



**KAUNO TECHNOLOGIJOS UNIVERSITETAS  
INFORMATIKOS FAKULTETAS**

**Martynas Ulinskas**

**ŽMOGAUS NUOVARGIO BŪSENOS ATPAŽINIMO TYRIMAS  
NAUDOJANT KOMPIUTERIO KLAVIATŪROS KLAVIŠŲ  
PASPAUDIMO DINAMIKOS DUOMENIS**

Baigiamasis magistro projektas

**Vadovas**

Prof. Dr. Robertas Damaševičius

**KAUNAS, 2017**

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**INFORMATIKOS FAKULTETAS**

**ŽMOGAUS NUOVARGIO BŪSENOS ATPAŽINIMO TYRIMAS  
NAUDOJANT KOMPIUTERIO KLAVIATŪROS KLAVIŠŲ  
PASPAUDIMO DINAMIKOS DUOMENIS**

Baigiamasis magistro projektas  
**Informatika (621I10003)**

**Vadovas**

Prof. dr. Robertas Damaševičius

**Recenzentas**

Prof. dr. Rytis Maskeliūnas

**Projektą atliko**

Martynas Ulinskas

**KAUNAS, 2017**



KAUNO TECHNOLOGIJOS UNIVERSITETAS

Informatikos fakultetas

(Fakultetas)

Martynas Ulinskas

(Studento vardas, pavardė)

Informatika, 621I10003

(Studijų programos pavadinimas, kodas)

Baigiamojo projekto „Pavadinimas“

## AKADEMINIO SAŽININGUMO DEKLARACIJA

20 17 m. gegužės 16 d.  
Kaunas

Patvirtinu, kad mano, **Martyno Ulinsko**, baigiamasis projektas tema „Žmogaus nuovargio būsenos atpažinimo tyrimas naudojant kompiuterio klaviatūros klavišų paspaudimo dinamikos duomenimis“ yra parašytas visiškai savarankiškai ir visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

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

\_\_\_\_\_  
(vardą ir pavardę įrašyti ranka)

\_\_\_\_\_  
(parašas)

Martynas, Ulinskas. Žmogaus nuovargio būsenos atpažinimo tyrimas naudojant kompiuterio klaviatūros klavišų paspaudimo dinamikos duomenimis. Magistro baigiamasis projektas / vadovas Prof. dr. Robertas Domaševičius; Kauno technologijos universitetas, Informatikos fakultetas.

Mokslo kryptis ir sritis: technologijos mokslai, informatikos inžinerija (07T)

Reikšminiai žodžiai: žmogaus – kompiuterio sąsaja, žmogaus būsenos atpažinimas.

Kaunas, 2017. 54 p.

## **SANTRAUKA**

Baigiamajame magistro darbe atliekamas žmogaus nuovargio nustatymo, naudojantis kompiuterio klaviatūros klavišų paspaudimo dinamikos duomenimis, tyrimas. Šiame tyrime analizuojama nuovargio įtaka žmogaus elgsenai ir savijautai bei apžvelgiami metodai, kurie taikomi nuovargio nustatymui. Remiantis, literatūros apžvalgos metu, nustatytais esamų sistemų trūkumais, pateikiamas naujas nuovargio nustatymo metodas, suprojektuojama ir realizuojama programinė įranga bei atliekamas sistemos tyrimas. Darbo pabaigoje, remiantis gautais tyrimo rezultatais, pateikiamos išvados.

Tyrimą sudaro trys dalys: analizės, projektavimo ir eksperimentinės. Analizės metu buvo tiriama kaip žmogui pasireiškia nuovargis, kokia jo įtaka elgsenai, kokie metodai naudojami nuovargio nustatymui. Remiantis surinkta informacija, buvo pasiūlyta nauja nuovargio nustatymo metodika, kuri galėtų nustatyti žmogaus nuovargį, darbo su kompiuteriu metu.

Projektavimo metu atliktas aplikacijos, kuri galėtų nustatyti nuovargį, projektavimas ir realizavimas. Aplikacijai sukurti buvo naudojami *Microsoft VisualStudio 2015* programavimo paketas ir *C#* programavimo kalba. Duomenų talpinimui ir saugojimui buvo naudojama *Microsoft SQL Server 2016* duomenų bazė.

Eksperimento metu keturi dalyviai dvi savaites rinko tekstą kompiuteriu, kuriam buvo fiksuojamos ir saugomos klavišų dinamikos. Naudojantis surinktais duomenimis buvo testuojamos skirtingų modifikacijų nuovargio nustatymo sistemos. Gauti rezultatai rodo, kad galima pasiekti 89,79 % nuovargio būsenų nustatymo tikimybę, su 4,99 % standartiniu nuokrypiu.

Ulinskas, Martynas. HUMAN'S FATIGUE DETECTION USING KEYSTROKE DYNAMICS  
Master's thesis in informatics engineering / supervisor assoc. prof. Robertas Damaševičius. The  
Faculty of Informatics, Kaunas University of Technology.

Research area and field: technological sciences, informatics engineering (07T)

Key words: keystrokes dynamics, human's fatigue detection

Kaunas, 2017. 54 p.

## **SUMMARY**

In this master thesis, we aimed to investigate humans' fatigue detection by using keystroke dynamics. The aim of this research was to analyze existing approaches of humans' fatigue detection, suggest a more human friendly detection method, implement and test it. Furthermore, we analyzed these systems' results and drew appropriate conclusions.

The research consists of three main parts: analysis, designing and implementing testing systems. In the analysis part, we analyzed how fatigue can affect humans' behavior, how it can be detected and also how current detection methods are implemented. Based on the background of our analysis, we hoped to suggest a more human friendly fatigue detection method.

In the designing part, we designed an application which can detect keystroke dynamics, store it in database and determinate humans' fatigue. We implemented this application using *Microsoft Visual Studio 2015* and *C#* programming language. To store our data, we used *Microsoft SQL server 2016* database.

In the experimental part, the research of the user's keystroke dynamics data classification is being carried out. Carrying out the experiment, we collected four participants text typing data for two weeks. With the collected dataset we evaluated to what precision we could identify human's fatigue. According to the results of our approach, we can achieve an accuracy of 89,79 % of identifying a human's fatigue.

# Turinys

Terminų ir santrumpų žodynas.....	7
ĮVADAS.....	8
1. LITERATŪROS APŽVALGA.....	9
1.1. Metodai naudojami nuovargio matavimui .....	9
1.2. Biometriniai duomenys .....	10
1.3. Klavišų paspaudimo dinamika .....	11
1.3.1. Klavišų paspaudimo savybės.....	11
1.3.2. Klavišų paspaudimo dinamikos sistemos .....	12
1.4. Klavišų paspaudimo dinamikos modelių kūrimo metodai.....	13
1.4.1. Dirbtinis Neuroninis tinklas.....	13
1.4.2. Atraminių vektorių mašina .....	14
1.5. Klasifikatorių mokymui naudojamų savybių atrinkimas .....	15
2. TEORINIS METODŲ PAGRINDIMAS.....	16
2.1. Dirbtinis neuroninis tinklas .....	16
2.1.1. Matematinis neurono modelis.....	16
2.1.2. Aktyvavimo funkcijos .....	16
2.1.3. Vienasluoksnis neuroninis tinklas .....	17
2.1.4. Daugiasluoksnis neuroninis tinklas .....	18
2.1.5. Neuroninio tinklo mokymas .....	18
2.2. Atraminių vektorių mašinos .....	20
2.2.1. Tiesinis binarinis klasifikatorius.....	20
2.2.2. Binarinis klasifikatorius netiesiškai atskiriamiems duomenims.....	22
2.2.3. Netiesinės atraminės vektorių mašinos.....	23
3. REALIZACIJA .....	24
3.1. Projekto planas .....	24
3.2. Funkciniai reikalavimai.....	24
3.3. Nefunkciniai reikalavimai .....	25
3.4. Projekto duomenų bazė .....	25
3.5. Duomenų gavyba.....	26
3.6. Duomenų klasifikavimas.....	28
3.7. Duomenų normalizavimas.....	29
3.8. Parametrų atrinkimas.....	29
3.9. Dvejetaisiais būsenų nustatymo metodas.....	30
3.9.1. Trijų klasifikatorių sistema .....	30
3.9.2. Klavišų porų klasifikatorių sistema .....	31
3.9.3. Trijų klasifikatorių sistema su klavišų porų svarbumo įvertinimu .....	33
3.10. Weka programinis įrangos paketas.....	34
4. EKSPERIMENTINĖ DALIS .....	35
4.1. Tyrimo dalyvių vidutinės KD savybių vertės.....	35
4.2. Branduolio funkcijos ir baudos koeficientų parinkimas.....	37
4.3. Klavišų dinamikos savybių atrinkimas.....	38
4.4. Trijų klasifikatorių sistemos testavimas .....	41
4.5. Klavišų porų klasifikatorių sistemos testavimas .....	43
4.6. Trijų klasifikatorių sistemos, su klavišų porų svarbumo įvertinimu, testavimas .....	45
4.7. Tyrimo rezultatų apibendrinimas ir apribojimai .....	47
5. IŠVADOS.....	51
6. LITERATŪROS ŠALTINIAI.....	52
7. PRIEDAI.....	54

## **Terminų ir santrumpų žodynas**

ANN – dirbtinis neuroninis tinklas (angl. Artificial Neural Network);  
BFS – tiesinio parinkimo algoritmu (angl. Best Forward Selection)  
EER – lygus paklaidos santykis (angl. Equal Error Rate);  
FAR – klaidingas patvirtinimo santykis (angl. False Acceptance Rate)  
FRA – klaidingas neidentifikavimo santykis (angl. False Rejected Rate)  
H – paspaudimo laikas (angl. Hold Time);  
HNN – paslėptieji Markovo modeliai (angl. Hidden Markov Mmodel)  
KD – klavišų dinamika (angl. Keystroke Dinamics);  
KNN – k-artimiausi kaimynai (angl. K-Nearest Neighbors)  
KP – klavišų pora;  
PCA – Svarbiausių komponentų analizė (angl. Principal Componnets Analysis)  
PĮ – programinė įranga;  
PP – paspaudimo – paspaudimo delsos laikas (angl. Press – Press latency);  
PR – paspaudimo – atleidimo delsos laikas (angl. Press – Release latency);  
RBK – radialinė branduolio funkcija (angl. Radial Basis Kernel)  
RP – atleidimo / paspaudimo delsos laikas (angl. Release -Press latency);  
RR – atleidimo – atleidimo delsos laikas (angl. Release – Release latency);  
SVM – atraminių vektorių mašina (angl. Support Vector Machine);

## ĮVADAS

Nuovargis yra įvardijamas kaip viena iš dažniausiai nelaimės lemiančių priežasčių. Bendros nuomonės laikosi ir mokslininkai, kurie atskaitose pateikia, kad miego trūkumas ir nepastovus miego ar poilsio ritmas yra pagrindinės nelaimingų atsitikimų priežastys [1]. Lager'io su komanda atlikta miego trūkumo pasekmių analizė [2] parodė, jog dėl nuovargio Jungtinėse Amerikos Valstijose įvyko apie 50 % motorinių transporto priemonių avarijų, kurių piniginė išraiška atitiktų 35 mlrd. dolerių nuostolius. Nuovargis yra įvardijamas ir tokių nelaimių priežastimi kaip Černobylio atominės jėgainės sprogimas ar Exxon Valez naftos išsiliejimas [3].

Dėl nuovargio gali didėti ne tik nelaimingų atsitikimų tikimybė, bet ir blogėti darbo ar mokslo pasiekimai. H.H. Sieverten'as su komanda [4] dvejus metus tyrė, kokią įtaką atliekamo standartizuoto testo laikas turi galutiniam mokinio įvertinimui, ir padarė išvadą, kad atliekant testą viena valanda vėliau testo rezultatai sumažėjo 0,9 %, o 20 – 30 min pertrauka testo rezultatus pagerino 1,7 %.

Siekiant nustatyti nuovargio poveikį žmogaus motorinei ir psichinei veiklai, dažnai nuovargis lyginamas su alkoholio poveikiu [5]. Alkoholis naudojamas ne tik todėl, kad tiesiogiai koreliuotas su nuovargiu (24 valandas nemiegojęs žmogus po testų rodo tokius pačius rezultatus, kaip asmuo, kurio kraujyje yra 0,08 % alkoholio[6]), tačiau ir dėl atliktų tyrimų, kurie išsamiai nusako alkoholio poveikį žmogaus motorinei ir psichinei veiklai.

### **Probleminė sritis:**

Šiame darbe nagrinėjama problema: objektyvus (kiekybinis) žmogaus nuovargio būsenos nustatymas.

### **Darbo tikslai ir uždaviniai:**

Darbo tikslas – objektyviai nustatyti žmogaus nuovargio būseną ir pateikti jos kokybinį įvertinimą. Tikslui pasiekti suformuoti šie darbo uždaviniai:

- Atlikti žmogaus nuovargio vertinimo metodų apžvalgą .
- Atlikti klavišo paspaudimo proceso charakteristikų analizę ir išskirti esminius požymius.
- Atlikti duomenų klasifikavimo metodų analizę.
- Pasiūlyti klaviatūros paspaudimų dinamikos sekimo sistemos žmogaus nuovargio būsenai nustatyti modelį.
- Realizuoti nuovargio būsenos sekimo sistemą.
- Atlikti nuovargio būsenos sekimo sistemos eksperimentinį tyrimą.



## 1. LITERATŪROS APŽVALGA

Nuovargis paprasta gali būti atpažįstamas pagal žmogaus veiksmus. Dėl energijos ir motyvacijos stokos, žmogus nebegali dirbti jam būdingu režimu, atsiranda užmaršumas, prastėja atmintis, ar negali galvoti apie tai, kas gali nutikti netolimoje ateityje. Be to, nuovargis gali lemti menką bendravimą ir budrumą, kurie yra labai svarbūs atliekant kasdienes darbus. S. Kim'as ir kt. [3] apibrėžė faktorius, kurie labiausiai lemia nuovargį ir nuovargio simptomus. Visi aptarti simptomai būdingi ir asmenims apsvaigusiems nuo alkoholio. Prie labiausiai nuovargį lemiančių veiksnių priskiriama:

- sutrikęs miegas;
- vitaminų trūkumas;
- amžius;
- fizinio aktyvumo trūkumas;
- monotoniškumas;
- nepastovus darbo grafikas.

Tokie veiksniai kaip fizinio aktyvumo trūkumas, monotoniškumas ar nepastovus darbo grafikas yra būdingi tam tikra veikla užsiimantiems asmenims. Pavyzdžiui monotoniškumo sąvoka dažnai naudojama vairavimo elgsenos tyrimuose, kuriuose yra nustatyta, kad nuovargis ir monotoniškumas mažina vairuotojo dėmesį [8]. Thiffault'as ir Bergeron'as [9] nustatė, kad monotoniškumo mažėjimas turi teigiamo efekto mažinant nuovargį, kuris susijęs su vairavimu.

Nuovargį galima atpažinti pagal tam tikrus ženklus ar simptomus. Dažniausiai pasireiškiantys nuovargio ženklai:

- užmaršumas;
- menkas bendravimas;
- žiovilys;
- nuolatinis išsekimas;
- nevalingas užsnūdimas;
- budrumo stoka;
- mieguistumas.

Nuovargiui matuoti nėra tiesioginio būdo, tačiau atlikti tyrimai parodė, jog nuovargis koreliuoja su alkoholio koncentracija kraujyje [7] ir turi tokius pat simptomus. Pagal Lietuvos Respublikos administracinių teisės pažeidimų kodeksą [10] asmuo negali vairuoti transporto priemonės, jeigu alkoholio koncentracija kraujyje viršija 0,041 %. Kodekse taip pat nurodomi atvejai, kuriais alkoholio koncentracija nėra toleruojama ir turi būti lygi nuliui. Tuo tarpu valstybinė darbo inspekcija tolerancinės alkoholio koncentracijos ribos nenurodo [11], tačiau ji, darbdavį įpareigoja nušalinti darbuotoją nuo pareigų, jeigu jis yra neblaivus. Kadangi, nuovargis turi tokias pačias pasekmes kaip ir alkoholio vartojimas, dirbantiems asmenims jis turi būti sekamas ir matuojamas pastoviai. Nustačius, jog jis viršija leistiną ribą, asmuo privalo daryti pertrauką, nes tolimesnis veiklos tęsimas gali kelti pavojų jam pačiam ir aplinkiniams.

### 1.1. Metodai naudojami nuovargio matavimui

Nuovargio nustatymui ir įvertinimui reikalingos jautrios ir veiksmingos priemonės, kurias būtų galima naudoti kasdieninėje aplinkoje. Šiam tikslui pasiekti dažnai naudoja programinę įrangą (PI), kuri per kompiuterio įvesties įrenginius fiksuoja asmens reakcijos laiką, tikslumą, atmintį ir kt.. A. M. Williamson'as ir komanda [6] išskiria aštuonis testus, kurių rezultatais galima įvertinti nuovargio poveikį žmogui:

1. Reakcijos laikas [12] – testas skirtas matuoti vizualinę reakciją. Šiame teste asmuo turi sekti prieš laikrodžio rodyklę judantį apskritimą ir klaviatūra fiksuoti apskritimo spalvos pasikeitimą iš geltonos į raudoną. Testas trunka dvi minutes, per kurias spalvos keičiasi 40 kartų. Testas matuoja kiek laiko užtrunka asmeniui sureaguoti į spalvos pasikeitimą ir skaičius spalvų pasikeitimų, į kurias asmuo nesureagavo.
2. Kitimo sekimas [6] – testas matuoja akių ir rankų koordinaciją, kurios esmė – išmatuoti, kaip tiksliai, pagal pateikiamą vizualinę informaciją, asmuo gali valdyti rankos judesius.

Testas sudarytas iš judančio geltono apskritimo ir mažesnio žalio apskritimo, kurį subjektas, naudodamasis kompiuterio pele, turėjo išlaikyti geltonajame apskritime. Testas, pagal taip, kaip subjektas tiksliai išlaiko žalią apskritimą geltoname, sumažina arba padidina geltono apskritimo judėjimo greitį kas penkias sekundes. Šiame teste matuojamas geltono apskritimo judėjimo nepastovumas.

