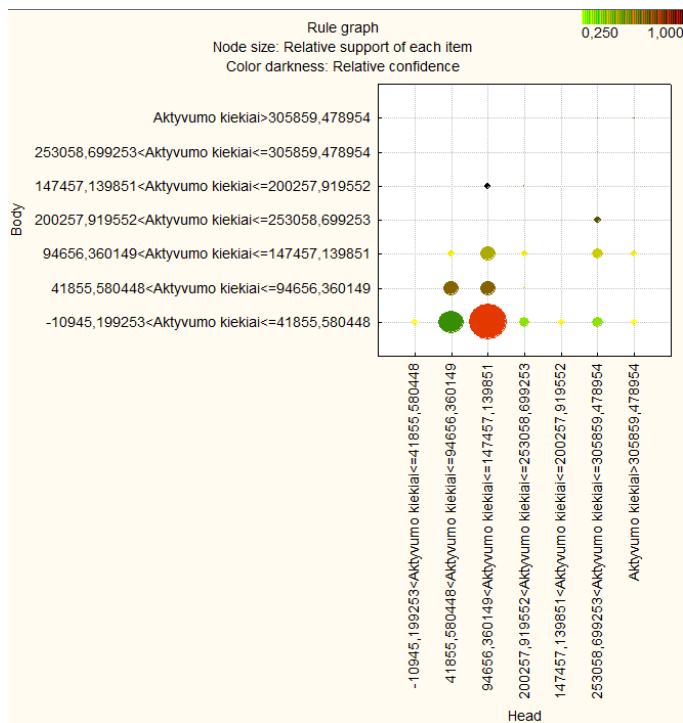


### 34. Vartotojų aktyvumo kiekio klasifikavimas *CHAID* metodu

metodu (Pav. 34) pastebima, kad pagal aktyvumo kiekį duomenys galima išskaidyti į tris pagrindinės dalis. Vartotojų aktyvumo kiekis iki 18935, vartotojų prisijungimo skaičių iki 144777, ir kai prisijungimų skaičius yra virš 144777. Pagal analizuojamų prisijungimų skaičių, pastebima, kad daugiausiai virš 144777 aktyvumo kiekio, sudaro aktyvumai pirmąjį mokslo metų pusmetį, kaip ir prisijungimų kiekio.

Analizuojant kiekvieno mėnesio nuspėjimo reikšmių duomenys (Priedas nr. 9), pastebėta, kad kiekvienų metų 2 ir 8 mėnesį vartotojų prisijungimų skaičiai yra stabilūs, taip pat 3 mėnesį aktyvumo kiekis paskutinius 3 metus yra stabilus su 57 % tikimybe.

Analizuojant prisijungimų skaičius Sequence, Association, and Link Analysis metodu pastebėta tendencija, kad aktyvumo kiekiai palengva keičiasi į ~94656 – 147457 rėžį, kuris puikiai atsispindi grafike (Pav. 35). Tai leidžia daryti prielaidą, kad vartotojų aktyvumo kiekiai yra tendencingi.



35. Vartotojų aktyvumo kiekių sekų analizė

## 3.2. Problemos ir jų sprendimo būdai

### 3.2.1 VMA Moodle sistemų duomenų bazių pildymo

1. Kai kurie studentų grupės vartotojai turi dėstytojų grupės teises, dažniausiai šios teisės yra priskiriamos studentų grupės vartotojams, kurie atlieka kokius nors tyrimus susijusius su dalyko redagavimu ar kūrimu, ši problema buvo išspręsta papildžius studentų grupės vartotojų filtravimą pagal tai, kad studento grupės vartotojų prisijungimo varde būtų du taškai, nes studentų grupės vartotojų prisijungimo vardų struktūrą yra „metai.pavardė.vardas“. Atsižvelgiant į tai, kad studentų grupės vartotojai gali turėti dėstytojų grupės teises, apsaugoti nuo atvirkštinio varianto, dėstytojų grupės vartotojų

prisijungimo vardas taip pat filtruojamas pagal tai, kad jų prisijungimo varde nebūtų dviejų taškų.

2. Pastebėta, kad visose *Moodle* duomenų bazėse dalykų pirmas įrašas yra *Moodle* sistemos domeno vardas. Todėl rašant užklausą buvo atsižvelgta ir įtraukiant visus studijuojamus dalykus iš visų duomenų bazių, nepaimti pirmo įrašo iš visų duomenų bazių.
3. Norint atlikti analizę pagal dalyko kodą, pastebėta problema, kad netaisyklingai dėstytojų grupės vartotojai pildo informaciją apie savo studijų dalykus. Pavyzdžiui, ne visi dėstytojų grupės vartotojai vienodai įrašo į dalyko informacijos kortelę dalyko kodą ir numerį, todėl dalyko analizė, pagal dalyko kodą tapo negalima.
4. Analizuojant duomenys pastebėta, kad vienu studentų grupės vartotojų vienodos grupės ar fakultetai priskirti mažosiomis raidėmis, pas kitus didžiosiomis raidėmis, ši problema išspręsta prieš, analizuojant duomenys pakeisti įrašus didžiosiomis raidėmis.
5. Analizuojant fakultetų įrašus, pastebėta, kad vieni fakultetai išreikšti skaitinę išraiška, o kiti raidinę. Ši problema išspręsta fakultetus verčiant skaitinę reikšmę.
6. Analizuojant studentų grupės vartotojų grupes, pastebėta, kad vieniems studentų grupės vartotojams prie grupės yra parašyta „nenurodyta“, kitiems tuščias laukelis, ši problema buvo išspręsta prieš analizuojant duomenys, atkreipiant dėmesį į tai, kad grupės gali būti nenurodytos, ir šias grupes priskirti į nenurodytų grupių kategoriją.
7. Analizuojant dėstytojų grupės vartotojus buvo pastebėta, kad kai kurie dėstytojų grupės vartotojai naudojami paskyromis, kuriose nėra nurodyti vidiniai kodai, dėl nebuvo negalima atlikti jų rūšiavimo. Todėl dėstytojų grupių vartotojų rūšiavimas atliktas, pagal prisijungimo vardą, tačiau susidurta su trimis problemomis:
  1. Pastebėta, kad vieno iš dėstytojo grupės vartotojo pavardė vienoje duomenų bazėse be lietuviškų raidžių, kitose su pvz. **e** ar **ė**.
  2. Pastebėta, kad vienas iš dėstytojų grupės vartotojų, vienoje *Moodle* sistemoje kuria dalykus naudojantis dėstytojo grupės vartotojo paskyra, kai kuriose sistemose administratoriaus grupės vartotojo paskyra, kurių prisijungimo vardai skiriasi.
  3. Taip pat filtruojant, pagal prisijungimo vardą, dėstytojų grupės vartotojų analizė tampa netiksli, pagal prisijungimo būdą.

Atsižvelgiant į iškilusias problemas, nuspręstą nedaryti išsamesnės dėstytojų grupės vartotojų analizės.

8. Studentų grupės vartotojų nurodytos grupės kartais nurodomos su pogrupiu. Nurodymai būna įvairūs: per brūkšnį, pasviręs brūkšnys į kairę ar į dešinę ir pan. Atsižvelgiant į tai, negalima daryti išsamesnės analizės pagal studentų grupės vartotojų grupes.

### **3.2.2 VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga**

1. JavaScript dvimačio masyvo išsaugojimas į CSV failą. Kadangi, *HTML Table Filter Generator* biblioteka parašyta JavaScript kalba, ir pasinaudojus jos funkcijomis gaunamas matomos ataskaitos dvimatis masyvas. Šį masyvą reikia išsaugoti į failą vartotojui patogiu būdu. Nors JavaScript veikia kliento pusėje, ir turi funkcijas failo rašymui bei skaitymui, tačiau susidurta su problema, kad JavaScript neturi tiesioginės galimybės tai padaryti, todėl surastas sprendimo būdas perduoti masyvą POST metodu į PHP formuojamą langą su masyvo turiniu, kurį po to vartotojui išskarto siūloma išsaugoti.
2. *HTML Table Filter Generator* bibliotekos funkcijos `getFilteredData()` netaisyklingas veikimas. Pastebėta, kad funkcija `getFilteredData()`, tam tikrais atvejais pateikia klaidingą dvimatį masyvą, kurio duomenys norima išsaugoti arba atvaizduoti diagramoje. Ši klaida jau yra paskelbta gamintojų puslapyje, tačiau klaida iki šiol neištaisyta. Šiai problemai sprendimo būdas rastas pasinaudojus kita šios bibliotekos funkcija `getColValues()`. Ši funkcija teisingai parodo stulpelio turinį, kadangi, mums yra žinomas stulpelių skaičius ataskaitose, tai galima surinkti dvimatį JavaScript masyvą, kurį po to galima išsaugoti į failą ar iš jo duomenų sukurti diagramas.
3. *JpGraph* bibliotekos didelių duomenų kiekio atvaizdavimo problema. Susidurta su problema, kad pasirinkta biblioteka, negali tinkamai atvaizduoti didelių duomenų kiekių, nors gamintojų puslapyje nurodyta, kad galima atvaizduoti labai didelius duomenų kiekius bei nėra nurodyta apribojimų. Ši problema buvo išspęsta apribojant atvaizduojamų duomenų kiekius.

## **3.3. Rekomendacijos**

### **3.3.1 VMA Moodle duomenų bazių pildymo**

1. Nustatyti dėstytojų grupės vartotojų taisykles, kaip teisingai pildyti dalykų korteles.

2. Sutvarkyti skriptus, kuriais naudojantis duomenys perduodami į *Moodle* duomenų bazę, kad nebeliktų vienodų grupių ar fakultetų, tačiau su skirtingų dydžių raidėmis. Šios problemos sprendimas padėtų išspręsti problemas dėl lietuviškų raidžių.
3. Išspręsti vartotojų duomenų atnaujinimą, kad vartotojų informacija atsinaujintų, ne tuomet, kai vartotojai prisijungia, o kai gaunama informacija apie vartotojus iš kito Institucijos padalinio.
4. Dėstytojų grupės vartotojams drausti naudotis paskyromis, kurioms nėra priskirti vidiniai kodai.
5. Sugriežtinti tvarką, dėl administratoriaus grupės vartotojų paskyrų naudojimo, ne *Moodle* sistemų administravimo darbams.
6. Išspęsti problemą, dėl pogrupių nurodymo grupės laukelyje, vienas iš problemos sprendimo būdų, galėtų būti suderinti prefiksą, pagal ką būtų nurodomas pogrupis.
7. Suvienodinti fakultetų duomenys paliekant juos arba skaitinius, arba raidinius.

### **3.3.2 VMA Moodle serveriams**

1. Administruojant VMA *Moodle* serverius, reikia ypatingą dėmesį skirti serveriams, kuriuose lankosi socialinių mokslų fakulteto studentų grupės vartotojai, kadangi jie sudaro ~50 % visų studijuojančių studentų grupės vartotojų.
2. Tobulinant ar atnaujinant serverius, reikia atkreipti dėmesį, kad studentų prisijungimų skaičius nuolat auga, todėl ateityje gali reikti galingesnių serverių, kurie atlaikytų didėjantį vartotojų prisijungimų skaičių.
3. Ypatingą dėmesį serverių apkrovai reikia skirti pirmąjį metų pusmetį, nes data *Moodle* vartotojų prisijungimų ir aktyvumų skaičiai yra didžiausi.

### **3.3.3 VMA Moodle sistemų vartotojams**

1. Dėstytojų grupės vartotojus skatinti vis daugiau kurti dalykus *Moodle* sistemose, kurie išnaudotų visas teikiamas *Moodle* sistemos galimybes, taip įtraukiant daugiau ir aktyviau studentų grupės vartotojus studijuoti dalykus nuotolinių būdų, nes dabar studentų grupės vartotojai vidutiniškai studijuoja po 2 dalykus, o jų aktyvumas beveik neauga.
2. Skatinti studentų grupės vartotojus aktyviau naudotis *Moodle* sistemomis, nes pagal klasterinės analizės rezultatus skaidant į 5 klasterius, net ~72% studentų pateko į klasterį, kuris visomis aktyvumo kategorijomis yra pasyviausias.