3. Dviguba užduotis [12] – susideda iš anksčiau minėtų reakcijos laiko ir kitimo sekimo testų. Testas seka, kaip žmogus sugeba atlikti dvi užduotis lygiagrečiai viena kitai. Šiame teste subjektas, naudodamasis pele, turi sekti judantį apskritimą ir kartu klaviatūra fiksuoti apskritimo spalvos apsikeitimus. Šiame teste matuojama reakcijos laikas į spalvos apsikeitimus, nesureaguotus spalvų pasikeitimus ir judėjimo netolygumus.
4. *Mackwoth* laikrodžio atidumo testas [13] – matuoja gebėjimą monotoniškose situacijose išlaikyti koncentraciją. Testas sudarytas iš 24 taškų, kurie nutolę vienodais atstumais ir sudaro apskritimo formą. Taškai, vienas paskui kitą, blyksteli kas 500 ms. Atsitiktiniu intervalu (vidutiniškai kartą per minutę) vienas sekos taškas yra praleidžiamas. Subjekto užduotis fiksuoti, kada sekoje yra praleidžiamas taškas. Testas trunka 15 minučių per kurias matuojama, kaip greitai asmuo sureagavo į praleistą tašką, teisingų ir klaidingų reakcijų skaičiai.
5. Skaitmeninis simbolių kodavimas [14] – neuropsichologinis testas, kuris matuoja žmogaus smegenims padarytą žalą, depresiją ar dimensiją. Testas sudarytas iš dešimties skaičių ir simbolių porų, kurie atvaizduojami kompiuterio ekrano viršuje. Subjektui ekrano apačioje, atsitiktine seka, pateikiama dešimt simbolių. Jo užduotis – kuo greičiau, pagal matomus simbolius ekrano apačioje, surašyti jų skaičių atitikmenis. Testo atlikimui skiriama 90 sekundžių. Testas fiksuoja kiek buvo suvesta simbolių atitikmenų, gerai ir klaidingai įvestų simbolių santykį procentais ir vidutinį atitikmens įvedimo greitį.
6. Vizualinė paieškos užduotis [6] – matuoja kaip greitai subjektas gali rasti ir identifikuoti fundamentalią vizualinę informaciją. Užduotį sudaro dvylika iteracijų. Vienoje iteracijoje parodomos maža ir didelė (60 simbolių) imtys. Maža imtis gali būti dviejų tipų – dviejų ir šešių simbolių. Subjekto užduotis – identifikuoti ar mažos imties simboliai taip pat yra ir didelėje imtyje. Testas fiksuoja asmens imčių įvertinimo laiką ir teisingų įvertinimų skaičių.
7. Nuoseklios sekos atkartojimas [6] – atminties testas, kuris kompiuterio ekrane atvaizduoja  $3 \times 3$  kvadratų matricą. Kvadratai atsitiktine seka užsidega 250 ms. Subjektas, pasibaigus sekai, privalo su kompiuteriu pele, laikantis anksčiau matytos sekos, paspausti atitinkamą kvadratą matricoje. Testas tęsiamas kol asmuo suklysta atkartodamas matytą seką. Testas kartojamas tris kartus per kuriuos fiksuojama atkartotos sekos ilgis.
8. Loginis samprotavimas [15] – protinių gabumų testas, kuriame pateikiami 64 sakiniai kurie nusako dviejų raidžių (A ir B) seką. Kiekvieno sakinio gale nurodomos minėtos raidės ir vartotojas turi identifikuoti ar sakinyje nurodoma seka atitinka pavaizduotų raidžių seką. Pavyzdžiui: B seka A ar B neina pirma A teiginiai teisingi sekai AB ir neteisingi BA. Testu fiksuojama teisingų ir neteisingų atsakymų skaičių ir laiką reikalingą pateikti teisingą atsakymą.

A. M. Williamson'o tyrimas parodė, jog miego trūkumas turėjo įtakos testų rezultatams. Įtaka rezultatams ypač buvo pastebima monotoniškame ir pasyviame *Mackwoth* laikrodžio atidumo teste ir reakcijos laiko teste, kuriame reikalinga greita reakcija į vizualinius pokyčius. Tačiau visi šie testai reikalavo atsitraukti nuo tiesioginės veiklos ir atlikti testus. Bendras testams atlikti reikalingas laikas yra 26 minutės nepridedant dviejų testų, kurių laikas nėra ribojamas[6].

Žmogaus ar jo būsenų identifikavimui gali būti naudojami biometriniai asmens duomenys. Šie duomenys plačiai naudojamas asmens identifikavimui arba autentikavimui [16], tačiau yra atliktų tyrimų [17], kuriuose teigiama, kad stresas, nuovargis, išsekimas ir kt. keičia asmens biometrinius duomenis – rašymo ritmą, eiseną ar balsą.

## 1.2. Biometriniai duomenys

Biometriniai duomenys yra fiziniai ar elgsenos bruožai, kurie kiekvieną žmogų padaro unikaliu ir gali būti panaudoti kaip asmenybės patvirtinimo priemonė. Biometrinių duomenų negalima pamesti,

pavogti ar nugirsti. Net, jeigu biometriniai duomenys būna kaip fiziškai pažeidžiami, jie vis tiek garantuoja šimtaprocentinį patikimumą. Pirštų antspaudai yra viena iš populiariausiai ir paprasčiausiai panaudojamų biometrinių duomenų charakteristikų.

Biometrinius duomenis galima suskirstyti į fizines ir elgsenos charakteristikas. Fizinėms charakteristikoms galima priskirti pirštų anspaudus, akies rainelę, veidą. Visos fizinės charakteristikos gali keistis tik įvykus fiziniam pažeidimui. Prie elgsenos charakteristikų priskiriama klaviatūros klavišų dinamika (angl. *Keystroke Dynamics*) (KD), eiseną, balsą, kompiuterinės pelės naudojimo bruožai, asmens parašas. Šie veiksniai turi pastovią ir kintamą dedamąsias. Pastovi elgsenos dedamoji priklauso nuo fizinių žmogaus duomenų ir nekinta. Tuo tarpu kintama dedamoji, priklausomai nuo žmogaus psichologinės būsenos, gali kisti [18].

### 1.3. Klavišų paspaudimo dinamika

Idėja, kad kiekvienas asmuo turi savitą rašymo ritmą, kilo dar dvidešimtajame amžiuje, kai telegrafo operatoriai pradėjo atpažinti vienas kitą, pagal persiunčiamas žinutes. Persiunčiamos žinutės turėdavo specifinius spausdinimo šablonus, pagal kuriuos buvo galima identifikuoti operatorių.

Klavišų paspaudimo dinamika yra viena iš biometrinių charakteristikų, kuri remiasi prielaida, kad skirtingi žmonės turi skirtingas rašymo manieras. Neuro-fiziologinis veiksnys lemia unikalų asmens parašą. Šis veiksnys taip pat lemia skirtingų žmonių skirtingas rašymo klaviatūra charakteristikas [19].

#### 1.3.1. Klavišų paspaudimo savybės

Asmeniui rašant klaviatūra galima išskirti keletą savybių, kurios gali būti naudojamos kaip biometriniai duomenys:

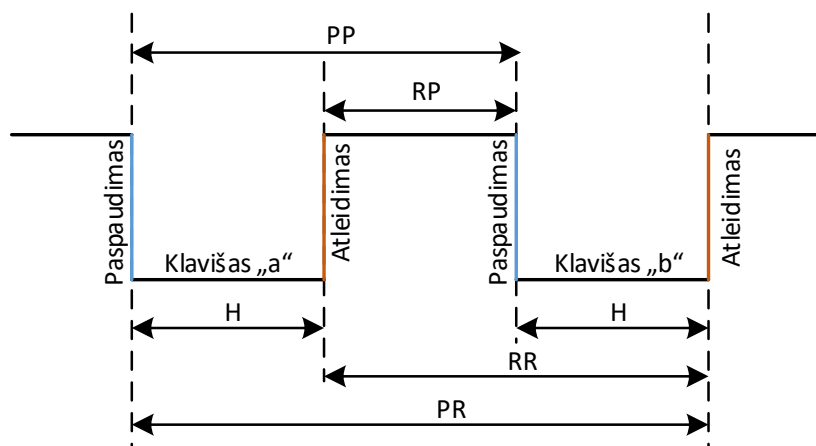
- Klavišų paspaudimo laikas;
- Slėgis, kuriuo spaudžiami klavišai rašant (šis metodas reikalauja papildomos įrangos, kuri fiksuotų slėgį, kuriuo spaudžiamas klavišas);
- Klavišo dalis, kuria pirštas jį nuspaudžia (reikalinga kamera, kuri fiksuotų pirštų padėtis ant klaviatūros);
- Pirštas kuriuo spaudžiamas konkretus klavišas (taip pat reikalinga kamera, kuri fiksuotų pirštų padėtis ant klaviatūros) [19].

Papildoma įranga, tokia kaip – kameros ar specifinės klaviatūros, kurios gali matuoti spaudžiamo klavišo slėgį, reikalauja papildomų investicijų ir apsunkina vartojimą prietaisu, todėl praktikoje jos nėra naudojamos.

Dažniausiai klavišų dinamika aprašoma naudojantis  $N = (t_1; t_2 \dots t_n)$  duomenų seka, kur  $t_i$  yra dvimatis vektorius sudarytas iš absoliučių klavišo paspaudimo  $t^p$  ir atleidimo  $t^a$  laikų –  $t_i = [t_i^p; t_i^r]$ . Norit sudaryti subjekto rašymo šabloną, būtina normalizuoti duomenų seką  $N$ . Sekos normalizavimas atliekama absoliučius sekos  $t_n$  laikus pakeičiant nuosekliais laiko intervalais  $t_i^n = t_{i-1}^x - t_i^x$ , kur  $x$ , nepriklausomai nuo  $t_{i-1}$  ir  $t_i$  gali būti paspaudimo arba atleidimo laikas [20].

A. Morales'as ir komanda [20] išskiria penkias KD savybes, kurios geriausiai charakterizuoja subjekto KD šabloną (1.1 pav.):

- paspaudimo laikas (angl. *Hold Time*) (H)– tai laiko skirtumas tarp klavišo  $t_i$  klavišo paspaudimo ir atleidimo:  $t_i^{pp} = t_i^p - t_i^r$ ;
- atleidimo – paspaudimo delsos laikas (angl. *Release -Press latency*) (RP)– tai laiko skirtumas tarp klavišo atleidimo ir sekančio klavišo paspaudimo:  $t_i^{rp} = t_i^p - t_{i-1}^r$ ;
- paspaudimo – paspaudimo delsos laikas (angl. *Press – Press latency*) (PP) – tai laiko skirtumas tarp  $t_{i-1}$  ir  $t_i$  klavišų paspaudimų laikų:  $t_i^{pp} = t_i^p - t_{i-1}^p$ ;
- atleidimo – atleidimo delsos laikas (angl. *Release – Release latency*) (RR) – tai laiko skirtumas tarp  $t_{i-1}$  ir  $t_i$  klavišų atleidimų laikų:  $t_i^{rr} = t_i^r - t_{i-1}^r$ ;
- paspaudimo – atleidimo delsos laikas (angl. *Press – Release latency*) (PR) – tai laiko skirtumas tarp  $t_{i-1}$  klavišo paspaudimo ir  $t_i$  klavišo atleidimo laikų:  $t_i^{pr} = t_i^r - t_{i-1}^p$ ;



1.1 pav. Normalizuotos klavišų paspaudimo dinamikos savybės

A. Morales'as taip pat tyrė, koki tikslumą galima pasekti su įvairiais klasifikatoriais, kai naudojamos skirtingos KD savybės. Tyrimo rezultatai pateikti 1.1 lentelėje. Tiksliausias identifikavimas buvo pasiektas naudojantis RP ir H klavišų dinamikos savybes. Tikslumui matuoti buvo naudojama lygus paklaidos santykis (angl. *Equal Error Rate*) (EER).

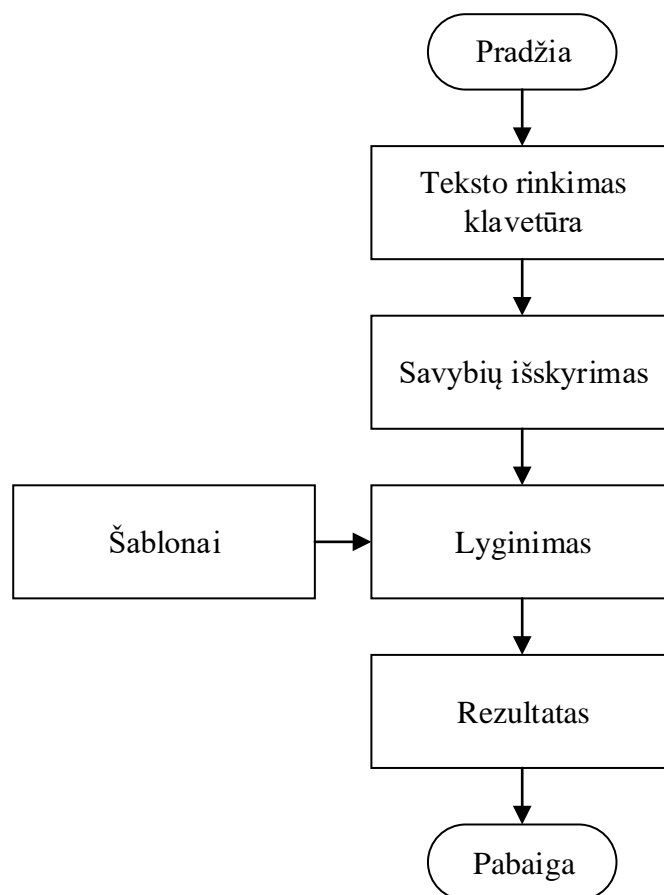
1.1 lentelė Savybių EER vidurkis[20]

KD savybė	Skaliuotas Manhattan atstumas	Apjungtas Manhattan-Mahalanobis atstumas	Mahalanobis + artimiausias kaimynas	Modifikuotas skaliuotas Manhattan atstumas
P - laikas	0.110	0.140	0.159	0.086
PA - delsa	0.063	0.130	0.121	0.048
AA - delsa	0.074	0.101	0.104	0.054
PP - delsa	0.069	0.116	0.110	0.045
AP - delsa	0.073	0.097	0.106	0.042
AP - delsa, P-laikas	0.044	0.097	0.089	0.023

### 1.3.2. Klavišų paspaudimo dinamikos sistemos

Biometrinė sistema yra automatinis procesas, kai naudojantis KD savybėmis galima identifikuoti subjektą ar nustatyti KD pokyčius. 1.2 paveiksle pavaizduotas KD sistemos veikimo principas. Nepriklausomai nuo to, ką bandoma nustatyti naudojantis klavišų dinamikos savybėmis, sistemos veikimo principinis nesikeičia:

1. Sistemai perduodama KD duomenys, kurie užfiksuojami asmeniui klaviatūra renkant tekstą;
2. KD duomenys normalizuojami ir išskiriamos savybės;
3. Išskirtos savybės lyginamos su sistemos turimais modeliais;
4. Pateikiamas rezultatas.



1.2 pav. Principinis klavišų dinamikos sistemos modelis [19]

Klavišų dinamikos sistemos skirstomos į dvi grupes: statines ir dinamines [18]:

**Statinė KD sistema** – sistema, kuri identifikacijai naudoja statinį tekstą. Pavyzdžiui prisijungimui prie sistemos naudojamas vartotojo vardas ar slaptažodis.

**Dinaminė KD sistema** – sistema, kuri identifikaciją vykdo visos vartotojo sesijos metu. Ši sistema identifikuota ir reaguoja į KD savybių pasikeitimas realiu laiku.

#### 1.4. Klavišų paspaudimo dinamikos modelių kūrimo metodai

Tam, kad KD sistema gebėtų identifikuoti subjektus ar pokyčius su toleruotina paklaida, reikalingas korektiškas šablonas. Šablonai gali naudoti įvairius klasifikatorius ar savaime apsimokančias sistemas. KD sistemų šablonų kūrimui dažniausiai naudojama dirbtiniai neuroniniai tinklai (angl. *Artificial Neural Network*) (ANN), atraminių vektorių mašina (angl. *Support Vector Machine*) (SVM), k-artimiausi kaimynai (angl. *K-Nearest Neighbors*) (KNN), paslėptieji Markovo modeliai (angl. *Hidden Markov Mmodel*).

##### 1.4.1. Dirbtinis Neuroninis tinklas

Ting-Yi Chan'as [21] tyrime analizavo KD savybių pritaikymą didinant apsaugą tradiciniam slaptažodžiui. Tyrimo metu įvedamam slaptažodžiui buvo išskiriamos KD savybės, kurios, kaip papildomas identifikatorius, buvo naudojamos papildomam vartotojo identifikavimui, jungiantis prie sistemos. Norit sudaryti korektišką vartotojo slaptažodžio šabloną, tyrimo dalyvių buvo paprašyta slaptažodį pakartoti penkis kartus. Normalizavus gautus duomenis buvo sudaryti keturi vektoriai:

$$t_i^{pr} = (t_1^{pr}, t_2^{pr}, \dots, t_q^{pr}),$$

$$t_i^{pp} = (t_1^{pp}, t_2^{pp}, \dots, t_q^{pp}),$$

$$t_i^{rp} = (t_1^{rp}, t_2^{rp}, \dots, t_{q-1}^{rp}),$$

$$t_i^{rr} = (t_1^{rr}, t_2^{rr}, \dots, t_{q-1}^{rr}),$$

kur  $q$  yra slaptažodį sudarančių simbolių skaičius.

Šablono sudarymui buvo sukonstruotas ANN kuris turėjo  $N = 4q - 3$  neuronų įėjimo sluoksnyje,  $M = q$  neuronus išėjimo sluoksnyje ir  $K = N + \sqrt{q}$  neuronus paslėptame sluoksnyje. Aktyvavimo funkcija buvo pasirinkta sigmoidinė funkcija.

Šablono patikimumo įvertinimui buvo naudojamas klaidingas patvirtinimo santykis (angl. *False Acceptance Rate*) (FAR), kuris nusako klaidingai patvirtintų vartotojų prisijungimų santykį su visais apgaulingais prisijungimais, ir teisingas atmetimo santykis (angl. *False Rejected Rate*) (FRR), kuris nusako klaidingai nepatvirtintų prisijungimų santykį su visais teisingų prisijungimų bandymais.

Teste dalyvavo dvidešimt trys asmens, kurių slaptažodžiams buvo sukurta KD šablonai. Kiekvieną slaptažodį bandė nulaužti 60 asmenų, vesdami jį dvidešimt kartų. Tyrimo rezultatai pateikti 1.2 lentelėje.

**1.2 lentelė** Ting-Yi Chan'o tyrime tirtų slaptažodžiu FRA ir FAR vidurkiai [21]

Slaptažodis	FAR (%)	FRR (%)
2ruxgrrx	9,42	0,0
2ugigggg	10,26	5,0
54jo3g4196aj4	3,72	1,0
123gemtek	0,72	7,0
11097	9,2	0,0
630404	72,42	0,0
97332005	19,57	0,0
d94321002	16,54	0,0
dago1013	8,3	0,0
f125812888	25,68	3,0
hwluchi	2,96	10,0
leon0708	6,21	10,00
mandylin	6,18	0,00
mux5605	2,16	10,05
sammi67423	6,72	10,00
sharelfu	4,26	0,00
still531	3,84	29,0
superdragon	0,01	15,0
thfn0821	1,68	8,5
ting0408	1,97	0,0
tmac520	11,83	11,50
yuju123	2,68	0,00
9oShiftARYjtq	0,00	0,00
.tie5ShiftRoanla	4,250	5,00
<b>Vidurkis:</b>	9,61	5,25

Gauti rezultatai parodė, jog FRR ir FAR priklauso nuo slaptažodžio simbolių parinkimo – slaptažodžiams sudarytiems iš skaičių, ANN šablono FAR santykis siekia 72 %.

#### 1.4.2. Atraminių vektorių mašina

Atraminių vektorių mašina taip pat dažnai naudojama KD sistemose. Alaa Darabseh'os komanda [22] atliko tyrimą, kuriame analizavo subjekto identifikavimo galimybes naudojant KD sistemą. Klavišų dinamikos identifikavimo šablonui kurti tyrime buvo naudojama SVM klasifikatorius.