3. Visus vartotojus supažindinti su *Moodle* sistemomis ir jų teikiamomis galimybėmis, nes pastebėta iš aktyvumo sekų analizės, kad kiekvienų mokslo metų pradžioje vartotojų prisijungimų skaičius auga didesniu tempu, nei aktyvumas, tai leidžia daryti prielaidą, kad ypatingai mokslo metų pradžioje vartotojai dažnai prisijungia prie Moodle sistemų, tačiau atlieka mažai veiksmų, viena iš priežasčių, kad nauji vartotojai, nemoka naudotis Moodle sistemos teikiamomis galimybėmis

### **3.4. Darbo rezultatų analizė**

#### **3.4.1. Moodle sistemai statistikos ir ataskaitos įskiepių analizė**

Analizuojant *Moodle* 1.9.x versijos sistemai statistikos ir ataskaitos įskiepius, buvo nustatyta, kad statistikos ir ataskaitos įskiepių, kurie patalpinti *Moodle* oficialiame puslapyje, yra nesuderinami su naujausia *Moodle* sistema 2.x versija, pagrindinė priežastis yra tai, kad su šia naujausia *Moodle* sistemos versija pasikeitė duomenų bazės struktūra. Taip pat pastebėta, kad *Moodle* oficialiame tinklalapyje nėra sukurtų statistikos ir ataskaitos įskiepių, skirtų *Moodle* 2.x sistemai. Naujausioje *Moodle* 2.x versijoje yra tik baziniai statistikos ir ataskaitos įrankiai, kurie parodo *Moodle* vartotojų veiksmus (angl. *log*) bei įvairių testų rezultatus ir studento pažymius. Kol kas nėra nei vieno oficialioje *Moodle* svetainėje patalpinto statistikos ir ataskaitos įskiepio skirto *Moodle* 2.x versijai, yra tik galimybė įdiegti pakoreguotus senesnio tipo *Moodle* versijoms sukurtus statistikos ir ataskaitos įskiepius. Taip pat pastebėta, kad *Moodle* sistemos negali pasigirti statistikos ir ataskaitos gausa, o duomenų gavybos įskiepių išvis nėra. Nors šiais laikais vis daugiau yra akcentuojama, kad duomenų gavybos technologijų panaudojimas VMA yra labai svarbus norint sukurti modernią, šiuolaikinę, besimokančiųjų VMA, tačiau, kad ir kaip būtų gaila *Moodle* sistemos kūrėjai visai neskiria dėmesio šiai sričiai.

#### **3.4.2. Institucijos padalinio VMA Moodle sistemų duomenų bazių analizė**

Darbo metu išanalizuoti Institucijos padalinio *Moodle* sistemų duomenų bazės ir jų duomenys, bei tarpusavio sąryšiai. Remiantis šios analizės metu rezultatais, suprojektuota programinė įranga, susisteminanti VMA *Moodle* duomenų bazių duomenys. Taip pat rastą *Moodle* sistemų duomenų bazių pildymo problemų, kurių ištaisymas gali užtrukti ilgą laiką. Ši

analizė parodo, jog svarbu yra tinkamai pildyti duomenų bazės duomenimis, kurie vėliau gali būti surinkti bei paruošti analizei.

### **3.4.3. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga**

Darbo metu sukurta programinė įranga susisteminanti VMA Moodle duomenų bazių duomenys, kuri yra universali ir gali būti naudojama analizuojant visas Moodle 1.9.x ir Moodle 2.x duomenų bazių versijas. Su šia programine įranga galima peržiūrėti įvairias ataskaitas, bei jas filtruoti ir rūšiuoti, pagal pasirinktus kriterijus, taip pat atlikti nedidelę statistinę analizę, kurios rezultatus bus galima atvaizduoti pasirinktinai skritulinėje, stulpelinėje diagramoje ar atitinkamai išsaugoti CSV failo formatu.

Ši programinė įranga, dėl duomenų pildymo problemų negali paimiti visos informacijos susijusios su Moodle sistemos vartotojais.

### **3.4.4. Gautos duomenų bazės duomenų gavybą programa STATISTICA 7**

Analizuojant duomenys, gautus iš sukurtos VMA Moodle duomenų bazių sisteminančios programinės įrangos, pastebėta, kad fakultetas 06 turi didžiausią studentų grupės vartotojų skaičių, kuris atitinką socialinių mokslų fakultetą. Taip pat pastebėta, kad studentų grupės vartotojai dažniausiai studijuoja po du dalykus.

Analizuojant studentų grupės vartotojų aktyvumus pastebėta, kad bendram studentų grupės aktyvumui didžiausią įtaką turi diskusijų forume peržiūra bei resursų peržiūra.

Pastebėta tendencija, kad nuo 2006 metų iki 2012 metų gegužės mėnesio vartotojų prisijungimų skaičius pastebimai augo, o vartotojų aktyvumas nežymiai kito. Taip pat pastebėta, kad vartotojai tiek prisijungimo kiekiu, tiek aktyvumo kiekiu pirmą mokslo metų pusmetį yra aktyvesni, lyginant su antru mokslo metų pusmečiu.

Analizuojant vartotojų prisijungimų kiekius, pastebėtos tam tikros tendencijos, kad vartotojų prisijungimų kiekiai kiekvienų metų 7 ir 8 mėnesiais yra panašūs. Analizuojant vartotojų aktyvumo kiekius pastebėta, kad kiekvienų metų 2 ir 8 mėnesį jau šešis metus iš eilės aktyvumai yra panašūs. Taip pat pastebėta tendencija, kad vartotojų prisijungimų kiekiai, kiekvienų mokslo metų pradžioje auga greičiau, negu vartotojų aktyvumo kiekiai.

## Išvados

1. Išanalizavus VMA *Moodle* statistikos ir ataskaitos įskiepius, galima teigti, kad nėra daug statistikos ir ataskaitos įskiepių, o duomenų gavybos įskiepių visiškai nėra.
2. Remiantis visų institucijos padalinio VMA *Moodle* duomenų bazių analizės rezultatais, suprojektuota programinė įranga susisteminanti VMA *Moodle* duomenų bazių duomenys.
3. Sukurta programinė įranga, kuri susistemina VMA *Moodle* duomenų bazių duomenys, pagal pasirinktus kriterijus.
4. Remiantis duomenų analizės programa STATISTICA 7 rezultatais, galima konstatuoti, kad buvo gauta daug ir įvairių duomenų, kurie padeda geriau suvokti ne tik Institucijos padalinio *Moodle* sistemų vartotojų, bet ir *Moodle* sistemų serverių apkrovų tendencingumus.
5. Pateiktos rekomendacijos Institucijos padaliniui, administruojančiam VMA *Moodle* sistemas, kurios padėtų išspręsti daugelį kilusių problemų, tiek dėl duomenų bazių pildymo, tiek dėl serverių administravimo bei vartotojų aktyvumo skatinimo.



## Informacinių šaltinių ir literatūros sąrašas

1. *Nuotolinis mokymasis norintiems nuolat tobulėti*: Prieiga per internetą: <http://www.anonsas.lt/portal/categories/27/1/0/1/article/220>, Žiūrėta [ 2012 03 12].
2. *Nuotolinis mokymasis*: Prieiga per internetą: <http://www.vipt.lt/cms/app?service=external/index&sp=1944&sp=1442>, Žiūrėta [ 2012 03 12].
3. *Mokymosi visą gyvenimą memorandumas*: Prieiga per internetą: [http://www.lssa.smm.lt/docs/Memorandumas\\_2001.doc](http://www.lssa.smm.lt/docs/Memorandumas_2001.doc), Žiūrėta [ 2012 03 12].
4. *Pasirengimas vertinti neformaliai ir savaime įgytą kompetenciją*: Prieiga per internetą: [http://www.leidykla.vu.lt/fileadmin/Acta\\_Paedagogica\\_Vilnensia/20/78-89.pdf](http://www.leidykla.vu.lt/fileadmin/Acta_Paedagogica_Vilnensia/20/78-89.pdf), Žiūrėta [ 2012 03 12].
5. *Moodle*: Prieiga per internetą: [http://moodle.ktu.lt/file.php/1/Moodle\\_vadovas.pdf](http://moodle.ktu.lt/file.php/1/Moodle_vadovas.pdf), Žiūrėta [ 2012 03 12].
6. *Virtuali mokymosi aplinka mokyklai*: Prieiga per internetą: <http://www.mtp.smm.lt/dokumentai/InformacijaSvietimui/MethodinesRekomendacijos/200506VMA.doc>, Žiūrėta [ 2012 03 12].
7. *Virtuali mokymosi aplinka*: Prieiga per internetą: <http://www.vipt.lt/cms/app;jsessionid=DC0A9710F41EE047CC382F00FD737652.worker?service=external/index&sp=2313&sp=2030>, Žiūrėta [ 2012 03 12].
8. Sakalauskas Leonidas. *Duomenų gavyba*. Vilnius, 2009.
9. *Duomenų gavybos technologijų taikymas išskirstytų serverių darbui gerinti*: Prieiga per internetą: [http://isd.ktu.lt/it2009//material/Research/ITA/ITA\\_3.pdf](http://isd.ktu.lt/it2009//material/Research/ITA/ITA_3.pdf), Žiūrėta [ 2012 03 15].
10. Saulius Preidys, Leonidas Sakalauskas. *Nuotolinių studijų vartotojų poreikių analizė*. Vilnius, 2009.
11. Margaret H. Dunham. *Data Mining. Introductory and Advanced Topics*. 2006.
12. G. Dzemyda, P. Gudynas, V. Saltenis, V. Tiesis. *Lietuvos peagogai ir moksleiviai: analizė ir prognozė*. Vilnius, 2001.
13. Turskienė S., Kulvietis G., Burbaitė R. *Duomenų gavybos technologijų panaudojimas jaunųjų kompiuterininkų mokyklos veiklos analizei, Lietuvos matematikos rinkinys*. 2003, t. 43, spec.nr., p. 297-303.

14. Saulius Preidys, Leonidas Sakalauskas. *Nuotolinio mokymosi stilių personalizavimas. Informacijos mokslai*. Vilnius, 2011.
15. Jolita Bernatavičienė. *Vizualios žinių gavybos metodologija ir jos tyrimas*, Daktaro disertacija, Vilnius, 2008.
16. Jelena Mačenko. *Duomenų gavybos technologijų taikymas išskirstytų serverių darbui gerinti*, Daktaro disertacija, Vilnius, 2008.
17. „Fotono“ istorija. Prieiga per internetą: <http://www.fotonas.su.lt/istorija.phtml>, Žiūrėta [ 2012 04 12].
18. *Reconsidering Moore's Transactional Distance Theory*. Prieiga per internetą: <http://www.eurodl.org/?article=374>, Žiūrėta [ 2012 03 22].
19. Armando Rocha Trindade. *The conceptual panorama of distance education and training*. 1993.
20. Diego Kuonen. *A statistical perspective of data mining*. 2004.
21. *STATISTICA Data Mining Solutions, Text Mining and Predictive Analytics Software*. Prieiga per internetą: <http://www.statsoft.com/products/data-mining-solutions>, Žiūrėta [ 2012 03 27].
22. Yunyue Zhu. *High Performance Data Mining in Time Series: Techniques and Case Studies*. 2004.
23. *Course management system*. Prieiga per internetą: <http://net.educause.edu/ir/library/pdf/DEC0302.pdf>, Žiūrėta [ 2012 05 01].
24. Castro F. et al. (2007). Applying Data Mining Techniques to e-Learning Problems. *Studies in Computational Intelligence (SCI)*, vol. 62, p. 183–221.
25. Stravinskienė Auksė, ir kt. *Duomenų gavybos įrankių pritaikymas mažose įmonėse*. Kaunas, 2010.
26. D. Tamašauskas, T. Bersėnas, doc. S. Turskienė, *Duomenų gavybos technologijos nuotolinių studijų kursų duomenų bazių analizei, 7-oji konferencija, Studentų moksliniai darbai*, Šiauliai, 2012.
27. Dr Ray Hoare, *Using CHAID for classification problems*, 2004 m.
28. Multivariate Adaptive Regression Splines (MARSplines). Prieiga per internetą: <http://www.statsoft.com/textbook/multivariate-adaptive-regression-splines/>, Žiūrėta [2012 05 31].

# Anotacija

Lietuviškai

Autorius: Donatas Tamašauskas

Tema: Duomenų gavybos technologijos nuotolinių studijų kursų duomenų bazės analizei

Šiaulių universitetas 2012

Šiuo metu nuotolinės studijos tampa vienu populiariausių studijų būdų Lietuvoje ir pasaulyje, todėl svarbu atlikti nuotolinių studijų proceso veiklų analizę. Duomenų gavybos technologijos padeda analizuoti studijų procesą, jo dalyvių veiklas, besimokančiųjų elgesį ir pan.