Klasifikatoriaus mokymo imtį sudarė keturių tipų duomenys:

1. Dvidešimties dažniausiai pasikartojančių anglų kalbos raidžių paspaudimo laikas ( $e, a, r, i, o, t, n, s, h, d, l, c, u, m, w, f, g, y, p, b$ )

2. Dvidešimties dažniausiai pasikartojančių angliško alfabeto raidžių porų paspaudimo – paspaudimo delsa (*in, th, ti, on, an, he, at, er, re, nd, ha, en, to, it, ou, ea, hi, isor, te*)
3. Dvidešimties dažniausiai pasikartojančių angliško alfabeto raidžių porų atleidimo – paspaudimo delsa
4. Dvidešimties populiariausių angliškų žodžių parašymo laikas (*for, and, the, is, it, you, have, of, be, to, that, he, she, this, they, will, i, all, a, him*).

Maksimalus imties dydis, naudojantis visais duomenų tipais, neviršijo 80. Tyrime buvo naudojami įvairius paieškos algoritmai, kurie atrinkdavo svarbiausius imties elementus. Tyrime taip pat buvo vertinama imties dydžio įtaka SVM subjektų identifikavimo tikslumui. Tyrimo rezultatai pateikti 1.3 lentelėje.

**1.3 lentelė** SVM klasifikatoriaus tikslumas [22]

Imties atrinkimo metodas	Mokymo imtis	Tikslumas %	Imties sumažėjimas procentais %
Pilna imtis	80	94,44	-
GFS	30	85,77	62,5
GBE	73	92,54	0,8
BFS	44	91,66	45
BBS	74	92,78	0,7
GA	59	91,07	26,25
PSO	51	89,88	36,25

Iš gautų rezultatų matyti, kad apmokant klasifikatorių su maksimalia duomenų imtimi pasiekiamas 94 % subjekto atpažinimo tikslumas. Atrinkus imties elementus su geriausio tiesinio parinkimo algoritmu (angl. *Best Forward Selection*) (BFS), imtis sumažėja 45 %, tikslumas –2 %.

### 1.5. Klasifikatorių mokymui naudojamų savybių atrinkimas

Pokyčius stebimame procesuose lemia tam tikrų savybių ar įėjimų pasikeitimai. Šiuolaikine įranga galima fiksuoti daugybę parametrų, kurie vienaip ar kitaip lemia pasikeitimus stebimuose procesuose. Dažnai, pagal proceso pobūdį, ekspertai gali suranguoti įėjimo parametrus pagal jų svarbumą, tačiau didėjant matuojamų parametrų skaičiui, ekspertinis parametrų rangavimas tampa komplikotas.

Svarbiausių komponentų analizė (angl. *Principal Components Analysis*) (PCA) yra vienas iš dažniausiai naudojamų parametrų mažinimo metodų. PCA, turimai  $n$  parametrų imčiai, ieško tiesinės  $d$  – matės erdvės, kur  $d < n$ .

Tiesinė erdvė gali būti apibūdinama  $d$  ortogonalinių vektorių, kurie sudaro naują svarbiausių komponentų (angl. *Principal Components*) koordinačių sistemą [23]. Hotelling'as PCA apibūdino kaip imties vektorių  $x_i$ ,  $i \in 1 \dots t$ ,  $d$  svarbiausių ašių, kurių kryptimi yra fiksuojamas didžiausias išsibarstymas [24].

## 2. TEORINIS METODŲ PAGRINDIMAS

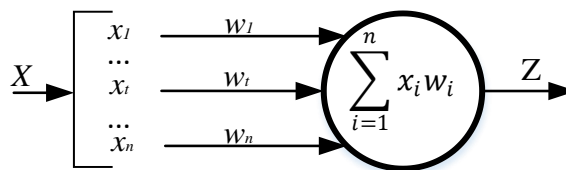
Mus supančioje aplinkoje egzistuoja begalės užduočių, kurias kompiuteris paprastai išsprendžia, tačiau žmogui tai padaryti gana sudėtinga, pavyzdžiui: daugyba, dalyba, kėlimas kvadratu ar kitos matematinės operacijos. Kitą vertus, tokios užduotys kaip objekto nustatymas erdvėje, objektų identifikavimas, erdvės suvokimas yra paprastos žmogui, bet sudėtingos kompiuteriui. Kiekvienas žmogus lengvai gali nustatyti objekto padėtį erdvėje, identifikuoti nuotraukų turinį. Buvo sukurtas matematinis skaičiavimo modeliai, kurie suteikė galimybę kompiuteriui atlikti žmogui lengvai suprantamas užduotis.

### 2.1. Dirbtinis neuroninis tinklas

Pirmą dirbtinio neuroninio tinklo modelį 1943 metais suformavo W.S. McCulloch'as ir W. Pitts'as. Pagal jų pasiūlytą modelį, smegenys yra sudaryti iš daug neuronų, kurie turi vienodą struktūrą. Neuroną galima išivaizduoti kaip slenkstinę funkciją – tik viršijus slenkstinę vertę, suformuojamas išėjimo signalas.

#### 2.1.1. Matematinis neurono modelis

Dirbtinis neuronas imituota biologinio neurono veikimo principą. Neurono įėjimas yra  $X$  vektorius, kurį mes gauname kaip pradinius duomenis arba, kaip kitų neuronų išėjimus. Kiekviena  $X$  vektoriaus komponentė  $x_i$  dauginama iš svorio reikšmės  $w_i$ . Visų sandaugų suma lemia neurono aktyvavimo lygį. Dirbtinio neurono modelis parodytas 2.1 paveiksle. Įėjimo vektorius  $X = (x_1, x_2, \dots, x_n)$  dauginamas iš atitinkamo įėjimo svorio koeficiento  $W(w_1, w_2, \dots, w_n)$  ir visos gaunamos sandaugos sudedamos.

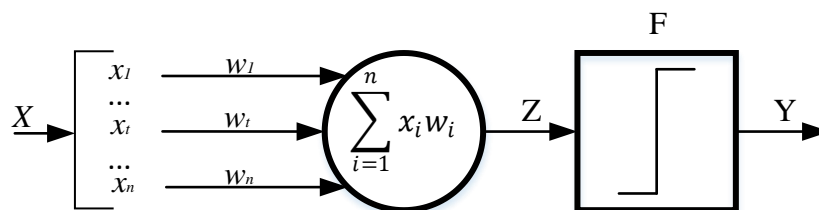


2.1 pav. Dirbtinio neurono modelis [25]

Įėjimo vektorius  $X = (x_1, x_2, \dots, x_n)$  dauginamas iš atitinkamo įėjimo svorio koeficiento  $W(w_1, w_2, \dots, w_n)$  ir visos gaunamos sandaugos sudedamos.

#### 2.1.2. Aktyvavimo funkcijos

Pagal sužadinimo signalą  $Z$ , naudojant aktyvavimo funkciją  $F$ , gaunamas neurono išėjimo signalas  $Y$  (2.2 pav.).



2.2 pav. Dirbtinio neurono modelis su aktyvavimo funkcija [25]

Dažniausiai naudojami aktyvavimo funkcijos:

- slenkstinė funkcija (2.1) (2.3 pav., b),
- sigmoidinė funkcija (2.2) (2.3 pav., a),
- hiperbolinis tangentas (2.3) (2.3 pav., c),
- Tiesinė funkcija (2.4) (2.3 pav., d).

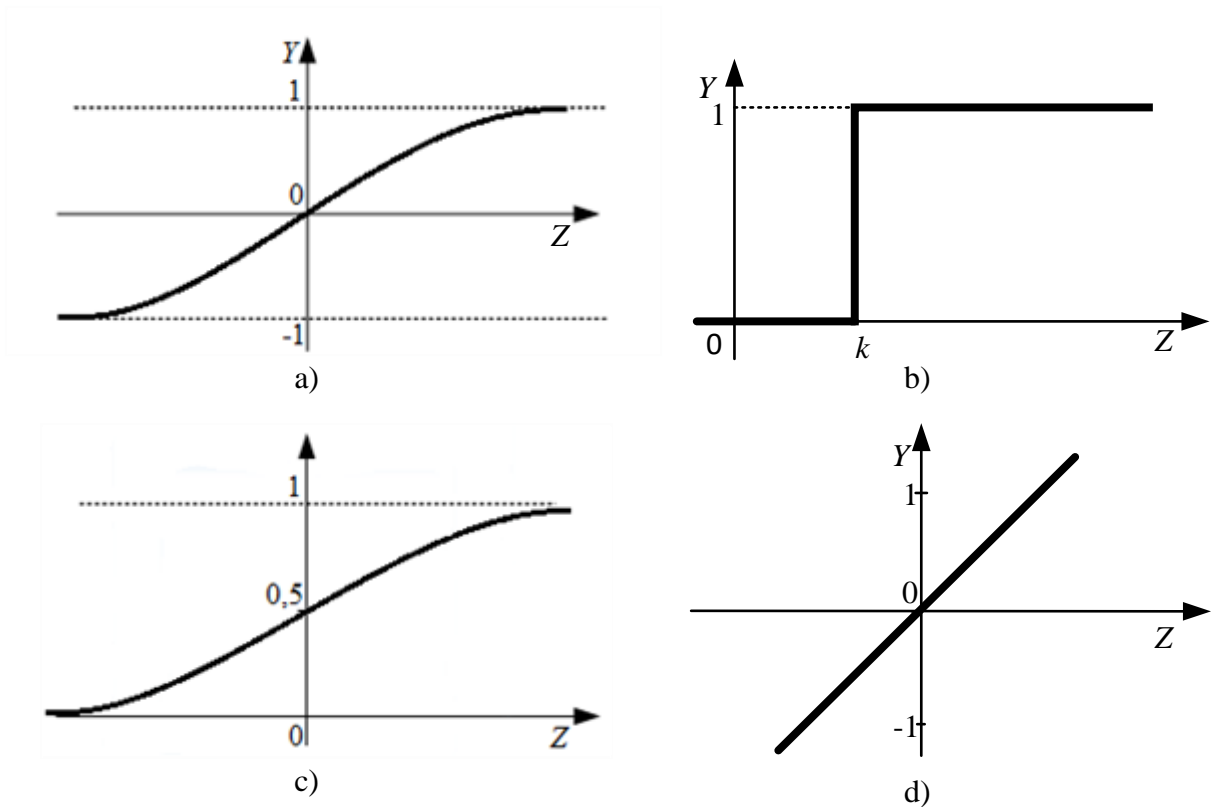
$$Y = F(Z) = \begin{cases} 1, & \text{jeigu } Z \geq k; \\ 0, & \text{jeigu } Z < k; \end{cases} \quad (2.1)$$



$$Y = F(Z) = \frac{1}{1 + e^{-Z}} \quad (2.2)$$

$$Y = F(Z) = \frac{e^Z - e^{-Z}}{e^Z + e^{-Z}} \quad (2.3)$$

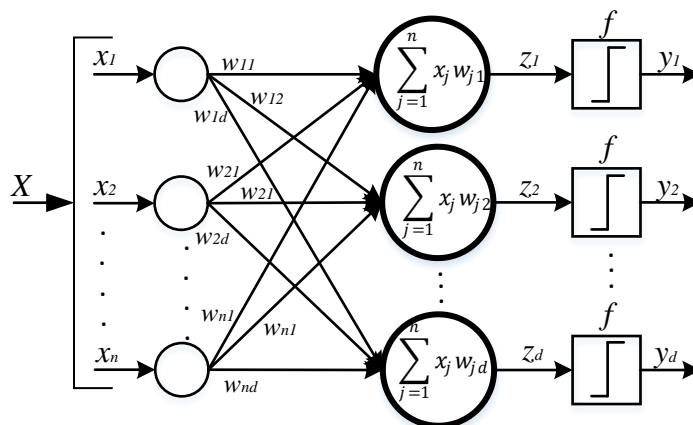
$$Y = F(Z) = kZ \quad (2.4)$$



2.3 pav. Aktyvavimo funkcijos: a) sigmoidinė, b) slenkstinė, c) hiperbolinis tangentas, d) tiesinė [25]

### 2.1.3. Vienasluoksnis neuroninis tinklas

Teoriškai vienas dirbtinis neuronas gali atlikti paprastus klasifikavimo uždavinius, bet sujungus neuronus į tinklą, klasifikacijos galimybes didėja. Paprasčiausias tinklas, kurį sudaro vienas sluoksnis neuronų, vadinamas perceptronu (2.4 pav.). Kairės pusės apskritimai yra įėjimo signalų skirstytuvai, nevykdantys jokių skaičiavimų. Kiekvienas įėjimo sluoksnio X elementas turi atitinkamą svorio koeficientą  $w_{jk}$ , kur  $jk$  nusako ryšį tarp j-tojo įėjimo ir k-tojo neurono.



2.4 pav. Vienasluoksnis neuroninis tinklas [25]

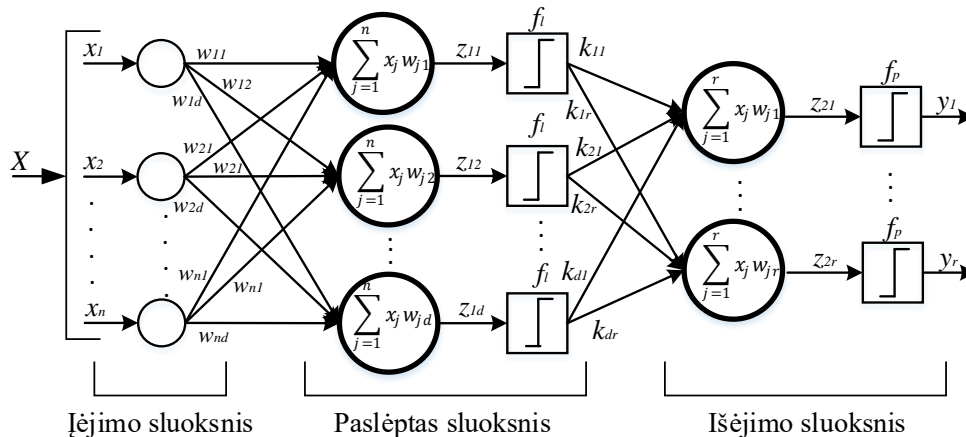
Kiekvienas išėjimas  $y_j, j = 1, \dots, d$ , yra įėjimo vektoriaus  $X = (x_1, x_2, \dots, x_n)$  funkcija, kuri apskaičiuojama pagal formulę:

$$y_j = f(z_j) = f\left(\sum_{k=0}^n x_{jk} w_k\right), \quad (2.5)$$

čia  $d$  – neuroninio tinklo išėjimų skaičius,  $f$  – aktyvavimo funkcija (2.1, 2.2, 2.3, 2.4).

#### 2.1.4. Daugiasluoksnis neuroninis tinklas

Neuroninis tinklas turintis daugiau nei vieną neuronų sluoksnį ir, kuriame nėra grįžtamojo ryšio, vadinamas daugiasluoksniu tiesioginio sklidimo neuroniniu tinklu (angl. *multilayer feedforward neural networks*) (2.5 pav.).



2.5 pav. Daugiasluoksnis neuroninis tinklas [25]

2.5 paveiksle  $w_{ij}$  yra pirmojo neurono sluoksnio svorių koeficientai,  $k_{ij}$  yra antrojo neuronų sluoksnio svorių koeficientai. Taip pat skirtingiems neuronų sluoksniams galima taikyti skirtingas aktyvavimo funkcija.

Manykime, kad turime daugiasluoksnį neuroninį tinklą, kuriame yra  $L$  sluoksnių,  $l = 0, 1, \dots, L$ , kur  $l = 0$  yra įėjimų sluoksnis o  $l = L$  – išėjimų sluoksnis. Kiekvienas  $l$  sluoksnis turi  $n_l$  neuronų. Pirmoko sluoksnio išėjimo reikšmės  $y_j$  yra apskaičiuojamos pagal (2.5) formulę.  $l$  – ojo sluoksnio įėjimai į neuronus yra  $(l - 1)$  – ojo sluoksnio išėjimai. Kiekvieno  $j$  – ojo neurono išėjimo reikšmė  $y_j$   $l$  – jame sluoksnyje yra apskaičiuojama:

$$y_j = f(z_j) = f\left(\sum_{k=0}^{n_{l-1}} w_{jk} y_k\right), j = 1, \dots, n_l \quad (2.6)$$

#### 2.1.5. Neuroninio tinklo mokymas

Neuroninio tinklo mokomas tam, jog būtų minimizuoti neatitikimai tarp neuroninio tinklo ir realių rezultatų. Neuroninis tinklas gali būti mokomas trejopai:

- mokomas su norimomis išėjimo reikšmėmis;
- mokomas be išėjimo reikšmių;
- mišrus mokymas.

Mokant su norimomis išėjimo reikšmėmis, įėjimo vektorius  $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$  susiejamas su norimomis išėjimo reikšmėmis  $T_i = (t_{i1}, t_{i2}, \dots, t_{id})$   $i = 1, \dots, h$ . Tinklo išėjimo reikšmės, skaičiuojamos kiekvienam įėjimo vektoriui  $X_i$ , yra tiesiogiai susijusios su norimomis tų išėjimų reikšmėmis. Tinklas koreguojamas keičiant svorių vektorių reikšmes ir siekiant gauti kiek galima mažesnę paklaidą, t.y. ieškoma tokių svorių, kad skirtumas tarp norimų išėjimo reikšmių ir reikšmių gautų išmokius neuroninį tinklą būtų kiek galima mažesnės.

Kartais norimos gauti tinklo reikšmės nėra žinomos, tinklas mokomas be išėjimo reikšmių. Šio tipo metoduose tinklas mokomas ieškoti koreliacijų ar panašumų tarp mokymo aibės įėjimų. Čia nėra

grįžtamojo ryšio, pasakančio, kuris atsakymas yra teisingas. Turima tik mokymo vektorių aibė  $X_i$ . Metodo tikslas suskirstyti mokymo duomenis į tam tikras klases arba rasti juose kokius nors reguliarumus arba ypatumus. Mišrus mokymas apima mokymą su norimais duomenimis ir be jų.

Mokant vienasluoksnį tinklą, svorių koeficientai keičiami taip, jog išėjimo vektorius  $Y_i = (y_{i1}, y_{i2}, \dots, y_{in})$ , gautas į įėjimą pateikus vektorių  $X_i$ , būtų kiek galima artimesnis vektoriui  $T_i$ , t.y. tinklo veikimo paklaida būtų kiek įmano mažesnė.

Paklaida  $E(W)$  apibrėžiama kaip, svorių matricos  $W = \{w_{ij}, i = 1, \dots, d, j = 0, \dots, n\}$ , funkcija. Paklaidos funkcijos  $E(W)$  minimumas randamas gradientiniais optimizavimo metodais. Dažniausiai naudojamas gradientinis nusileidimo algoritmas, o paklaida apskaičiuojama taikant kvadratinių nuokrypių sumą:

$$E_i(W) = \frac{1}{2} \sum_{j=1}^d (y_{ij} - t_{ij})^2. \quad (2.7)$$

Mokant daugiasluoksnį neuroninį tinklą reikia žinoti paslėptų neuronų išėjimų reikšmes. Jei paklaidos ir aktyvacijos funkcijos yra diferencijuojamos, ieškant minimalios paklaidos gali būti naudojama gradientinio nusileidimo metodas. Metodas, kuris naudoja gradientinį nusileidimą tinklo mokymui vadinamas „klaidos sklidimo atgal“ algoritmu (angl. *back-propagation learning algorithm*).