Atlikus Šiaulių universiteto *Moodle* duomenų bazės analizę, kurios rezultatais remiantis sukurta Universiteto virtualios mokymosi aplinkos *Moodle* duomenų apdorojimo programinė įranga. Šia programine įranga gaunami įvairūs *Moodle* duomenų bazių duomenys, kurie analizuojami duomenų gavybos technologijomis keliais pjūviais.

English

Author: Donatas Tamašauskas

Subject: Data Mining Technologies for the Analysis of Distance Learning Course Database

Šiaulių University 2012

Nowadays, studies by distance learning are becoming one of the most popular learning methods in Lithuania and the whole world. Therefore it is important to analyze the process of learning by distance activities. Data mining technologies are helping to analyze the process of studies, the activities of its participants, behavior of the learners, etc.

After completing the analysis of the *Moodle* database of the Siauliai university, the University's virtual learning environment *Moodle* data processing software, was created based on its results. With this software a various *Moodle* data is received, which is analyzed by data mining technologies by different means.

# Priedai

## 1. Kompaktinio disko turinys

- Šakniniame kataloge yra elektroninės magistro darbo aprašymo versijos (rinkmenos „aprasymas.doc“ ir „aprasymas.pdf“);
- Kataloge „Diagramos“ įdėtos visos su projektu susijusios diagramos;
- Kataloge „Analizuoti duomenys“ įdėti visi analizuoti duomenys;
- Kataloge „Programinė įranga“ patalpintas nuotolinio studijų centro pradinių duomenų apdorojimo programinės įrangos išeities kodai.

## 2. VMA Moodle sistemų duomenų bazės susisteminančios programinės įrangos diegimo instrukcijos

Norint įdiegti VMA Moodle sistemų duomenų bazės susisteminančią programinę įrangą, reikia programinės įrangos kataloge „blocks“ pakoreguoti failą „bd.php“, kuriame reikia nurodyti šio įrankio duomenų bazės duomenys.

Tam, kad VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga funkcionuotų tinkamai, reikia suteikti galimybę serveryje keisti vykdymo ciklą laiką.

## 3. VMA Moodle sistemų duomenų bazių susisteminanti programinė įranga

Susisteminančią VMA Moodle sistemų duomenų bazių duomenis programinę įrangą galite rasti adresu <http://193.219.168.18/skaiciai/>. Peržiūrėtojo teisėmis yra sukurtas vartotojas, jo duomenys yra: prisijungimo vardas: demo, slaptažodis: demo.

## 4. Studentų mokslinių darbų programa



ŠIAULIŲ  
UNIVERSITETAS  
TECHNOLOGIJOS  
FAKULTETAS

7 – oji konferencija

### STUDENTŲ MOKSLINIAI DARBAI

2012 m. gegužės 16 d.

Šiauliai

## KONFERENCIJA „STUDENTŲ MOKSLINIAI DARBAI“

### I SEKCIJA: ELEKTRONIKOS IR ELEKTROS INŽINERIJA

ŠU Technologijos fakultetas, Vilniaus g. 141, 109 a.

**I posėdis** – 2012-05-16, 10 val. Elektronikos inžinerija

Pranešimo trukmė – 5-10 min.

1. Egidijus Kiuras, magistrantas (ŠU Elektronikos katedra),  
*Informacinių produktų vertinimas pagal žvilgsnio krypties duomenis*  
Vadovas – prof. habil. dr. Vincas Laurutis
2. Tomas Karčauskas, magistrantas (ŠU Elektronikos katedra).  
*Trumpai pateikiamų Miulerio -Laerio iliuzijų tyrimas šuoliniais akių judesiais*  
Vadovas – prof. habil. dr. Vincas Laurutis
3. Gytenis Sudintas, magistrantas (ŠU Elektronikos katedra).  
*Taikinio su nematomais intarpais sekimo tyrimas*  
Vadovas – prof. habil. dr. Vincas Laurutis
4. Sergėj Liaškėvič, magistrantas (ŠU Elektronikos katedra).  
*Žmonių figūrų paieška vaizduose*  
Vadovas – prof. dr. Gintautas Daunys
5. Renatas Montvilas, magistrantas (ŠU Elektronikos katedra).  
*Operatoriaus veido sekimas valdant kameros kryptį*  
Vadovas – doc. dr. Donatas Dervinis
6. Evaldas Borcovas, magistrantas (ŠU Elektronikos katedra).  
*Galvos pozicijos nustatymas ekrano atžvilgiu, panaudojant ant galvos pritvirtintą kamerą*  
Vadovas – prof. dr. Gintautas Daunys
7. Artūras Nakvosas, magistrantas (ŠU Elektronikos katedra).  
*Galvos pozicijos nustatymas kompiuterine regu*  
Vadovas – prof. dr. Gintautas Daunys
8. Andrius Tamašauskis, magistrantas (ŠU Elektronikos katedra).  
*Ekrano valdymas rankų gestais*  
Vadovas – prof. dr. Gintautas Daunys
9. Arūnas Šlenderis, bakalaurantas (ŠU Elektronikos katedra).  
*Vaizdo filtravimas LPLM įrenginiu*  
Vadovas – prof. dr. Gintautas Daunys

II posėdis – 2012-05-16, 12 val.  
Pranešimo trukmė – 5-10 min.

Elektros inžinerija

1. Arvydas Valys, magistrantas (ŠU Elektros inžinerijos katedra)  
***Elektrinio ožinio krano energetinio efektyvumo tyrimas***  
Vadovas – lekt. dr. Arūnas Grigaitis
2. Antanas Šiaulytis, magistrantas (ŠU Elektros inžinerijos katedra)  
***Kintamojo dažnio elektros pavaros valdymo, matavimų ir analizės sistema***  
Vadovas – doc.dr. Gediminas Valiulis
3. Kęstutis Kalkis, magistrantas (ŠU Elektros inžinerijos katedra)  
***Skirtuminių relių charakteristikų tyrimas***  
Vadovas – doc. dr. Enrikas Vilimantas Nevardauskas
4. Remigijus Baltušis, magistrantas (ŠU Elektros inžinerijos katedra)  
***Fotoelektrinės parametrų tyrimas***  
Vadovas – doc. dr. Leonardas Buivis
5. Žilvinas Kripaitis, magistrantas (ŠU Elektros inžinerijos katedra)  
***Elektros kokybės gerinimo priemonių tyrimas***  
Vadovas – doc. dr. Enrikas Vilimantas Nevardauskas

## II SEKCIJA: INFORMATIKOS INŽINERIJA, INFORMATIKA

ŠU Technologijos fakultetas, Vilniaus g. 141, 318 a.

Posėdis – 2012-05-16, 10 val.

Pranešimo trukmė – 5-10 min.

1. Donatas Tamašauskas, magistrantas (ŠU Informatikos katedra).  
*Duomenų gavybos technologijos nuotolinių studijų kursų duomenų bazių analizei*  
Bendraautorai: T. Bersėnas, doc. S. Turskienė
2. Denas Pavlavičius, magistrantas (ŠU Informatikos katedra).  
*Informacinių sistemų inventorius valdymo sistema*  
Bendraautorė doc. Sigita Turskienė
3. Ernestas Stravinskas, magistrantas (ŠU Informacinių technologijų katedra).  
*Programinių agentų kūrimo technologijų analizė*  
Bendraautorius doc. dr. Egidijus Paliulis
4. Renata Baronienė, magistrantas (ŠU Informacinių technologijų katedra).  
*Reliacinių duomenų bazių valdymo sistemų charakteristikų analizė*  
Bendraautorius doc. dr. Egidijus Paliulis
5. Kiril Griazev, magistrantas (ŠU Informacinių technologijų katedra).  
*Struktūrinių duomenų išgavimas nestruktūrizuotose informacinėse sistemose*  
Bendraautorė lekt. Simona Ramanauskaitė
6. Kęstutis Visialga, magistrantas (ŠU Informacinių technologijų katedra).  
*Projektų valdymo modelių analizė*  
Bendraautorė doc. dr. Asta Slotkienė
7. Žilvinas Gasiūnas, magistrantas (ŠU Informacinių technologijų katedra).  
*Sprendimų priėmimų modeliavimas naudojant kontekstinius grafus*  
Bendraautorė doc. dr. Asta Slotkienė
8. Donatas Veikutis, bakalaurantas (ŠU Informacinių technologijų katedra).  
*Optimizavimo paieškos sistemoms metodų analizė*  
Bendraautorė lekt. Simona Ramanauskaitė
9. Lorentis Sverdiolas, bakalaurantas (ŠU Informacinių technologijų katedra).  
*Nutolusio serverio mobilaus valdymo sistema*  
Bendraautorius lekt. Simona Ramanauskaitė



**III SEKCIJA: APLINKOS INŽINERIJA, APLINKOTYRA**  
ŠU Technologijos fakultetas, Vilniaus g. 141, 410 a.

Posėdis – 2012-05-16, 10 val.

Pranešimo trukmė – 5-10 min.

1. Kristiāna Siliņa, Daugavpils University, Environmental science.  
*Biogeninių elementų koncentracijos tyrimas Dauguvos upėje*  
*Concentration of biogenic elements in river Daugava during spring drift expedition.*  
Dāvis Gruberts
2. Ginta Janoviča, Daugavpils University, Environmental science.  
*Pavojingų invazinių augalų rūšys Jaungulbens apylinkėje*  
*Dangerous invasive plant species in Jaungulbens parish territory*  
Santa Rutkovska
3. Anna Brokāne, Daugavpils University, Environmental science.  
*Pavojingų invazinių augalų rūšių paplitimas Subato teritorijoje*  
*Distribution of dangerous invasive plant species in Subate's territory*  
Santa Rutkovska
4. Sanita Zunda, Daugavpils University, Ecology.  
*„Kryževnikų“ poligone esančios atliekų perdirbimo įmonės duomenų analizė*  
*Исследование базы данных хозяйственного полигона «Крижевники»*  
Solvita Mičiule
5. Mantas Svetikas, magistrantas (KTU Aplinkos apsaugos vadyba ir švarioji gamyba)  
*Kai kurių akustikos (garso-triukšmo) parametrų skirtingas poveikis į aplinką*  
*Some acoustic (sound-noise) parameters of the different effects on the environment*  
Prof. habil. dr. Danielius Gužas
6. Mantas Svetikas, magistrantas (KTU Aplinkos apsaugos vadyba ir švarioji gamyba)  
*Monitoringo modelių pritaikymas tiriant aplinkos kenksmingus veiksnius*  
*The monitoring for models of application the investigation of environmental factor*  
Prof. habil. dr. Danielius Gužas
7. Monika Jurgita Brazaitė, bakalaurantė (ŠU Aplinkotyros katedra).  
*Naminės pelėdos (Strix aluco L.) mitybos tyrimai Šilutės rajone*  
*Research of the owl (Strix aluco L.) nutrition in the district of Šilutė*  
Algirdas Dumčius
8. Almeda Šinušaitė, bakalaurantė (ŠU Aplinkotyros katedra).  
*Meteorologinių parametrų įtaka beržo, alksnio ir dilgėlės genčių žiedadulkių koncentracijai 2005-2011 m. Vilniaus mieste*

*The effect of meteorological factors on birch, alder and nettle airborne pollen concentration during the years of 2005-2011 in Vilnius*