Šis algoritmas tinklą moko su norimomis išėjimų reikšmėmis, paklaida  $E_i(W)$  apskaičiuojama pagal (2.7) formulę o svoriai keičiami pagal (2.8).

$$\Delta w_{jk}^i(t) = -\eta \frac{\partial E_i(t)}{\partial w_{jk}}, \quad (2.8)$$

čia  $\eta$  yra teigiamas daugiklis (mokymo greitis), kuriuo reguliuojamas gradientinio optimizavimo žingsnio ilgis,  $t$  – iteracijos numeris. Pirmame algoritmo žingsnyje įėjimų vektoriui  $X_i$  apskaičiuojama išėjimų vektorius  $Y_i$ . Įvertinama paklaidos funkcija  $E_i(W)$  išėjimų sluoksnyje  $L$ . Jei paklaidos funkcija  $E_i(W)$  nelygi nuliui, jungčių svoriai keičiami pagal (2.8) formulę.

Dalinės išvestinės išreiškiamos formule:

$$\frac{\partial E_i}{\partial w_{jk}} = \frac{\partial E_i}{\partial z_{jk}} * \frac{\partial z_i}{\partial w_{jk}} \quad (2.9)$$

Iš (2.6) gaunama

$$\frac{\partial z_i}{\partial w_{jk}} = y_{ik}. \quad (2.10)$$

Tegul

$$\delta_{ij} = \frac{\partial E_i}{\partial z_{ij}} \quad (2.11)$$

Įstatę (2.11) ir (2.10) į (2.9) ir (2.8) formules gauname:

$$\Delta w_{jk}^i = -\eta \delta_{ij} y_{ik}, \quad (2.12)$$

čia  $j$ -asis neuronas priklauso  $l$  –ajam o  $k$  –asis neuronas priklauso  $(l - 1)$ -ajam sluoksniui. Išėjimų sluoksnyje

$$\delta_{ij} = \frac{\partial E_i}{\partial z_{ij}} = f'(a_{ij})(y_{ij} - t_{ij}) \quad (2.13)$$

Kur  $j$  – asis neuronas priklauso  $L$  – ajam išėjimų sluoksniui.

Naudojant dalinės išvestines  $\frac{\partial E_i}{\partial w_{jk}}$  apibendrinę gauname :

$$\delta_{ij} = \frac{\partial E_i}{\partial w_{ij}} = \sum_{s=1}^{n_{l+1}} \frac{\partial E_i}{\partial z_{is}} * \frac{\partial z_{is}}{\partial z_{ij}} \quad (2.14)$$

čia  $n_{l+1}$  žymi neuronų  $(l + 1)$  –ajame sluoksnyje skaičių. Išraiška  $\frac{\partial E_i}{\partial z_{is}}$  yra lygi dydžiui  $\delta_{ij}$ , apibrėžtam s-ajam neuronui  $(l + 1)$  –ajame sluoksnyje. Atsižvelgiant į (2.6) gauname:

$$\frac{\partial z_{is}}{\partial z_{ij}} = f'(a_{ij})w_{is} \quad (2.15)$$

Tada paslėptųjų  $j$  –ųjų neuronų

$$\delta_{ij} = f'(a_{ij}) \sum_{s=1}^{n_{l+1}} w_{is} \delta_{is} \quad (2.16)$$

čia  $j$ -asis neuronas priklauso sluoksniui  $l < L$ ,  $s$ -asis neuronas priklauso sluoksniui  $l + 1$ .

Iš pradžių apskaičiuojamas  $\delta_{ij}$  reikšmės išėjimų sluoksnyje  $L$  pagal (2.14). Tada pagal (2.16) palaipsniui skaičiuojamos  $\delta_{ij}$  reikšmės tarpiniuose sluoksniuose. Visi svoriai keičiami pagal (2.12). Tada į tinklą paduodamas naujas vektorius  $X_i$  ir mokymo algoritmas kartojamas iš naujo Algoritmo sustojimo kriterijus yra arba iš anksto nustatyta paklaidos funkcijos slenksčio reikšmė, arba atitinkamas atliktų iteracijų (mokymo žingsnių) skaičius [26].

## 2.2. Atraminių vektorių mašinos

Atraminių vektorių mašinos yra vienas iš dažniausiai taikomų kvalifikavimo metodų. SVM metodas buvo pasiūlytas Vladimir'o N.Vapnik'o 1995 metais [27]. Šis metodas remiasi idėja, kad duomenis esančius  $n$  –matėje erdvėje, galima atskirti  $n$ -mate plokštuma.

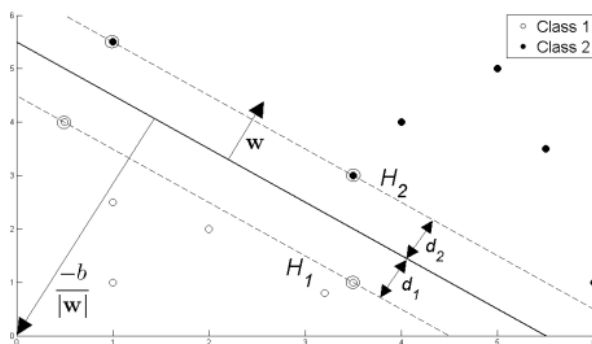
### 2.2.1. Tiesinis binarinis klasifikatorius

Tarkime, kad turime  $L$  mokymo taškų imtį, kur kiekvienas  $x_i$  įėjimas turi  $D$  atributų ir priklauso vienai iš dviejų klasių  $y_i = -1$  arba  $y_i = 1$ . Tokią mokymo imtį galime aprašyti kaip  $\{x_i, y_i\}$ , kur  $i = 1 \dots L, y_i \in \{-1; 1\}, x \in R^D$ .

Priėmus, kad imtį  $L$  sudaro dvi, tiesiškai atskiriamos klasės, jas galime padalinti su tiese, kai  $D = 2$ , arba daugiamate plokštume, kai  $D > 2$ . Tokia daugiamatė plokštuma gali būti aprašoma lygtimi

$$w * x + b = 0, \quad (2.17)$$

ur  $w$  yra daugiamatės plokštumos normalė,  $\frac{b}{\|w\|}$  yra atstumas nuo koordinačių pradžios iki skiriamosios plokštumos (2.6 pav.)



2.6 pav. Hiperbolinė plokštuma su dviem tiesiškai atskirtomis klasėmis[28]

SVM tikslas – taip parinkti  $w$  ir  $b$  reikšmės, kad skiriamoji plokštuma būtų labiausiai nutolusi nuo abiejų klasių. Tokiu atveju mokymo imtis  $L$  aprašoma:

$$x_i * w + b \geq +1, \text{ kai } y_i = +1 \text{ ir} \quad (2.18)$$

$$x_i * w + b \leq -1, \text{ kai } y_i = -1, \quad (2.19)$$

arba bendra išraiška:

$$y_i(x_i * w + b) - 1 \geq 0 \forall_i. \quad (2.20)$$

Jeigu priimame prielaidą, kad atraminiai vektoriai yra taškai, kurie yra arčiausiai skiriamosios plokštumos, tada  $H_1$  plokštumą, ant kurios yra šie taškai, galima aprašyti:

$$x_i * w + b = +1, \quad (2.21)$$

$H_2$  plokštuma aprašoma:

$$x_i * w + b = -1. \quad (2.22)$$

Maksimalaus atstumo, tarp plokštumos ir atraminių vektorių radimui, ieškomas didžiausias atstumas  $r = d_1 + d_2$ . Remiantis vektorių teorija, atstumas yra  $r$  yra lygus

$$r = \frac{1}{\|w\|}. \quad (2.23)$$

$r$  maksimizavimui, reikia rasti mažiausią  $w$  vertę tenkinančią 2.20 išraišką:

$$\min \|w\| = \frac{1}{2} * \|w\|^2, \text{ kai } y_i(x_i * w + b) - 1 \geq 0 \forall_i \quad (2.24)$$

Šiam veiksmui atlikti ieškomi Lagrandžo daugikliai  $\alpha$ , kur  $\alpha_i \geq \forall_i$ :

$$L_P \equiv \frac{1}{2} \|w\|^2 - \alpha [y_i(x_i * w + b) - 1 \forall_i] \quad (2.25)$$

$$\equiv \frac{1}{2} \|w\|^2 - \sum_{i=1}^L \alpha_i [y_i(x_i * w + b) - 1] \quad (2.26)$$

$$\equiv \frac{1}{2} \|w\|^2 - \sum_{i=1}^L \alpha_i y_i(x_i * w + b) + \sum_{i=1}^L \alpha_i \quad (2.27)$$

$w$  ir  $b$  reikšmių, kurios minimizuoja ir  $\alpha$  – kuri maksimizuoja (2.17) išraišką, suradimui, pagal  $w$  ir  $b$  kintamuosius, ieškomos  $L_P$  dalines išvestines:

$$\frac{\partial L_P}{\partial w} = 0 \Rightarrow w = \sum_{i=1}^L \alpha_i y_i x_i \quad (2.28)$$

$$\frac{\partial L_P}{\partial b} = 0 \Rightarrow \sum_{i=1}^L \alpha_i y_i = 0 \quad (2.29)$$

Įstačius 2.28 ir 2.29 į 2.27 gauname naują išraišką, kuri priklauso tik nuo  $\alpha$  parametro:

$$L_D \equiv \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i x_j, \text{ kai } \alpha_i \geq 0 \forall_i, \sum_{i=1}^L \alpha_i y_i \quad (2.30)$$

$$\equiv \sum_{i=1}^L \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i H_{ij} \alpha_j, \text{ kur } H_{ij} \equiv y_i y_j x_i x_j \quad (2.31)$$

$$\equiv \sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T H \alpha, \text{ kai } \alpha_i \geq 0 \forall_i, \sum_{i=1}^L \alpha_i y_i. \quad (2.32)$$

Maksimalaus  $L_D$  išvestinės dydžio radimui, ieškomas didžiausia  $\alpha$  vertė:

$$\max_{\alpha} \left[ \sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T H \alpha \right], \text{ kai } \alpha_i \geq 0 \forall_i, \sum_{i=1}^L \alpha_i y_i \quad (2.33)$$

Apskaičiuavus  $\alpha$  iš 2.28 lygties gaunama  $w$ .

Bet kuris, atraminis vektorius  $x_s$ , kuris tenkina 2.29 turi

$$y_s(x_s * w + b) = 1 \quad (2.34)$$

formą, kur  $w$  pakeitus 2.28 išraiška, gauname:

$$y_s \left( \sum_{m \in S} \alpha_m y_m x_m * x_s + b \right) = 1. \quad (2.35)$$

Čia  $S$  yra kintamųjų aibė, kuri apibūdinanti atraminius vektorius.  $w$  ir  $b$  reikšmių gavimui, kurios apibūdina daugiamatį skiriamąjį paviršių, apskaičiuojamas  $x_s$  imties vidurkis:

$$b = \frac{1}{N_s} \sum_{s \in S} \left( y_s - \sum_{m \in S} \alpha_m y_m x_m * x_s \right) \quad (2.36)$$

### 2.2.2. Binarinis klasifikatorius netiesiškai atskiriamiems duomenims

Tam, kad SVM metodas galėtų atskirti duomenis, kurie nėra pilnai tiesiškai atskiriami, 2.18 ir 2.19 išraiškas reikia papildyti teigiamu skverbties (angl. *slack*) kintamuoju  $\xi_i$ , kur  $i = 1 \dots L$ :

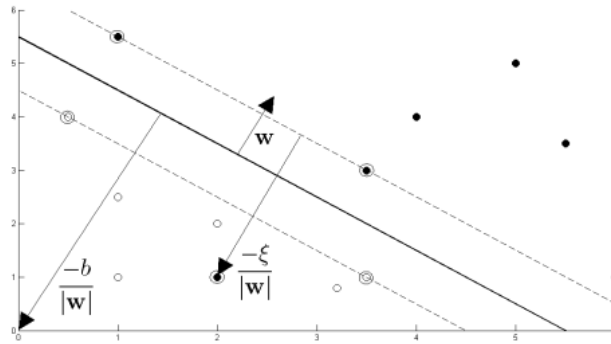
$$x_i * w + b \geq +1 - \xi_i, \text{ kai } y_i = +1 \text{ ir} \quad (2.37)$$

$$x_i * w + b \leq -1 + \xi_i, \text{ kai } y_i = -1 \quad (2.38)$$

Abi išraiškos gal būti išreikštos bendra išraiška:

$$y_i(x_i * w + b) - 1 + \xi_i \geq 0 \forall_i. \quad (2.39)$$

Dvimatis, netiesiškai atskiriamų, duomenų aibės atskyrimas, pateiktas 2.7 paveiksle.



2.7 pav. daugiamatė plokštuma su dviem, netiesiškai atskiriamomis klasėmis [28]

Duomenys, kurie yra įsiterpę į kitos klasės erdvę, turi nuobaudą (angl. *Penalty*), kuri priklauso nuo to, kaip giliai taškas įsiskverbęs į ją. Nesuklasifikuotų duomenų kiekio mažinimui 2.24 formulę papildome skverbties kintamuoju  $\xi_i$ :

$$\min ||w|| = \frac{1}{2} * ||w||^2 + C \sum_{i=1}^L \xi_i, \text{ kai } y_i(x_i * w + b) - 1 + \xi_i \geq 0 \forall_i. \quad (2.40)$$

Čia  $C$  yra nuobauda, kuri nusako santykį tarp skverbties kintamojo ir skiriamosios juostos. Šio atveju, Lagrandžo daugiklių paieškai naudojama modifikuota 2.27 formulė:

$$L_P \equiv \frac{1}{2} ||w||^2 + C \sum_{i=1}^L \xi_i - \sum_{i=1}^L \alpha_i [y_i(x_i * w + b) - 1 + \xi_i] + \sum_{i=1}^L \mu_i \xi_i, \quad (2.41)$$

kuriai ieškome dalinių išvestinių pagal  $w$  (2.28),  $b$  (2.29; 2.28; 2.29; 2.29) ir  $\xi$ :

$$\frac{\partial L_p}{\partial \xi} = 0 \Rightarrow c = \alpha_i + \mu_i \quad (2.42)$$

Gautas dalines išvestines įstačius į 2.6 formulę gaunama 2.14 išraiška.

$$\max_{\alpha} \left[ \sum_{i=1}^L \alpha_i - \frac{1}{2} \alpha^T H \alpha \right], \text{ kai } 0 \leq \alpha_i \leq C \ \forall_i \text{ ir, } \sum_{i=1}^L \alpha_i y_i = 0 \quad (2.43)$$

$b$  koeficientas apskaičiuojamas pagal 2.36 formulę, kai  $0 < \alpha_i \leq C$ .

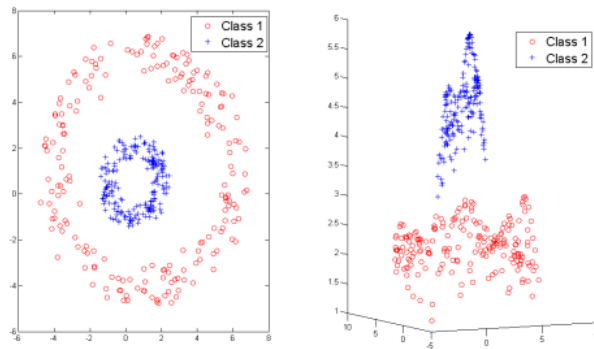
### 2.2.3. Netiesinės atraminės vektorių mašinos

Atraminų vektorių mašinos pritaikymui klasifikuoti netiesiškai atskiriamus duomenims, sudaroma duomenų imties matrica  $H$ :

$$H_{ij} = y_i y_j k(x_i, x_j) = x_i * x_j = x_j^T x_i. \quad (2.44)$$

Čia  $k(x_i, x_j)$  yra funkcijų branduolys (*angl. Kernel Functions*). Branduolio funkcijų rinkinį sudaro vidiniai vektorių porų kintamieji (2.45). Jeigu, naudojantis netiesinių funkcijų susiejimo savybėmis  $x \rightarrow \phi(x)$ , funkcija gali būti pervedama į aukštesnę matavimo erdvę, kurioje reikia nurodyti tik susietų įėjimų vidinius kintamuosius.

Branduolio funkcijos naudojamos tada, kai klasifikavimo problemos yra tiesiškai neatskiriamos erdvėje su įėjimais  $x$ , bet gali būti atskiriamos aukštesnėje erdvėje su tinkamu  $x \rightarrow \phi(x)$  susiejimu.



**2.8 pav.** Imties perorientavimas į aukštesnę erdvę naudojant spindulio paretą branduolio funkciją [28]

2.8 paveiksle pateiktų duomenų imtis dvimatėje erdvėje nėra tiesiškai atskiriama (kairys grafikas), tačiau tiesiškai atskiriamos aukštesnėje erdvėje, kuri yra apibrėžiama spindulio paretą branduolio funkcija (*angl. Radial Basis Kernel*)(RBK):

$$k(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2}. \quad (2.45)$$

Kitos populiarios branduolio funkcijos:

Polinominio branduolio (*angl. Polynomial Kernel*):

$$k(x_i, x_j) = (x_i * x_j + a)^b. \quad (2.46)$$

Sigmoidinio branduolio (*angl. Sigmoidal Kernel*):

$$k(x_i, x_j) = \tanh(ax_i * x_j - b), \quad (2.47)$$

kur  $a$  ir  $b$  yra parametrai apibūdinantys branduolio savybes [28].

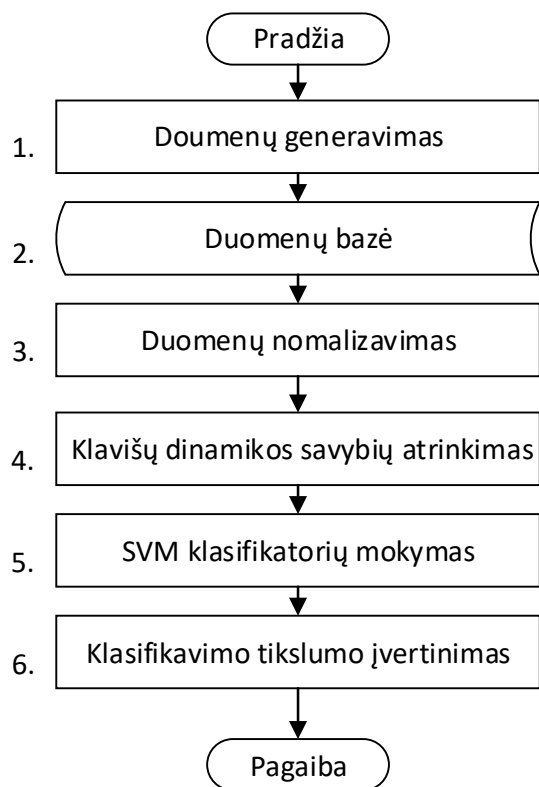
### 3. REALIZACIJA

Šiame skyriuje aptariama eiga, kuri reikalinga tyrimui atlikimui ir rezultatų gavimui. Skyriuje taip pat įvardijami sistemos funkciniai, bei nefunkciniai reikalavimai, darbo kokybės kriterijai. Aprašoma kompiuterio klaviatūros klavišų dinamikos sekimui realizuota PĮ ir duomenų bazės struktūra reikalinga duomenų saugojimui.