Lekt. dr. Laura Veriankaitė

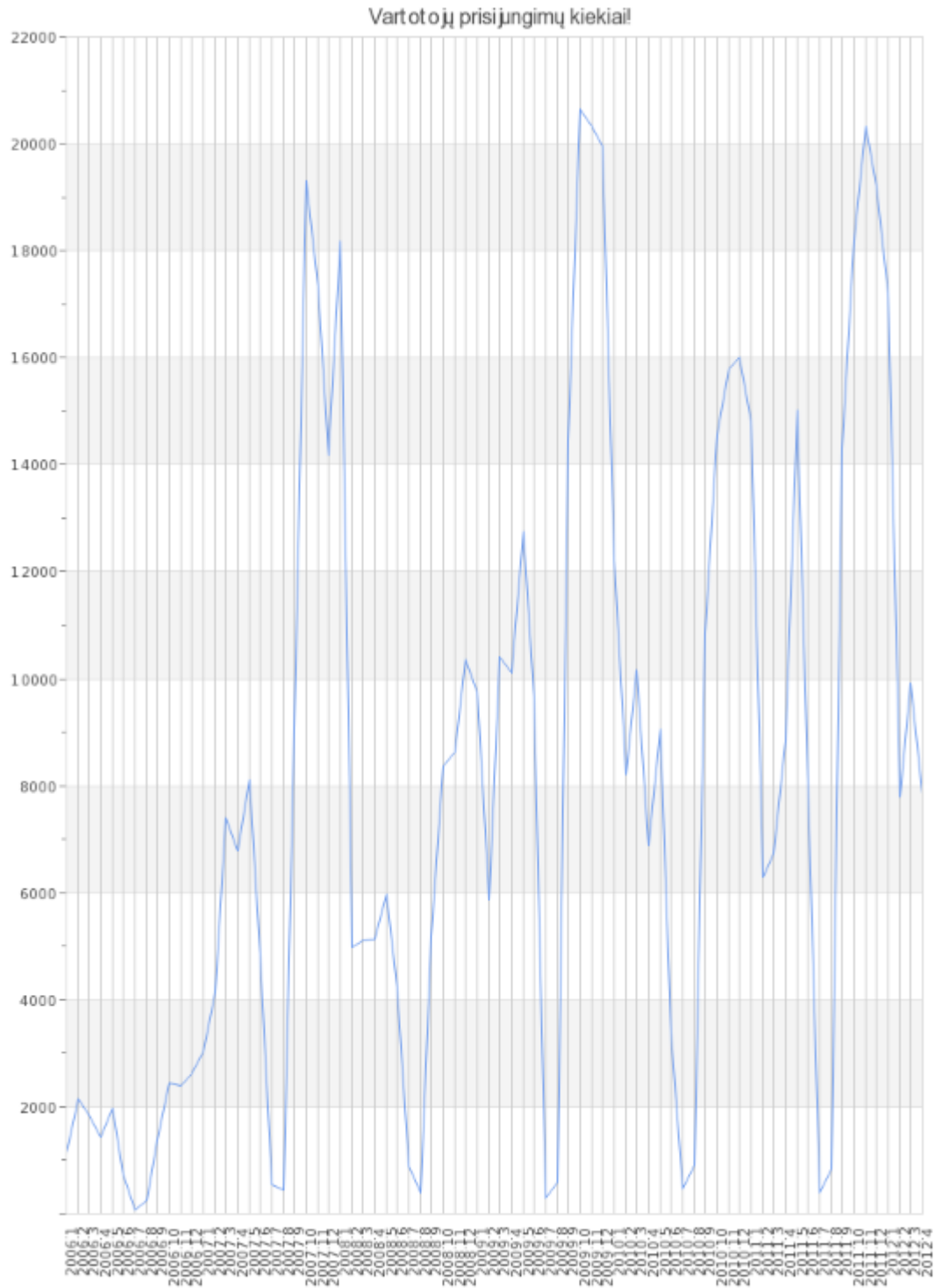
9. Viktorija Songailaitė, bakalaurantė (ŠU Aplinkotyros katedra).  
*Meteorologinių parametrų įtaka beržo, lazdyno ir kiečio augalų genčių žiedadulkių sklaidai Šiaulių mieste 2004-2011 metais*  
*The effect of meteorological parameters on birch, hazelnut and wormwood plants genus airborne pollen concentration in Šiauliai city during the years of 2004-2011*  
Lekt. dr. Laura Veriankaitė
10. Raimonda Pocevičiūtė, bakalaurantė (ŠU Aplinkotyros katedra).  
*Titnagdumblių įvairovė Šiaulių ir Šilalės rajonų paviršiniuose vandens telkiniuose*  
*Diversity of diatoms from surface water bodies in Šiauliai and Šilalė district*  
Lekt. dr. Martynas Kazlauskas
11. Rūta Vaitkutė, bakalaurantė (ŠU, Statybos inžinerijos katedra).  
*Ančios upės vandens taršos tyrimas*  
*Ančia river water pollution research*  
Prof. dr. Vaclovas Tričys
12. Aušra Sadauskaitė, bakalaurantė (ŠU, Statybos inžinerijos katedra).  
*Kietųjų dalelių koncentracijos tyrimas Tauragės miesto ore*  
*Research of solid particle concentrations in the air of Tauragė town*  
Prof. dr. Vaclovas Tričys
13. Neringa Dimšaitė, bakalaurantė (ŠU, Statybos inžinerijos katedra).  
*Fizikinės taršos tyrimas Kelmės mieste*  
*Research of Physical pollution in the city of Kelme*  
Prof. dr. Vaclovas Tričys
14. Greta Karavackaitė, bakalaurantė (ŠU, Statybos inžinerijos katedra).  
*Autotransporto taršos poveikio įvertinimas Gegužių gatvės aplinkoje*  
*Environmental impact of motor vehicle pollution in Gegužiai street*  
Lekt. Robertas Klimas
15. Aistė Karpušenkaitė  
*Radviliškio rajono savivaldybės aplinkos oro kokybės valdymo programos projekto parengimas*  
*The Project Preparation of Ambient Air Quality Management Programme in Radviliskis District Municipality*  
Lekt. Robertas Klimas
16. Eventas Šegžda, bakalaurantas (ŠU, Statybos inžinerijos katedra).  
*Kietųjų dalelių koncentracijos tyrimas Šiaulių miesto centrinėje dalyje*  
*Research of solid particulate concentrations in the central part of Šiauliai city*  
Prof. dr. Vaclovas Tričys

**IV SEKCIJA: MECHANIKOS INŽINERIJA, STATYBOS INŽINERIJA**  
ŠU Technologijos fakultetas, Vilniaus g. 141, 304 a.

Posėdis – 2012-05-16, 10 val.  
Pranešimo trukmė – 5-10 min.

1. Marius Reizgevičius, doktorantas (VGTU).  
***Keturių dimensijų projektavimo modelio efektyvumas***  
Bendraautorė Rasa Reizgevičienė
2. Justas Labanauskas, bakalaurlant (ŠU Statybos inžinerijos katedra).  
***Langų įtakos pastato šiluminėms savybėms vertinimas***  
Bendraautorius lekt. Marius Reizgevičius
3. Mantas Tankūnas, Vytautas Urbonas, bakalaurlantai (ŠU Statybos inžinerijos katedra).  
***Sijų stiprumo statmename pjūvyje apskaičiavimas pagal kreivinę Euronormų gniuždomo betono įtempių diagramą***  
Bendraautorius prof. Ipolitas Židonis
4. Laura Reizgevičiūtė, bakalaurlantė (ŠU Statybos inžinerijos katedra).  
***Ikimokyklinės paskirties pastatų funkcinių zonų modeliavimas įvedant dirbtinį intelektą***  
Bendraautorius lekt. Marius Reizgevičius
5. Arūnas Venckus, magistrantas (ŠU Mechanikos inžinerijos katedra).  
***Dviračio miniklio švaistiklio stiprumo ir standumo tyrimas***  
Bendraautorius doc. dr. Artūras Sabaliauskas
6. Imantas Norvilis, magistrantas (ŠU Mechanikos inžinerijos katedra).  
***Detalių konstrukcinių elementų parametru tikslumo tyrimas***  
Bendraautorė doc. dr. D. Čikotienė
7. Gintaras Stašys, magistrantas (ŠU Mechanikos inžinerijos katedra).  
***Centrinio šildymo kieto kuro katilo „KALVIS – 4D“ parametru tyrimas***  
Bendraautorius prof. J. Bareišis
8. Ramūnas Šernius, magistrantas (ŠU Mechanikos inžinerijos katedra).  
***Dinamininių apkrovų poveikio dviračio vairui tyrimas***  
Bendraautorius doc. dr. S. Rimovskis
9. Arnas Slotkus, magistrantas (ŠU Mechanikos inžinerijos katedra).  
***Žmogiškojo veiksnio įtakos dviračių surinkimo proceso kokybei tyrimas***  
Bendraautorė doc. dr. D. Čikotienė

## 5. Vartotojų prisijungimų kiekiai kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio

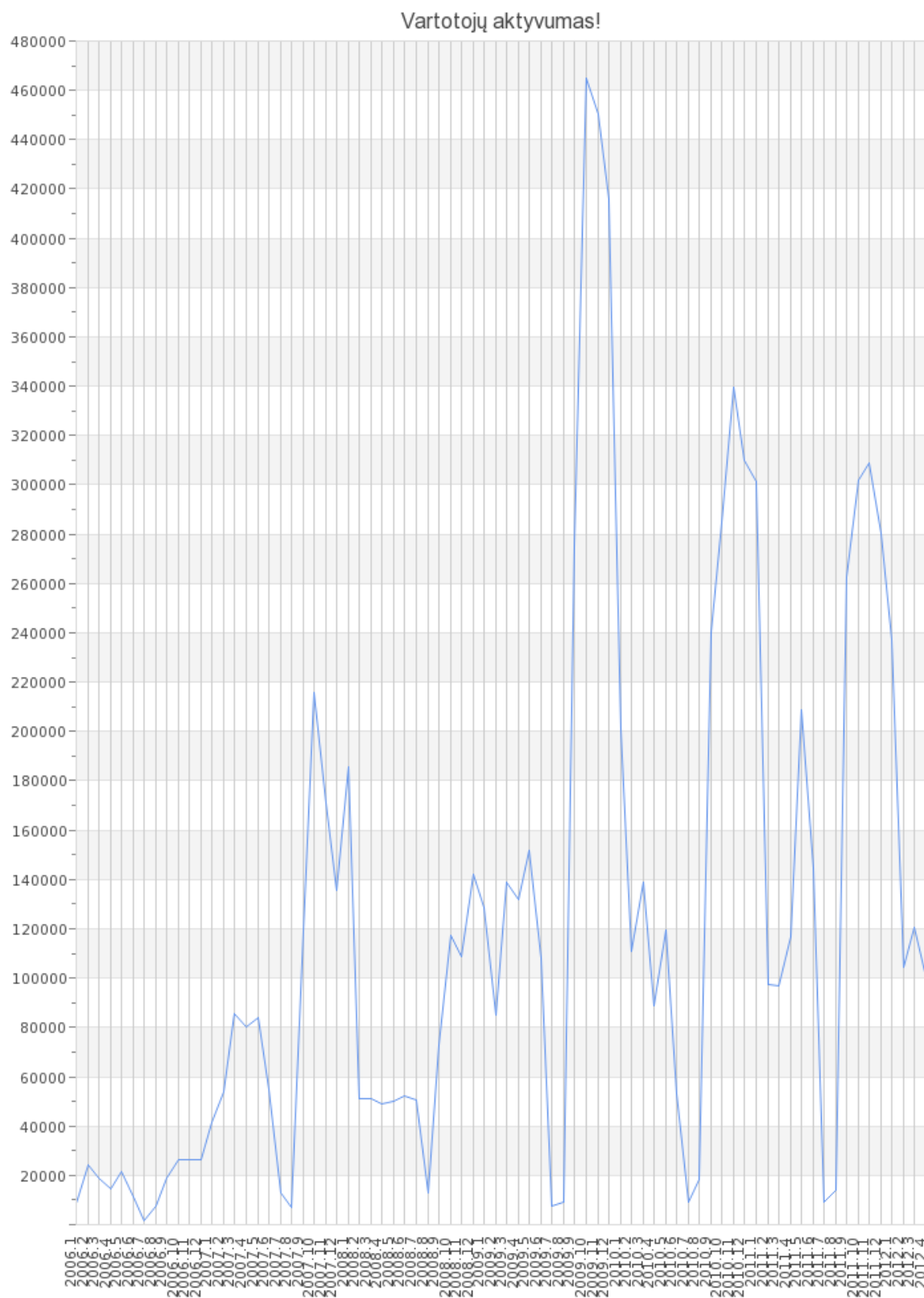








## 8. Vartotojų aktyvumas kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio







# 10. Sekų duomenyse paieška, pagal prisijungimų kiekį kas mėnesį nuo 2006 m. iki 2012 m. gegužės mėnesio

Summary of sequential rules (Vartotojų aktyvumo kiekiai pagal pasirinkta intervala)		Head		Support(%)	Confidence(%)
Min. support = 20.0%, Min. confidence = 10.0%					
Max. size of an itemset = 10					
	Body		Head		
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26					
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51					
52					
53					
54					
55					
56					
57					
58					
59					
60					
61					
62					
63					
64					
65					
66					
67					
68					
69					
70					
71					
72					
73					
74					
75					
76					
77					
78					
79					
80					
81					
82					
83					
84					
85					
86					
87					
88					
89					
90					
91					
92					
93					
94					
95					
96					
97					
98					
99					
100					
101					
102					
103					
104					
105					
106					
107					
108					
109					
110					
111					
112					
113					
114					
115					
116					
117					
118					
119					
120					
121					
122					
123					
124					
125					
126					
127					
128					
129					
130					
131					
132					
133					
134					
135					
136					
137					
138					
139					
140					
141					
142					
143					
144					
145					