#### 3.1. Projekto planas

Žmogaus nuovargio nustatymo, pagal kompiuterio klavišų dinamikas, tyrimui atlikti reikalinga duomenų imtis. Tyrimo duomenų imtis sukuriama realizuotoje PĮ, tyrimo dalyviams, renkant standartinį tekstą. Prieš atliekant duomenų klasifikavimo uždavinį, duomenys normalizuojami, išskiriamos KD savybės, nustatomi savybių svarbumą duomenų imčiai ir sukuriama SVM klasifikatoriaus mokymo imtis. Apibendrinti tyrimo etapai pateikti tyrimo atlikimo schemeje (3.1 pav.). Išskiriami šeši etapai, kurie reikalingi tyrimo atlikimui:

- klavišų dinamikos fiksavimo PĮ realizavimas;
- KD duomenų saugojimas duomenų bazėje;
- duomenų normalizavimas programine įranga;
- KD savybių atrinkimas;
- SVM klasifikatorių mokymas;
- SVM nuovargio klasių klasifikavimo tikslumo įvertinimas.



3.1 pav. Nuovargio nustatymo tyrimo schema

#### 3.2. Funkciniai reikalavimai

Šio tyrimo realizavimui keliami šie funkciniai reikalavimai:

- vartotojas gali įvesti tekstą, kurį jam perrašant programinė įranga seks klaviatūros klavišų dinamikos duomenis;
- klavišų dinamikos duomenys turi būti talpinami *Microsoft SQL Server* serveryje;
- išsaugoti vartotojo įvedamo teksto rinkimo stiliaus klavišų dinamikos duomenis;
- vartotojas gali pasirinkti paros laiką (rytas, diena, vakaras) kada vedamas tekstas;
- vartotojas gali pasirinkti klasifikavimo imties dydį;



- vartotojas gali pasirinkti, duomenų klasifikavimo tipą.

### 3.3. Nefunkciniai reikalavimai

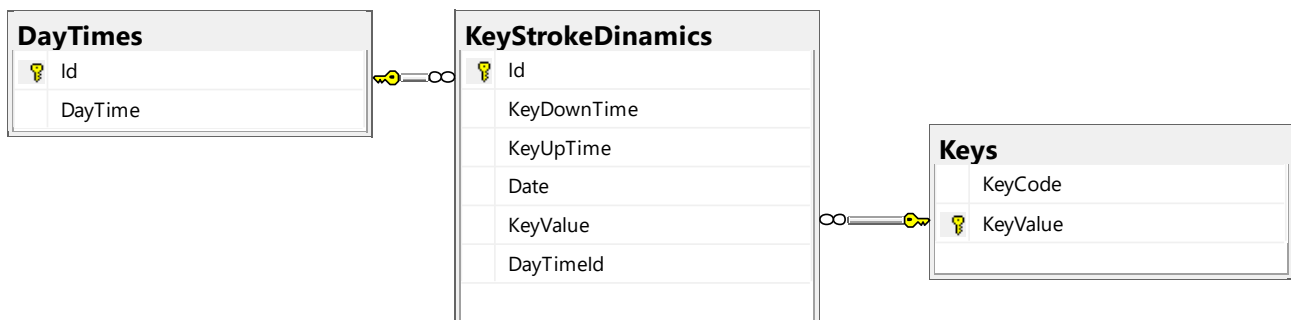
Šio tyrimo realizavimui keliami šie nefunkciniai reikalavimai:

- programinė įranga turi veikti su *Windows* operacine sistema;
- PĮ treniravimo ir testavimo rezultatus pateikia pulto lange.

### 3.4. Projekto duomenų bazė

Tyrimo naudojamos duomenų bazės schema pateikta 3.2 paveiksle. Ją sudaro trys lentelės:

- „DayTimes“ – lentelėje yra talpinama paros laikas (3.1 lent.);
- „KeyStrokeDinamics“ – lentelėje saugo klaviatūros klavišų dinaminis duomenis (3.2 lent.)
- „Keys“ – lentelėje talpinama klaviatūros klavišai ir jų skaitiniai atitikmenys (3.3 lent.);



3.2 pav. Duomenų bazės schema

Toliau pateikiamos lentelės, kuriose aprašomi duomenų bazės schemą sudarančių lentelių laukai.

#### 3.1 lentelė „DayTimes“ lentelės laukų paaiškinimai

Lauko pavadinimas	Lauko tipas	Lauko tipo atitikmuo <i>MS SQL Server</i> serveryje	Paskirtis
Id	Sveikas skaičius	integer	Įrašo identifikatorius
DayTimes	Tekstas	nvarchar(25)	Paros laikas

#### 3.2 lentelė „KeyStrokeDinamics“ lentelės laukų paaiškinimai

Lauko pavadinimas	Lauko tipas	Lauko tipo atitikmuo <i>MS SQL Server</i> serveryje	Paskirtis
Id	Sveikas skaičius	integer	Įrašo identifikatorius
KeyDownTime	Laikas	time	Klaviatūros klavišo paspaudimo laikas. Pateikiamas „HH:MM:SS“ formatu, $10^{-7}$ sekundės tikslumu
KeyUpTime	Laikas	time	Klaviatūros klavišo atleidimo laikas. Pateikiamas „HH:MM:SS“ formatu, $10^{-7}$ sekundės tikslumu
Date	Laikas	date	Įrašo generacijos data, kurios formatas: "YYYY:MM:DD"
KeyValue	Sveikas skaičius	integer	Klaviatūros klavišo numeris, kuris įrašą susieja su lentelės „Keys“ atitikmeniu.
DateTimeId	Sveikas skaičius	integer	Paros laiko identifikatorius, kuris įrašą susieja su lentelės „DayTime“ atitikmeniu.

### 3.3 lentelė „Keys“ lentelės laukų paaiškinimai

Lauko pavadinimas	Lauko tipas	Lauko tipo atitikmuo MS SQL Server serveryje	Paskirtis
KeyValue	Sveikas skaičius	integer	Kompiuterio klaviatūros klavišo numeris
KeyCode	Tekstas	nvarchar(25)	Kompiuterio klaviatūros klavišo simbolis

### 3.5. Duomenų gavyba

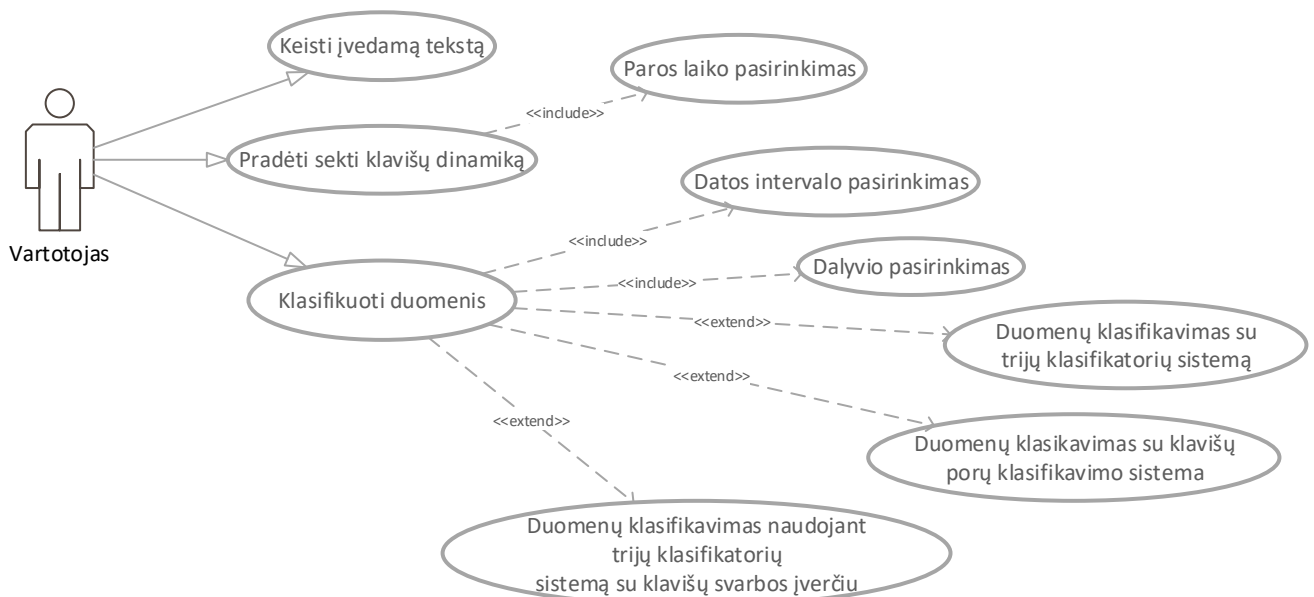
Eksperimentui atlikti reikalinga skirtingų klasių duomenys, kur kiekviena klasė reprezentuotų tam tikrą asmens nuovargį. Atliekant tyrimą buvo priimti, kad paros laikas koreliuoja su nuovargiu. Remiantis šia prielaida paros laikas suskirstomas į tris klases:

- *rytas* – asmuo pailsėjęs, nuovargio nėra;
- *diena* – asmuo yra po 8 valandų darbo dienos, jaučiamas vidutinis nuovargis;
- *vakaras* – asmuo pavargęs.

Atitinkamu paros metu (ryte, dieną ar vakare) asmuo, naudodamasis programine įranga, suveda standartinį tekstą. Teksto vedimo metu programinė įranga seka kompiuterio klaviatūros klavišų dinamiką, kuri vėliau išsaugoma duomenų bazėje.

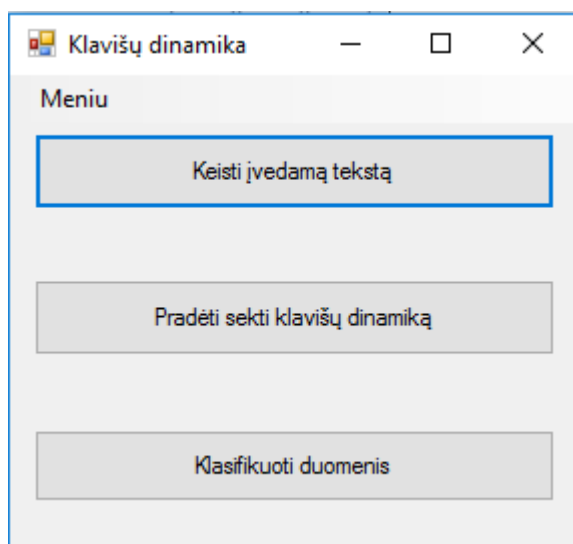
Programinė įranga realizuota C# programavimo kalba. Šios programinės įrangos UML panaudos atvejų diagrama pavaizduota 3.3 paveikslo schemeje. Pagal pateiktą schemas vartotojas gali atlikti tokius veiksmus:

- keisti įvesties tekstą;
- surinkti tekstą –
  - pasirinkti paros laiką;
- klasifikuoti duomenis –
  - datos intervalo pasirinkimas,
  - dalyvio pasirinkimas,
  - duomenų klasifikavimas su trijų klasifikatorių sistema,
  - duomenų klasifikavimas su klavišų porų klasifikavimo sistema,
  - duomenų klasifikavimas naudojant trijų klasifikatorių sistemą su klavišų svarbumo įvertinimu.



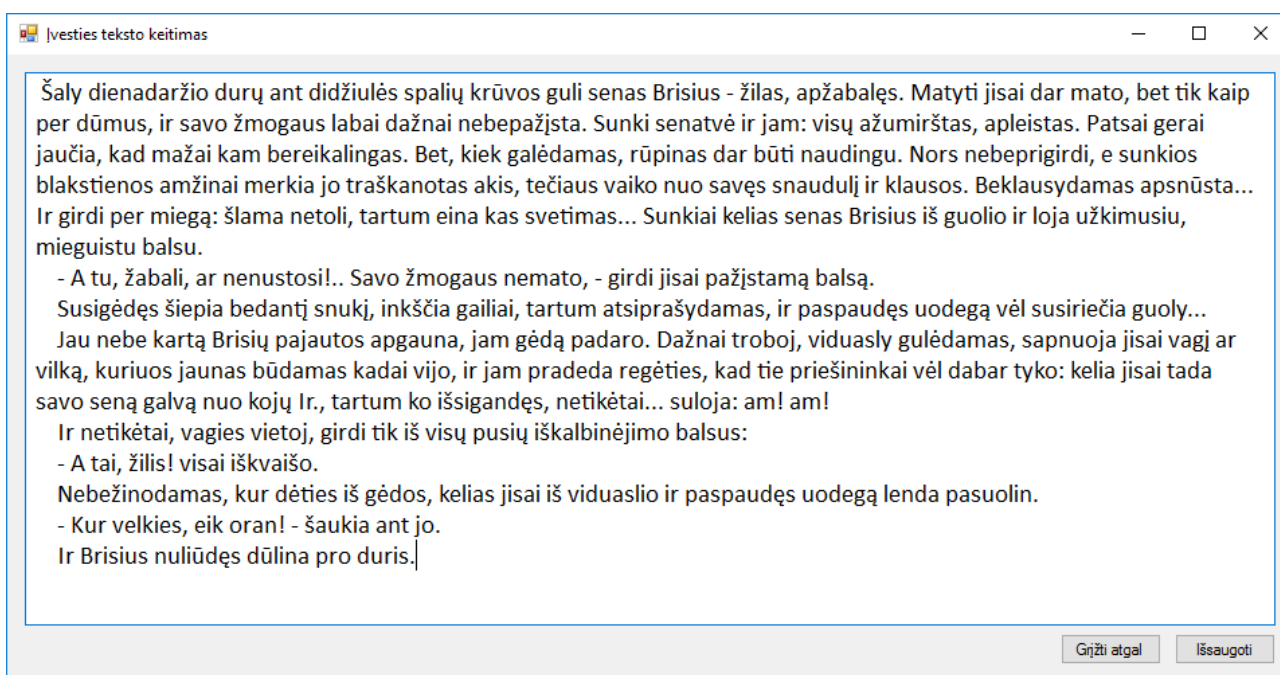
3.3 pav. Klavišų dinamikos sekimo programinės įrangos panaudos atvejų schema

Pradinis PĮ grafinės vartotojo sąsajos langas pateiktas 3.4 paveiksle



**3.4 pav.** Pradinis klavišų dinamikos sekimo PĮ grafinės vartotojo sąsajos langas

Pasirinkus „*Keisti įvesties tekstą*“ (3.4 pav.) yra atveriamas įvesties teksto keitimo langas (3.5 pav.), kuriame vartotojas gali įrašyti teiktą. Šis tekstas pateikiamas vartotojui, teksto rinkimo metu.



**3.5 pav.** Įvesties teksto keitimo grafinės vartotojo sąsajos langas

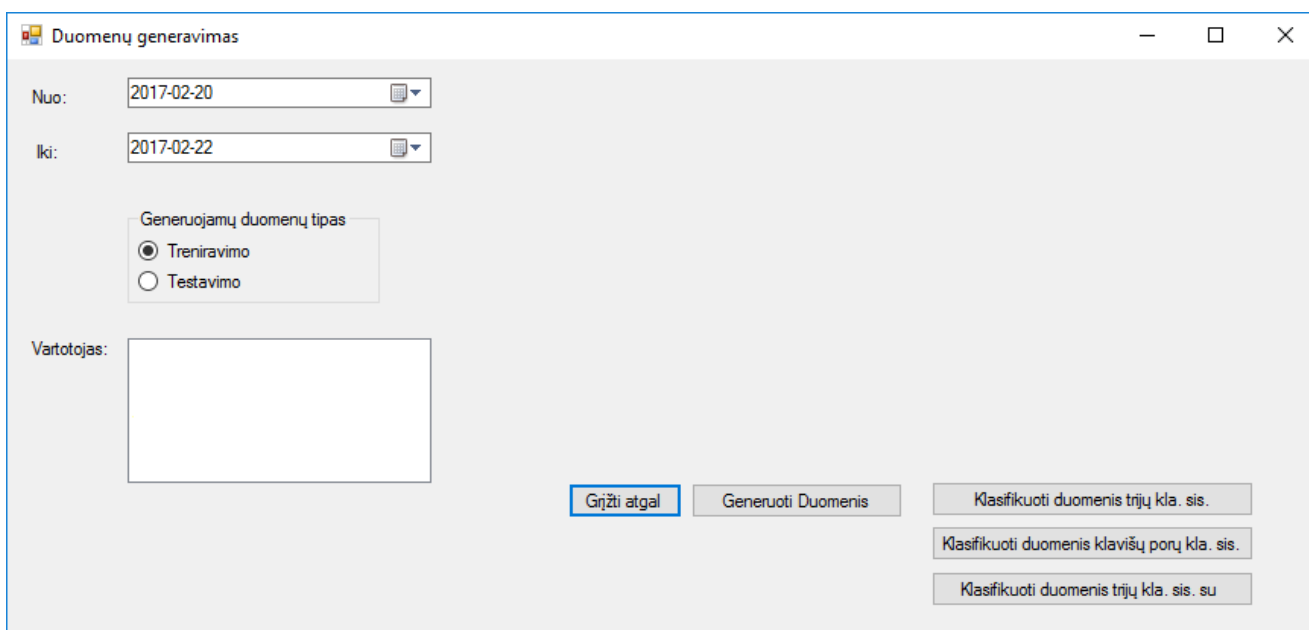
Pasirinkus „*Pradėti sekti klavišų dinamiką*“ yra atveriamas klaviatūros klavišų sekimo langas (3.6 pav.). Šiame lange, prieš pradėdant rinkti tekstą, būtina pasirinkti paros laiką, kada yra vedamas tekstas. Teksto vedimo metu sistema tikrina, ar suvestas tekstas atitinka bazinį tekstą ir, jeigu neatitinka, vartotoją informuoja apie esamą klaidą oranžiniais mirksinčiais ženkliukais. Tai padeda vartotojui identifikuoti tekste paliktas klaidas.



3.6 pav. Klaviatūros klavišų dinamikos sekimo grafinės vartotojo sąsajos langas

### 3.6. Duomenų klasifikavimas

Pasirinkus trečią pradinio lango pasirinktį „Klasifikuoti duomenis“ atveriamas duomenų klasifikavimo langas (3.7 pav.).



3.7 pav. Duomenų klasifikavimo grafinės vartotojo sąsajos langas

Šitas PĮ langas skirtas sugeneruoti duomenų imtį ir, pagal pasirinkta duomenų klasifikavimo metodą, apmokyti sistemą, nuovargio klasių nustatymui. Kadangi, bus ieškoma geriausios duomenų imties ir klasifikatorių struktūros, duomenų eksportavimas turi būti lankstu. Pagal 3.3 paveiksle pateikta UML panaudos atvejų schemą vartotojas prieš klasifikavimą privalo:

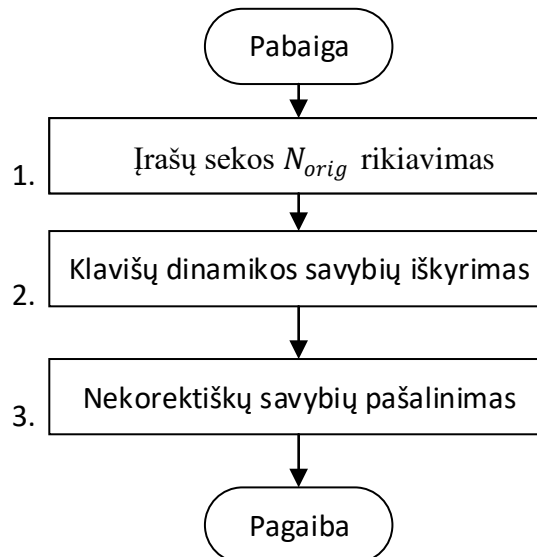
- pasirinkti datos intervalą iš kurio bus generuojama duomenų imtis;
- pasirinkti dalyvį, kuriam bus generuojama duomenų imtis;
- pasirinkti duomenų klasifikavimo metodą.

Duomenų klasifikavimui naudojama atvirojo kodo *LibSVM* biblioteka, kurioje realizuota atraminių vektorių mašinos klasifikavimo metodai [29].

### 3.7. Duomenų normalizavimas

Klaviatūros klavišų sekimo lange, renkant tekstą, sistema fiksuoja ir duomenų bazėje saugo absoliučius klavišo paspaudimo ir atleidimo laikus, datą, klavišo numerį ir paros laiko identifikatorių (3.3 lentelė). Taip suformuojama įrašų seka  $N_{orig} = \{n_1, n_2, \dots, n_i\}$ , kur  $n_i$  ir  $n_{i-1}$  nėra tarpusavyje susiję.

1.3.1 skyriuje aprašytų klavišų dinamikos savybių išskyrimui reikalingas įrašų sekos  $N_{orig}$  normalizavimas. Šame darbe taikomi duomenų normalizavimo etapai pateikta 3.8 paveiksle.



3.8 pav. Įrašų sekos N normalizavimas

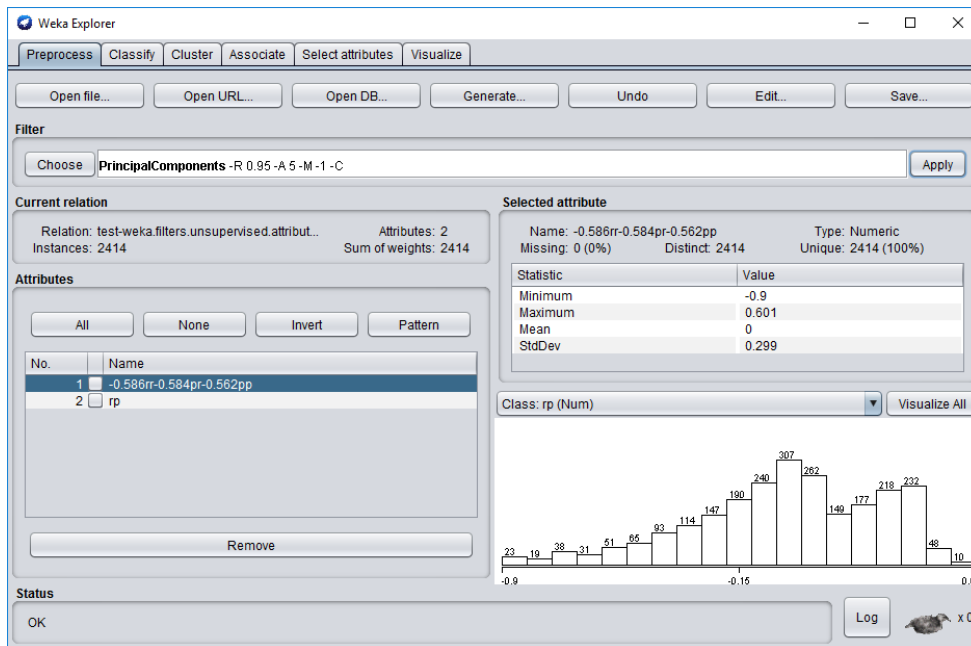
Normalizavimas susideda iš trijų etapų:

1. Įrašų sekos  $N_{orig}$  rikiavimas nuo mažiausio iki didžiausio pagal datos (*Date*) ir klavišo paspaudimo (*KeyDownTime*) duomenų lentelės lakus.
2. Klavišų dinamikos savybių išskyrimas.
3. Iš normalizuotos duomenų sekos  $N_{nor}$ , kurią sudaro paspaudimo, atleidimo – paspaudimo, paspaudimo – paspaudimo, atleidimo – atleidimo ir paspaudimo – atleidimo delsos laikai, pašalinama įrašai, kurių paspaudimo – paspaudimo laikai yra daugiau negu 0,4s.

### 3.8. Parametrų atrinkimas

Prieš pradėdant SVM klasifikatorių apmokymą, reikia suformuoti efektyviausią savybių rinkinį  $S = (s_1, s_2, \dots, s_3)$ , su kuriuo galima pasiekti tiksliausią klasifikavimą. Savybėms įvertinti panaudojamas 1.5 skyriuje aprašytas PCA metodas, kuris nusako kiekvienos savybės sklaidą imtyje ir SVM klasifikatorius, kuriuo įvertinama, koks klasifikavimo tikslumas gali būti pasiekiamas su atitinkamomis KD savybėmis  $s_n$ .

Savybių  $s_n$  sklaidos imtyje įvertinimui naudojama *Weka* programinė įranga, kuri plačiau aprašoma 3.10 skyriuje. PCA metodo taikymas, *Weka* PĮ pakete pavaizduotas 3.9 paveiksle, kuriame keturių dimensijų įėjimo vektorius  $W_{4d}(pp; pr, rr, rp)$  paverčiamas į dviejų dimensijų vektorių  $W_{2d}(-0,586rr - 0,584pr - 0,562pp; rp)$ . Savybių koeficientas  $q_i$  nusako konkrečios savybės sklaidą imtyje.



3.9 pav. Weka PĮ langas duomenų imties savybių mažinimui

### 3.9. Dvejetaisiais būsenų nustatymo metodas

2.2 skyriuje aptarta atraminių sektorių mašina gali klasifikuoti tik dvi klases (t.y. atlikti dvejetainį klasifikavimą). Dvejetainis klasifikavimas, turintis daugiau negu dvi klasifikuojamas klases, atliekamas vienu prieš visus (angl. *one vs all*) principu [30]. Šio metodo esmė – iš imties išskirti tik vieną klasę, o likusias klases apjungti į vieną.

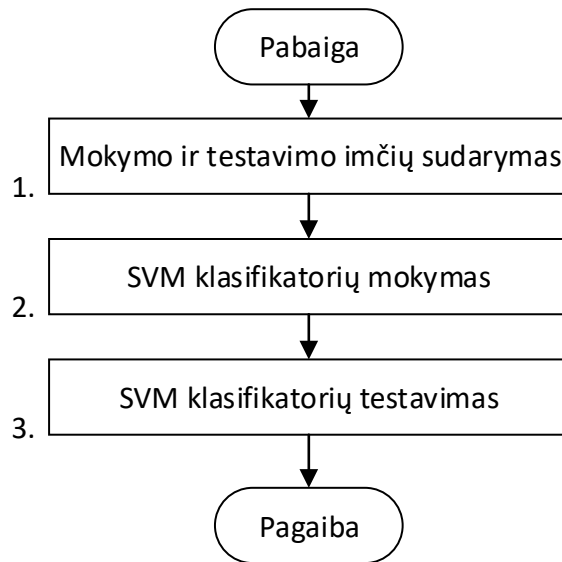
Mūsų tiriamas atvejis turi tris klases, todėl tyrime naudojama klasifikatorių sistema (3.11 pav.), kur kiekvienas klasifikatorius atskiria vieną konkrečią klasę nuo likusių. Klasifikatorių mokymo imčių formavimo lentelė pateikta 3.4 lentelėje.

3.4 lentelė. Klasifikatorių mokymo imtys

Klasifikatorius	Prognozuojama klasė	Apjungiamos klasės	
		Diena	Vakaras
1.	Rytas	Diena	Vakaras
2.	Diena	Vakaras	Rytas
3.	Vakaras	Diena	Rytas

#### 3.9.1. Trijų klasifikatorių sistema

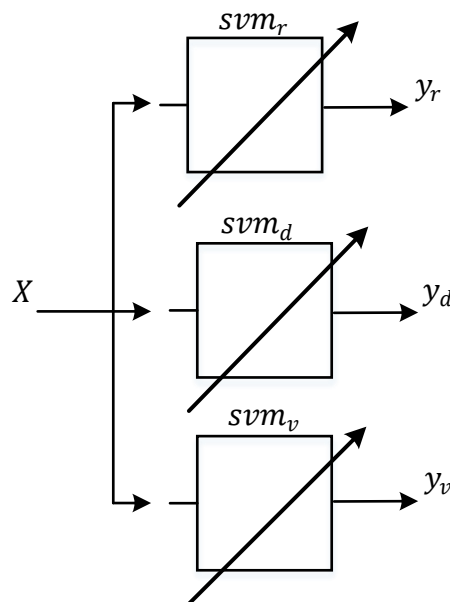
Trijų klasių atpažinimui sudaroma sistema su trimis klasifikatoriais, kur kiekviena, atraminių vektorių mašina, apmokoma su duomenų imtimis, sudarytomis pagal 3.4 lentelė. 3.10 paveiksle pateikta trijų klasifikatorių sistemos sekos diagrama, kuri naudojama sistemos nuovargio klasių klasifikavimo tikslumo įvertinimui..



**3.10 pav.** Trijų klasifikatorių sistemos sekos diagrama

Sistemą sudaro trys žingsniai:

1. *mokymo ir testavimo imčių sudarymas* – pasirinktam laiko intervalui sistema suformuoja mokymo ir testavimo duomenų imtis  $X_t = (x_{t1}, x_{t2}, \dots, x_{tn})$ ,  $t \in \text{rytas, diena, vakaras}$ .
2. *SVM klasifikatorių mokymas* – kiekvienos nuovargio klasės atpažinimui apmokomas SVM klasifikatorius  $svm_t$ , kuriam parenkamos tokios  $C$  ir  $\gamma$  reikšmėmis, jog klasifikavimo tikslumas  $P$  būtų maksimalus.
3. *Sistemos testavimas* – sistemos testavimui naudojamas trijų klasifikatorių modelis (3.11 pav.), kurio įėjimas yra savybių  $s_n$  vektorius  $x_t(s_1, s_2, \dots, s_n) \in X_t$ . Modulio išėjimas yra vektorius  $Y_t(y_r; y_d; y_v)$ , kur  $y$  – modelio prognozuojamos klasės tikimybė, išreikšta procentais. Nuovargio klasė parenkama pagal didžiausią vektoriaus  $Y_t$  klasės tikimybę  $y$ . Trijų klasifikatorių sistemos testavimas atliekamas su Visomis trimis duomenų imtimis  $X_t$ .

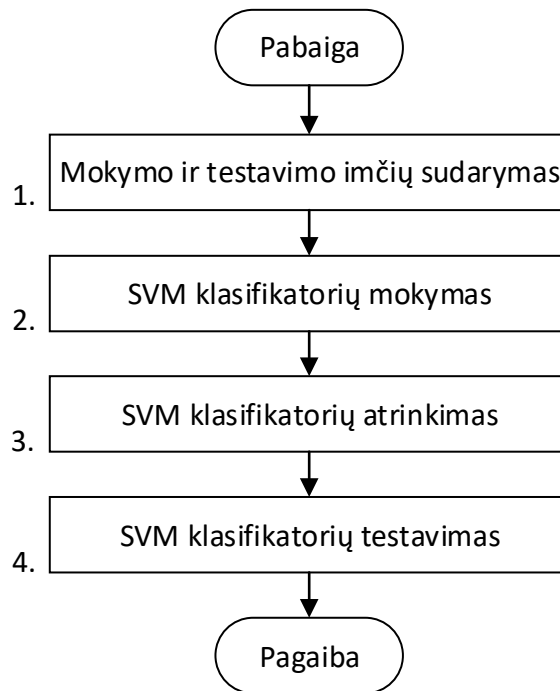


**3.11 pav.** Trijų klasifikatorių sistemos modelis

### 3.9.2. Klavišų porų klasifikatorių sistema

3.9.1 skyriuje aprašyto sistemos klasifikatorių įėjimo savybių vektoriaus  $x_t$  savybė  $s_1$  yra dviejų klavišų porų numerių (3.2 lentelė) kodas. Norint įvertinti nuovargio įtaką kiekvienai klavišų poros

dinamikai, sudaroma sistema, kurioje trijų klasifikatorių sistemos modelis bus sudaromas kiekvienai sistemos fiksuotai klavišų porai atskirai. Šios sistemos sekų diagrama pavaizduota 3.12 paveiksle.

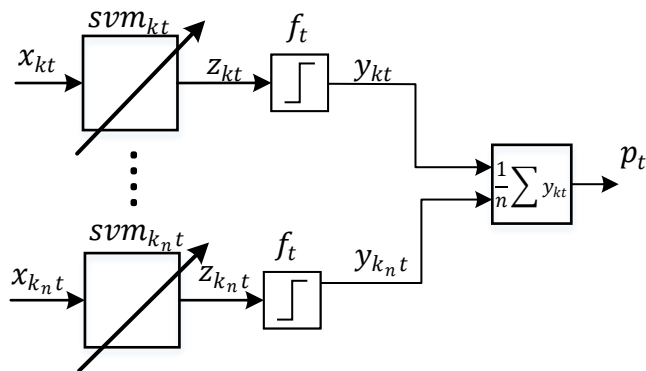


**3.12 pav.** Klavišų porų sistemos sekos diagrama

Sistemą sudarantys žingsniai:

1. *mokymo ir testavimo imčių sudarymas* – pasirinktam laiko intervalui sistema suformuoja mokymo ir testavimo duomenų imtis  $X_{kt} = (x_{t1}, x_{t2}, \dots, x_{tn})$ , kur  $k$  – klavišų porų kodas,  $t \in \{\text{rytas}; \text{diena}; \text{vakaras}\}$ .
2. *SVM klasifikatorių mokymas* – kiekvienai klavišų porai  $k$  apmokomas SVM klasifikatorius  $svm_t$ , kuriam parenkamos tokios  $C$  ir  $\gamma$  reikšmėmis, jog klasifikavimo tikslumas  $P$  būtų maksimalus. Sistemai grąžinamas masyvas  $T = \{(k_1, smv_1, p_1); (k_2, smv_2, p_2); \dots; (k_n, smv_n, p_n)\}$ , kur  $k$  – klavišų poros kodas,  $smv$  – SVM klasifikatorius,  $p$  – duomenų klasifikavimo tikslumas.
3. *SVM klasifikatorių atrinkimas* – pagal masyvo  $T$ , reikšmes  $p_n$  atrenkamos klavišų poros ir SVM klasifikatoriai, kurių mokymo etape klasifikavimo tikslumas buvo didesnis už slenkstinę reikšmę  $sl_1$ .
4. *Sistemos testavimas* – sistemos testavimui naudojamas 3.11 paveiksle pateiktas modelis, tačiau šiuo atveju  $svm_t$  klasifikatorius keičiamas į  $svm_k$  klasifikatorių posistemę, kurios modelis pateiktas 3.13 paveiksle. Posistemės išėjimas  $p_t$  yra klasifikavimo tikslumas, kuris nusako, kiek procentų  $svm_{kt}$  klasifikatorių duomenų klasifikavimo tikslumas  $z_{kt}$  tenkino slenkstinės funkcijos  $f_t$  reikšmę  $sl_2$ .



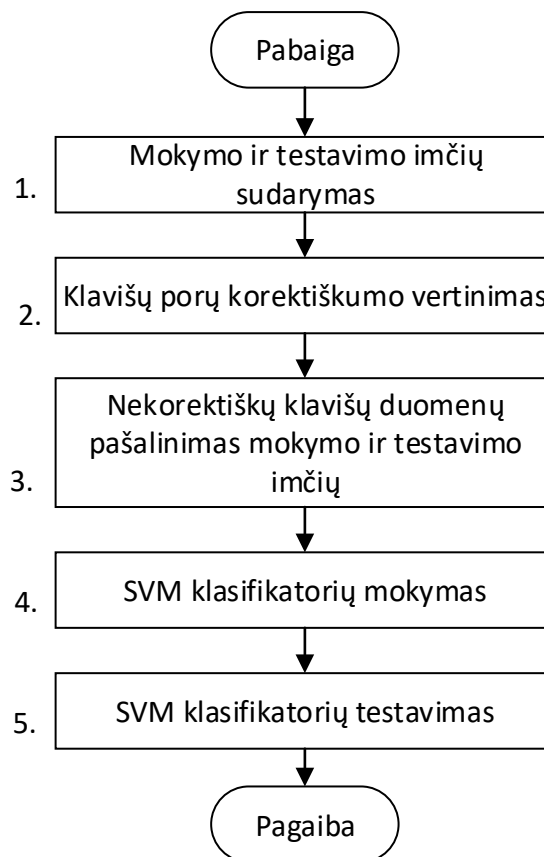


3.13 pav.  $svm_k$  klasifikatorių posistemės modelis

### 3.9.3. Trijų klasifikatorių sistema su klavišų porų svarbumo įvertinimu

3.9.1 skyriuje aprašytas sistemos modelis nevertina klavišų porų, kurių paspaudimo dinamikai nuovargis įtakos nedarė. Tokias klavišų poras galima įvardinti sistemos jėgimo imčių  $X_t$  triukšmais, kurie gali bloginti sistemos klasifikavimo tikslumą.

Trijų klasifikatorių sistemos papildyta nekorektiškų klavišų porų nustatymu ir pašalinimu iš duomenų imčių sekos diagrama pateikta 3.14 paveiksle.



3.14 pav. Trijų klasifikatorių sistemos modelio sekų diagrama su klavišų porų korektiškumo įvertinimu

Sistemą sudaro penki žingsniai:

1. *Mokymo ir testavimo imčių sudarymas* (plačiau aprašyta 3.9.1 skyriuje);
2. *Klavišų porų korektiškumo vertinimas* – Klavišų korektiškumui įvertinti panaudojama 3.9.2 skyriuje aprašyta metodika, pagal kurią iš bendros mokymo imties suformuojamos klavišų porų imtys  $X_{kt}$ , kiekvienai klavišų porai  $k$  apmokomas SVM klasifikatorius su tokiais  $C$  ir  $\gamma$  reikšmėmis, jog klasifikavimo tikslumas  $P$  būtų maksimalus. Sistemai grąžinama masyvas  $T = \{(k_1, p_1); (k_2, p_2); \dots; (k_n, p_n)\}$ .

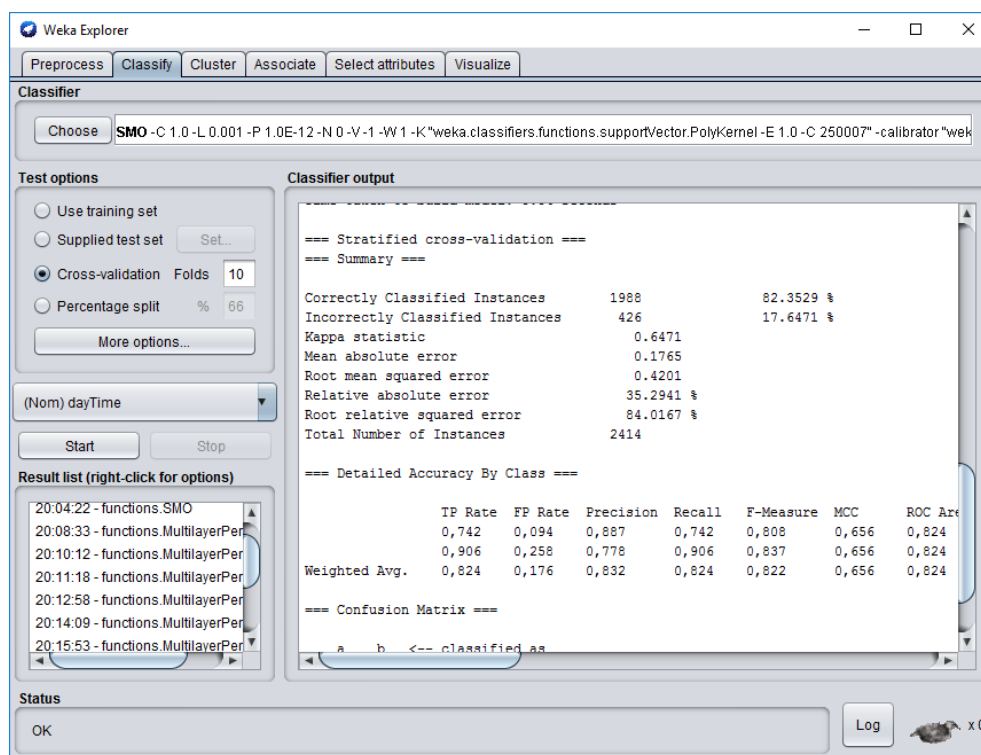
3. *Nekorektiškų klavišų duomenų pašalinimas iš mokymo ir testavimo imčių* – Pagal  $svm_{k_n t}$  klasifikatorių duomenų klasifikavimo tikslumą  $p_{k_n t}$  ir slenkstinės funkcijos  $f_t$  reikšmę atrenkamos triukšmingos klavišų poros ir jų duomenys pašalinami iš mokymo ir testavimo imčių.
4. *SVM klasifikatorių mokymas* (plačiau aprašyta 3.9.1 skyriuje);
5. *Sistemos testavimas* (plačiau aprašyta 3.9.1 skyriuje).

### 3.10. Weka programinis įrangos paketas

*Weka* paketas yra atvirojo kodo programinė įranga, kuri plėtojama Vaikato universiteto [31]. Joje yra realizuoti savaiminiai apsimokymo ir duomenų gavybos algoritmai. *Weka* PĮ paketas, pagal nutylėjimą, naudoja *ARFF* formato failus, tačiau duomenys taip pat gali būti užkraunami iš *CSV* failo arba tiesiogiai iš duomenų bazės. *Weka* PĮ pakete realizuota:

- išankstinis duomenų apdorojimas
- klasifikavimas;
- regresija;
- grupavimas;
- vizualizavimas;
- asociacijos taisyklių nustatymas.

*Weka* PĮ duomenų klasifikavimo grafinės vartotojo sąsajos langas pateiktas 3.15 paveiksle.



3.15 pav. *Weka* PĮ duomenų klasifikavimo grafinės vartotojo sąsajos langas

#### 4. EKSPERIMENTINĖ DALIS

Ekspimento metu buvo analizuojama keturių tyrime dalyvavusių asmenų duomenys. Trys tyrime dalyvavę asmenys buvo dirbantys, vienas – studentas. Jų amžius buvo tarp 22 iki 33 metų.

Tyrimas truko dvi savaites, per kurias, tyrime dalyvavę asmenys, tris kartus per dieną vedė PĮ pateikiamą tekstą. Tyrimo metu tekstas buvo keistas tris kartus. 4.1 lentelėje pateikti tyrimo dalyvių rinkti tekstai ir datos, kada dalyviai juos rinko.

##### 4.1 lentelė. Tyrimo metu dalyvių rinkti tekstai

Ištrauka iš Vytauto V. Lansbergio "Partizanas" [32]. Tekstas rinktas 2017.02.20 - 2017.02.23 dienomis.
Vienumoje dažnai galvoju apie žuvusius. Viešpatie mano, kiek jų daug. Mūsų tėvynė tokia maža, o žūva geriausi vyrai, jauni, stojiški, kupini pasiryžimo, drąsūs... Ir širdį kartais taip pradeda skaudėti, kad imu niekuo netikėti. Žmogus juk nesi geležinis, o ir geležis per trumpą laiką surūdija sutrupa. Ir todėl kartais pradėdi dvejoti, argi neišsėks mūsų jėgos? Kas iš tos Lietuvos, jeigu ji nustos geriausios savo dalies? Kas bus, jeigu atėjus lemiamam momentui nebus kam jos prikelti, išgydyti? Jaučiu, kad savo kraštą pamilstu vis labiau. Jeigu kas pasiūlytų laisvę Amerikoje, neišvažiuočiau. Geriau žūti čia garbingai kovojant negu rankas sudėjus laukti. Pagaliau mūsų kraujas nenueis veltui. Mes turėsime teisę visiems žiūrėti tiesiai į akis, nes mes savo tėvynės nepalikome. Ir kas gi mus nugalės, jei mes mirti nebijom, jeigu mes nugalėjome ir mirtį.
Ištrauka iš Haruki Murakami knygos "Dramblys pradingsta" [33]. Tekstas rinktas 2017.02.24 - 2017.03.01 dienomis.
Labiausiai aš bijau tų, kurie nesusimąstydami, nekritiškai priima tokių, kaip Aokis, kalbas ir jomis tiki. Tų, kurie patys iš savęs nieko nepagamina ir nieko nesupranta, o tik kaip kvailių avinų banda vadovaujasi patogia, gražiai skambančia kitų nuomone. Jie nė per nago juodymą nesusimąsto, kad gal patys yra neteisūs. Jie banda, kuriai net nešauna į galvą, kad galbūt ką nors beprasmiškai ir mirtinai žaloja. Jie net negalvoja prisiimti atsakomybės už savo veiksmų padarinius. Va tokių aš bijau. Ir va tokius aš sapnuoju naktimis. Sapne tvyro tylą. O žmonės be veidų. Tyla kaip vėsus vanduo pamažu ima skverbtis į atokiausias kerteles. Ir viskas toje tyloje ima pamažu tirpti. Aš tirpstu ir iš visų jėgų rėkiu, bet manęs niekas negirdi.
Šimtas dažniausiai lietuvių kalboje naudojamų žodžių, kurie sudaryti iš ne mažiau kaip trijų raidžių [34]. Tekstas rinktas 2017.03.02 - 2017.03.03 dienomis.
kad yra tai buvo kaip savo nuo jis bet apie tik jau dar darbo Lietuvos gali būti bei arba kai labai turi jos dėl čia per taip prie tačiau kas jie galima reikia nes metų dabar mūsų jei gamybos pagal todėl net iki būtų mes ant daug tai nėra bus kur jeigu daugiau pirmininkas taip pat nei seimo tada kiek gal prieš žmogus kurie žemės metu pats tarp jam kuris žmogaus Tomas kitų man įmonės laiko visi tas visai žmonių mano vėl juos jog respublikos ūkio kaip ten Lietuvoje gerai valstybės sistemos ypač nors pavyzdžiui kurios šis duomenų šios lietuvių.

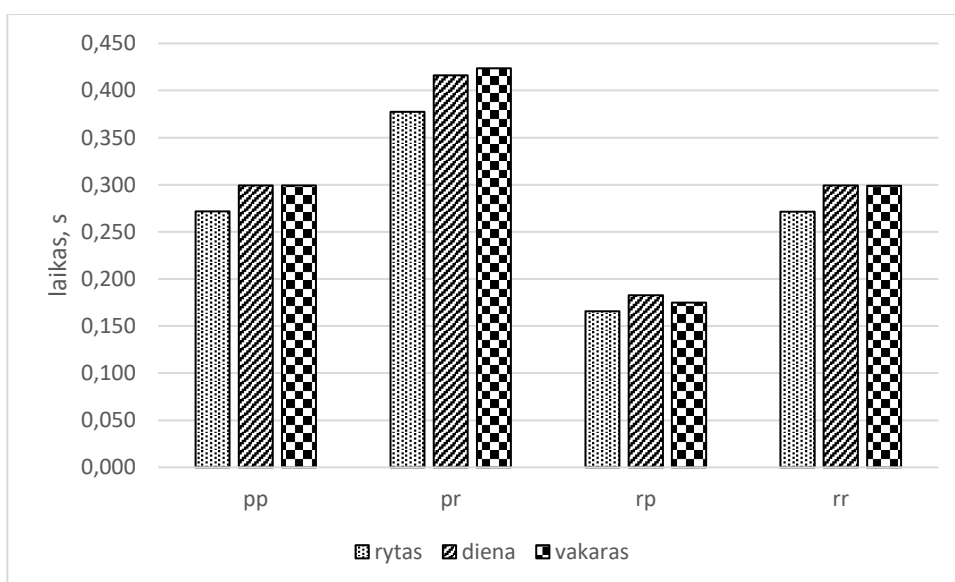
##### 4.1. Tyrimo dalyvių vidutinės KD savybių vertės

Tyrimas buvo atliekamas su prielaida, kad paros laikas, kada yra vedamas tekstas, koreliuoja su žmogaus nuovargiu. Laikoma, kad ryte vestas tekstas reprezentuoja asmenį, kuris nejaučia nuovargio, diena – vidutinišką nuovargį jaučiantį asmenį ir vakare – pavargusį asmenį. 4.2 lentelėje pavaizduoti kiekvieno tyrime dalyvavusio asmens viso tyrimo apibendrinti duomenys ir išvesti klavišų dinamikų vidurkiai atitinkamam paros laikui.

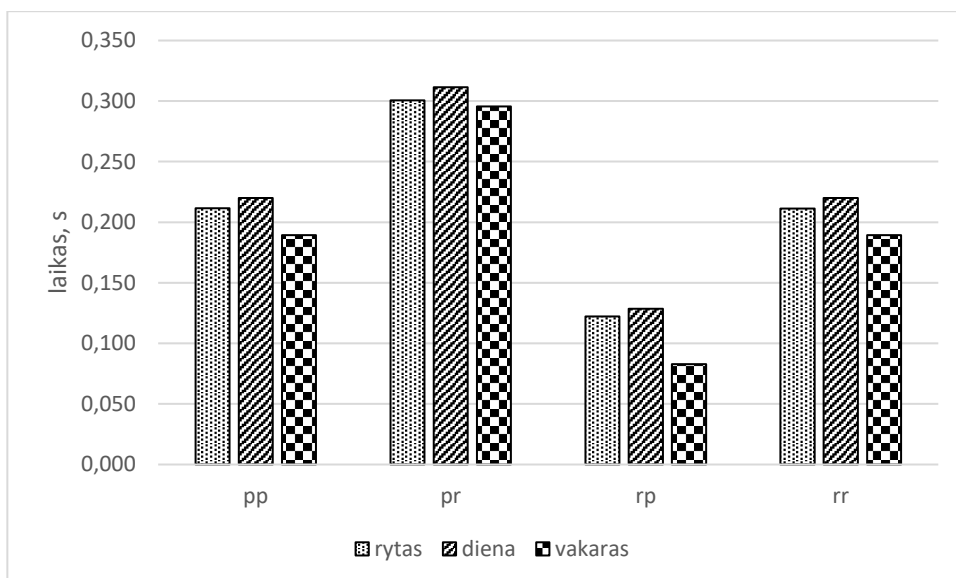
**4.2 lentelė.** Vidutinės tyrime dalyvavusių asmenų KD savybių reikšmės atitinkamam paros laikui

Dalyvis	Klasė	<i>pp</i>	<i>pr</i>	<i>rp</i>	<i>rr</i>
1	rytas	0,272	0,377	0,166	0,271
	diena	0,299	0,416	0,183	0,299
	vakaras	0,299	0,424	0,175	0,299
2	rytas	0,211	0,301	0,122	0,211
	diena	0,220	0,311	0,129	0,220
	vakaras	0,189	0,296	0,083	0,189
3	rytas	0,332	0,428	0,237	0,333
	diena	0,323	0,420	0,226	0,324
	vakaras	0,302	0,411	0,195	0,304
4	rytas	0,229	0,337	0,122	0,229
	diena	0,217	0,327	0,107	0,217
	vakaras	0,190	0,303	0,077	0,190

Dalyvių rinkto teksto KD savybių vidutinėse vertėse keičiasi priklausomai nuo teksto rinkimo laiko. Mūsų tiriamu atveju, teksto vedimo laikui būdinga atitinkama nuovargio būseną. Pirmojo ir antrojo dalyvių klavišų dinamikos vidurkiai pateikti atitinkami 4.1.pav. ir 4.2 pav. stulpelinėse diagramose.



**4.1 pav.** Dalyvio Nr. 1 klavišų dinamikos vidurkių stulpelinė diagrama



4.2 pav. Dalyvio Nr. 2 klavišų dinamikos vidurkių stulpelinė diagrama

Iš 4.1 ir 4.2 paveiksluose pateiktų stulpelinių diagramų matyti, jog absoliučios KD savybių vidurkių vertės, renkant tekstą esant skirtingam nuovargiui, kinta. Dalyvio Nr. 1 klavišų dinamikos savybių vidutinės reikšmės (4.1. pav.) neturi bendros tendencijos, t.y. savybių *pp* ir *rr* dienos ir vakaro vidurkiai yra lygūs, *pr* – didėja, o *rp* dienos vidurkis išauga lyginant su ryto, tačiau vakaro vidurkis sumažėja. Tuo tarpu, dalyvio Nr. 2 vidutinėms klavišų dinamikos savybės (4.2. pav.) turi bendrą didėjimo ir mažėjimo tendenciją – diena rinkto teksto KD savybių vidurkiai išauga, vakare – sumažėja. Iš 4.1 lentelėje 4.1 ir 4.2 paveiksluose pateiktų duomenų galime teigti, kad:

- tyrime dalyvavusių asmenų KD savybės neturi bendros tendencijos ir kiekvieno dalyvio KD savybių vidurkiai, keičiantis nuovargiui, kinta individualiai;
- KD savybių vertės, kai kurias atvejais (4.1 pav. *pp* ir *rr* savybių dienos ir vakaro reikšmės) gali būti vienodos.

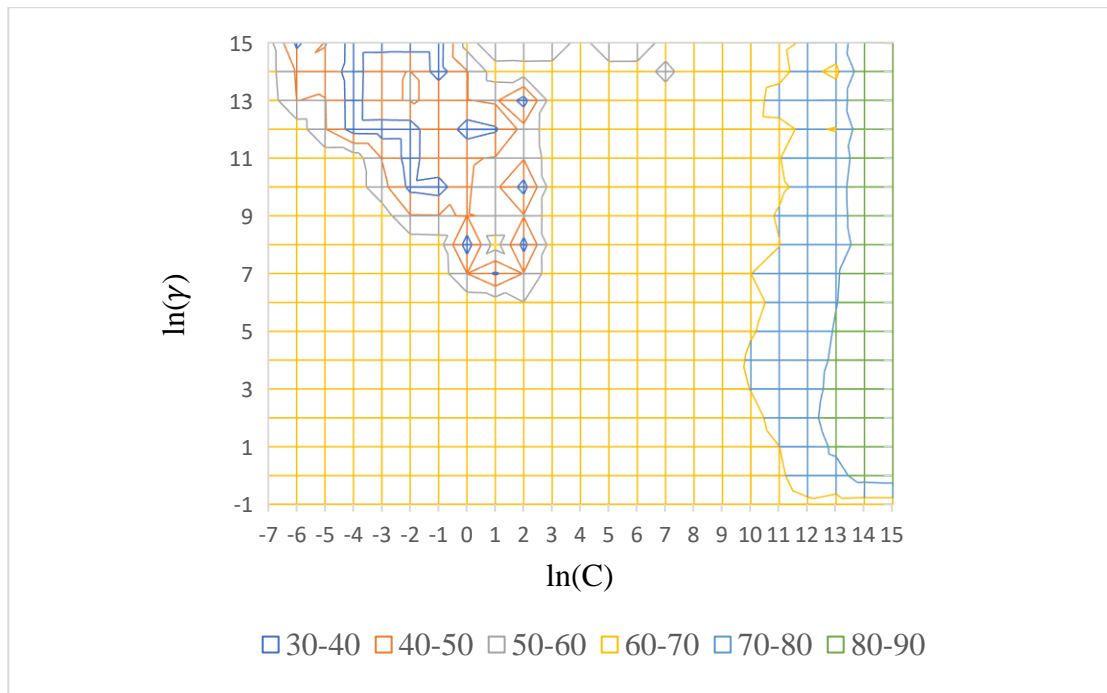
#### 4.2. Branduolio funkcijos ir baudos koeficientų parinkimas

Standartinis SVM klasifikatorius skirtas klasifikuoti tiesiškai atskiriamas duomenų imties klases, tačiau dažniai duomenų imties klasės yra atskiriamos netiesiškai. Klasifikuojant netiesiškai atskiriamus duomenis su SVM klasifikatoriumi, duomenų pervedimui į aukštesnę erdvę, reikia naudoti branduolio funkcijas. Šiame tyrime naudojama RBK (plačiau aprašyta 2.2.3 skyriuje) branduolio funkcija.

Naudojant SVM su RBK branduolio funkciją, reikia parinkti  $C$  ir  $\gamma$  reikšmes. Šios reikšmės nėra žinomos, todėl, kiekvienai duomenų imčiai, reikia rasti tokią minėtų parametų porą, su kuria nesuklasifikuotų duomenų skaičius būtų sumažintas iki minimalaus. Šiam tikslui pasiekti naudojama tinklo paieška (angl. *Grid Search*), kurios metu bandomos įvairios  $C$  ir  $\gamma$  parametų reikšmių poros. Parenkama ta parametų pora, su kuria SVM tiksliausiai klasifikavo duomenų imties klases. Klasifikavimo tikslumui įvertinti naudojamas gerai klasifikuotų duomenų ir visų klasifikuotų duomenų santykis išreikštas procentais:

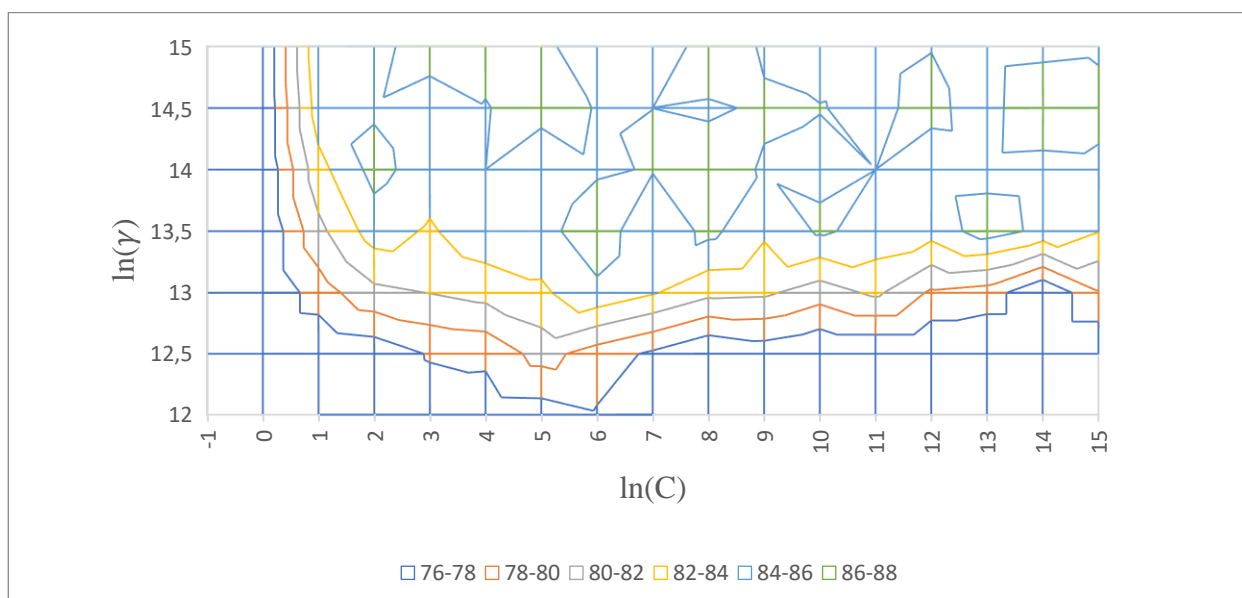
$$p = \frac{\sum \text{gerai klasifikuoti domenys}}{\sum \text{klasifikuoti duomenys}} * 100\% \quad (4.1)$$

Remiantis Chih-Wei Hsu'no [29]  $C$  ir  $\gamma$  parametų parinkimo metodika, buvo nuspręsta naudoti logaritminę skalę ir paiešką atlikti du kartus. Pirminė tinklo paiešką buvo atlikta naudojant eksponentiškai didėjančias  $C$  ir  $\gamma$  sekas:  $C = (10^{-15}, 10^{-14}, \dots, 10^{15})$ ,  $\gamma = (10^{-15}, 10^{-14}, \dots, 10^{15})$ , kur. 4.3 paveiksle pateikta pirminės tinklo paieškos rezultatai. Duomenų atvaizdavimui grafike naudojama logaritminė duomenų skalė.



**4.3 pav.** Duomenų klasifikavimo tikslumas  $p$  prie atitinkamų  $C$  ir  $\gamma$  reikšmių.

Iš gautų rezultatų matyti, jog intervaluose  $C = [C^{-7}; C^7]$  ir  $\gamma = [\gamma^6; \gamma^{15}]$  fiksuojamas 30 – 50 % klasifikavimo tikslumas. Maksimalios 80 – 90%.  $p$  reikšmės fiksuojamos  $C = [C^{12}; C^{15}]$  ir  $\gamma = [\gamma^{-1}; \gamma^{15}]$  intervaluose. Norit surasti globalų  $p$  maksimumą, minėtuose intervaluose, atliekama antrinė tinklo paieška, kuriai taikoma dvigubai mažesnis parametru kitimo žingsnis –  $y_n = y_0^{\ln(x_0)+0,5n}$ . Antrinės tinklo paieškos rezultatai pateikti 4.4 paveiksle.



**4.4 pav.** Antrinės tinklo paieškos duomenų klasifikavimo tikslumas prie atitinkamų  $C$  ir  $\gamma$  reikšmių

Atlikus antrinę tinklo paiešką matyti, jog su skirtingomis  $C$  ir  $\gamma$  parametru poromis pasiekiamas vienodas 86 – 88 % klasifikavimo tikslumas. Kadangi absoliutus maksimumas nebuvo rastas, tolimesniame tyrime, vykdant klasifikatoriaus apmokymą, bus naudojama tinklo paieška  $C = [C^4; C^{12}]$  ir  $\gamma = [\gamma^{13,5}; \gamma^{15}]$  režuose su  $y_n = y_0^{\ln(x_0)+0,5n}$  parametru kitimo žingsniu.

### 4.3. Klavišų dinamikos savybių atrinkimas

Kadangi, pasitaiko atveju, kai kintant nuovargiui tam tikros klavišų dinamikos savybės išlieka vienodos (4.1 pav.), reikia įvertinti kiekvienos savybės įtaką SVM klasifikatoriaus mokymui ir duomenų klasifikavimui. Įvertinti savybių įtaką, duomenų imtyse, naudojamas 1.5 skyriuje aprašytas

svarbiausių komponentų analizės metodas. *Weka* PĮ gauti KD savybių *pp*, *pr*, *rr* ir *rp* sklaidos koeficientų reikšmės  $q_i$  pateiktos 4.3 lentelėje.

**4.3 lentelė.** Svarbiausių komponentų analizės metodu duomenų imtims apskaičiuoti KD savybių sklaidos koeficientai

Dalyvis	Klasė	<i>pp</i>	<i>pr</i>	<i>rr</i>	<i>rp</i>
1	rytas	0,499	0,5	0,501	0,5
	diena	0,5	0,499	0,5	0,501
	vakaras	0,5	0,5	0,5	0,5
2	rytas	0,492	0,508	0,508	0,491
	diena	0,492	0,511	0,508	0,489
	vakaras	0,5	0,5	0,5	0,5
3	rytas	0,501	0,499	0,501	0,499
	diena	0,501	0,499	0,501	0,499
	vakaras	0,5	0,5	0,5	0,5
4	rytas	0,492	0,501	0,508	0,499
	diena	0,485	0,503	0,515	0,496
	vakaras	0,501	0,5	0,5	0,499

Iš gautų duomenų matyti, jog visos naudojamos KD savybės, nepriklausomai nuo dalyvio ir jaučiamo nuovargio, turi panašią  $q = 0,5 \pm 0,15$  sklaidą duomenų imtyje.

Norit įvertinti savybių įtaką, SVM klasifikavimo tikslumui, sudaromi SVM klasifikatoriai, kurie apmokomi su, iš skirtingų KD savybių sudarytomis, duomenų imtimis. Atraminių vektorių mašina duomenis klasifikuoja vienas prieš visus principu, todėl, trijų nuovargio klasių klasifikavimui, sudaromi trys SVM klasifikatoriai. Klasifikatorių dalyvių nuovargio klasių klasifikavimo tikslumo rezultatai pateiktas 4.4 lentelėje.

**4.4 lentelė.** SVM klasifikavimo tikslumas mokymo imtyje naudojant *pp*, *pr*, *rr* ir *rp* savybes

Dalyvis / Klasė	1	2	3	4
	<i>p</i> , %	<i>p</i> , %	<i>p</i> , %	<i>p</i> , %
rytas	68,59	65,82	66,58	70,54
diena	66,22	67,80	66,00	65,01
vakaras	66,73	70,42	66,37	64,51

Atraminių vektorių mašina, klasifikavimui naudojant *pp*, *pr*, *rr* ir *rp* klavišų dinamikos savybes, visais atvejais pasiekė ne mažesnę kaip 64% tikslumą. Maksimalus, 70,42% procentų klasifikavimo tikslumas buvo pasiektas prognozuojant antro dalyvio „vakaras“ klasę. Norit didinti nuovargio klasių klasifikavimo tikslumą, į klasifikavimo duomenų imtį reikia įtraukti savybes, kurios turėtų mažesnę sklaidą klasifikavimo imtyje, negu ją turi 4.3 lentelėje aprašytos savybės.

Šiam tikslui pasiekti panaudojamos KD savybių vidutinės reikšmės. Iš duomenų bazėje turimų duomenų apskaičiuojamos vidutinės *pp*, *pr*, *rr* ir *rp* nuovargio klasių savybių reikšmės:

- dienos – apskaičiuojamos konkrečios dienos vidutinės *pp*, *pr*, *rr* ir *rp* savybių reikšmės;
- klavišų poros – iš visos duomenų imties apskaičiuojamos konkrečios klavišų poros vidutinis *pp*, *pr*, *rr* ir *rp* savybių reikšmės;
- klavišų poros / dienos – apskaičiuojamos vidutinės *pp*, *pr*, *rr* ir *rp* klavišų porų savybių reikšmės kiekvienai dienai.

Įvertinti savybių sklaidą duomenų imtyse su papildomomis KD savybėmis vėl pritaikomas svarbiausių komponentų analizės metodas. klasifikavimo tikslumo įvertinimui apmokomos atraminių vektorių mašinos. Gauti rezultatai, pagal papildomas vidutines dienos, klavišų poros ir klavišų poros/dienos savybes, pateikti atitinkamai 4.5, 4,6 ir 4,7 lentelėse.

**4.5 lentelė.** KD savybių ir jų vidurinių dienos reikšmių sklaidos koeficientų  $q$  reikšmės su SVM klasifikavimo tikslumu

Dalyvis	Klasė	$q_{rr}$	$q_{rp}$	$q_{pp}$	$q_{pr}$	$q_{rrAvg}$	$q_{rpAvg}$	$q_{ppAvg}$	$q_{prAvg}$	p,%
1	rytas	0,510	0,506	0,490	0,493	0,001	0,002	0,002	0,003	75,28
	diena	0,500	0,501	0,500	0,499	0	0	0	0	67,72
	vakaras	0,500	0,500	0,500	0,500	0	0	0	0	80,78
2	rytas	0,515	0,498	0,485	0,502	0,005	0,003	0,005	0,007	68,72
	diena	0,509	0,494	0,490	0,506	0,011	0,009	0,012	0,014	71,58
	vakaras	0,500	0,500	0,500	0,500	0,000	0,000	0,000	0,000	95,57
3	rytas	0,507	0,508	0,491	0,492	0,003	0,003	0,003	0,003	76,97
	diena	0,504	0,507	0,496	0,493	0,003	0,002	0,003	0,003	67,71
	vakaras	0,499	0,501	0,501	0,050	0,000	0,000	0,000	0,000	90,15
4	rytas	0,507	0,502	0,493	0,498	0,000	0,000	0,000	0,000	77,47
	diena	0,516	0,500	0,484	0,499	0,002	0,001	0,002	0,003	66,67
	vakaras	0,500	0,501	0,500	0,499	0,000	0,000	0,000	0,000	88,90

**4.6 lentelė.** KD savybių ir jų vidurinių dienos reikšmių sklaidos koeficientų  $q$  reikšmės su SVM klasifikavimo tikslumu

Dalyvis	Klasė	$q_{rr}$	$q_{rp}$	$q_{pp}$	$q_{pr}$	$q_{rrAvg}$	$q_{rpAvg}$	$q_{ppAvg}$	$q_{prAvg}$	p,%
<b>Vidurkis apskaičiuotas atitinkamai dienai</b>										
1	rytas	0,410	0,408	0,392	0,394	0,307	0,305	0,290	0,291	68,63
	diena	0,460	0,461	0,460	0,459	0,196	0,197	0,196	0,195	66,10
	vakaras	0,430	0,431	0,431	0,430	0,254	0,255	0,255	0,254	69,11
2	rytas	0,419	0,405	0,390	0,404	0,308	0,297	0,279	0,289	67,07
	diena	0,393	0,383	0,377	0,387	0,327	0,319	0,311	0,318	66,91
	vakaras	0,381	0,382	0,382	0,382	0,322	0,323	0,323	0,323	75,70
3	rytas	0,392	0,392	0,379	0,378	0,325	0,326	0,312	0,311	66,37
	diena	0,392	0,394	0,385	0,382	0,319	0,325	0,311	0,310	66,12
	vakaras	0,447	0,449	0,449	0,447	0,221	0,223	0,223	0,021	68,67
4	rytas	0,373	0,368	0,360	0,364	0,347	0,343	0,335	0,338	72,34
	diena	0,399	0,386	0,371	0,384	0,332	0,321	0,305	0,317	64,23
	vakaras	0,359	0,360	0,360	0,359	0,347	0,348	0,348	0,347	63,83

**4.7 lentelė.** KD savybių ir jų vidurinių dienos reikšmių sklaidos koeficientų  $q$  reikšmės su SVM klasifikavimo tikslumu

Dalyvis	Klasė	$q_{rr}$	$q_{rp}$	$q_{pp}$	$q_{pr}$	$q_{rrAvg}$	$q_{rpAvg}$	$q_{ppAvg}$	$q_{prAvg}$	p,%
1	rytas	0,510	0,506	0,490	0,493	0,001	0,002	0,002	0,003	75,28
	diena	0,500	0,501	0,500	0,499	0	0	0	0	67,72
	vakaras	0,500	0,500	0,500	0,500	0	0	0	0	80,78
2	rytas	0,515	0,498	0,485	0,502	0,005	0,003	0,005	0,007	68,72
	diena	0,509	0,494	0,490	0,506	0,011	0,009	0,012	0,014	71,58
	vakaras	0,500	0,500	0,500	0,500	0,000	0,000	0,000	0,000	95,57
3	rytas	0,507	0,508	0,491	0,492	0,003	0,003	0,003	0,003	76,97
	diena	0,504	0,507	0,496	0,493	0,003	0,002	0,003	0,003	67,71
	vakaras	0,499	0,501	0,501	0,050	0,000	0,000	0,000	0,000	90,15
4	rytas	0,507	0,502	0,493	0,498	0,000	0,000	0,000	0,000	77,47
	diena	0,516	0,500	0,484	0,499	0,002	0,001	0,002	0,003	66,67
	vakaras	0,500	0,501	0,500	0,499	0,000	0,000	0,000	0,000	88,90



KD savybių vidurkiai su mažiausiomis skaidos reikšmėmis  $q$  gautos vidurkius skaičiuojant kiekvienai dienai (4.5 lent.). Mažos vidutinių KD savybių  $q$  reikšmės lėmė tikslesnį nuovargio klasių klasifikavimą. Geriausi klasifikavimo rezultatai, nepriklausomai nuo dalyvio, pasiekti vykdant „vakaras“ klasės klasifikavimą. Antram dalyviui siekė klasifikavimo tikslumas siekė 95,57 % ir buvo 25 % procentais aukštesnis, nei kad be vidutinių reikšmių.

Kitais dvejais atvejais (KD vidutinių reikšmių skaičiavimas pagal klavišų poras ir klavišų poras atitinkamai dienai) (4.6 lent. ir 4.7 lent.) KD savybių vidurkiai turėjo panašias skaidos koeficiento  $q$  reikšmes kaip ir klavišų dinamikos savybės ir klasifikavimo tikslumui svarios įtakos neturėjo. Fiksuoti -1 % – +5 % klasifikavimo tikslumo pokyčiai.

Geriausi klasifikavimo rezultatai gauti su papildomomis KD savybėmis, kurių skaida buvo nulis. Nulinė sklaida rodo, kad nefiksuoti parametrų svyravimai ar jie fiksuoti minimalūs. Parametrai su mažu svyravimu gali iškraipyti SVM klasifikavimo tikslumą. Tam, kad neiškreiptume tyrimo rezultatų, toliau bus tiriama visos trys papildomos savybių reikšmės atskirai.

#### 4.4. Trijų klasifikatorių sistemos testavimas

Trijų klasifikatorių sistema testuojama kiekvienam tyrimo dalyviui atskirai. Normalizavus dalyvio duomenis jie yra padalinami į treniravimo ir testavimo imtis. Santykis, koks duomenų kiekis naudojamas SVM treniravimui ir testavimui nurodomas programiškai.

Formuojant testavimo imtį, PĮ atsitiktine tvarka išrenka 20 % pradinės duomenų imties ir likusią imtį naudoja SVM klasifikatorių apmokymui. Mūsų tiriamu atveju testavimui naudojama 20 % visos duomenų imties duomenų. Iš pradžių papildomos vidutinės klavišų dinamikos reikšmės skaičiuojamos kiekvienai klavišų porai.

Ketvirto skyriaus pradžioje buvo minėta, jog viso bandymo metu, dalyviai rinko tris skirtingus tekstus. Norit įvertinti ar dalyvių renkamas tekstas turi įtakos nuovargio klasifikavimo tikslumui, žinodami datas, kada buvo renkami konkretūs tekstai, kiekvienam dalyvio vestam tekstui sudarome treniravimo ir testavimo imtis. Papildomai sudaroma ketvirta imtis iš visų rinktų tekstų duomenų.  $svm_{tk}$  klasifikatorių apmokymo rezultatai pateikti 4.8 lentelėje. Čia  $t$  – teksto identifikatorius,  $k$  – nuovargio klasė.

4.8 lentelė. SVM klasifikatorių apmokymo rezultatai

SVM \ Dalyvis	1	2	3	4
	$y$	$y$	$y$	$y$
$svm_{1r}$	84,12	85,05	83,77	84,60
$svm_{1d}$	84,62	85,83	83,83	83,45
$svm_{1v}$	85,47	85,44	85,13	86,79
$svm_{2r}$	84,47	86,63	86,17	83,41
$svm_{2d}$	86,78	85,79	85,01	75,54
$svm_{2v}$	86,99	86,99	85,88	73,38
$svm_{3r}$	75,50	77,08	76,66	-
$svm_{3d}$	77,48	75,95	75,90	-
$svm_{3v}$	77,48	75,76	74,76	-
$svm_{1,2,3r}$	91,48	91,51	92,19	89,75
$svm_{1,2,3d}$	92,75	93,01	92,83	89,79
$svm_{1,2,3v}$	92,80	93,34	93,02	90,27

Iš gautų SVM apmokymo rezultatų matyti, jog renkamas tekstas turi įtakos klasifikavimo tikslumui. Tekstai Nr. 1 ir Nr. 2 buvo rišlūs, tekstas Nr. 3 – nerišlus. Prognozuojant nuovargio klases iš pirmų dviejų tekstų, vidutinis klasių klasifikavimo tikslumas, pirmam, antram ir trečiam dalyviui siekė 85,44%. Klasifikuojant ketvirto dalyvio antrą tekstą fiksuotas prastesnis vakaro ir dienos nuovargio klasių klasifikavimo tikslumas ir atitinkamai siekė 73,38 %, 75,54 %. Klasifikuojant teksto Nr. 3 nuovargio klases, visiems dalyviams fiksuotas 10%  $\pm$  2% mažesnis klasifikavimo tikslumas ir vidutiniškai siekė 76,28 %. Trečias tekstas ketvirto dalyvio nebuvo vestas.

Geriausi klasifikavimo rezultatai pasiekti klasifikuojant nuovargio klases bendrai visiems tekstams. Klasifikavimo tikslumas vidutiniškai siekė 91,98%. Lyginant su antru ir trečiu tekstu buvo vidutiniškai 6,45 % didesnis.

Trijų klasifikatorių sistemos įvertinimui modeliuojama reali situacija, kai klasifikatorių  $svm_{tr}$ ,  $svm_{td}$  ir  $svm_{tv}$  įėjime paduodama duomenų imtis  $X_t$ , kurią sudaro vienos nuovargio klases duomenys, o išėjime sistema pateikia trimatį išėjimo vektorių  $Y_t(y_r; y_d; y_v)$ , kur  $y_k$  yra  $X_{tk}$  duomenų imties teisingai klasifikuotų duomenų skaičius, išreikštas procentais. Kiekvieno dalyvio rinkto teksto sistemos duomenų klasifikavimo rezultatai, pateikti 4.9, 4.10, 4.11 ir 4.12 lentelėse.

**4.9 lentelė.** Trijų klasifikatorių sistemos testavimo rezultatai su teksto Nr. 1 duomenimis, kai naudojama papildomos vidutinės klavišų poros KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	72,26	0,00	0,00	89,24	0,00	1,36	69,86	0,70	0,00	80,00	0,00	0,00
$y_d$	0,00	68,87	0,00	1,27	79,37	0,00	0,68	73,94	0,00	0,00	69,93	0,00
$y_v$	0,00	0,00	77,70	0,63	0,00	85,71	0,00	0,00	81,56	0,00	0,00	80,61

**4.10 lentelė.** Trijų klasifikatorių sistemos testavimo rezultatai su teksto Nr. 2 duomenimis, kai naudojama papildomos vidutinės klavišų poros KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	78,05	0,00	0,00	80,37	0,00	0,83	77,68	0,00	0,00	3,33	0,00	0,00
$y_d$	0,00	80,16	0,00	0,00	72,65	0,00	0,00	78,99	0,00	0,00	53,68	0,88
$y_v$	0,81	1,59	84,62	0,00	0,00	77,69	0,00	0,00	68,29	3,33	0,00	52,21

**4.11 lentelė.** Trijų klasifikatorių sistemos testavimo rezultatai su teksto Nr. 3 duomenimis, kai naudojama papildomos vidutinės klavišų poros KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	62,96	0,00	0,00	72,06	0,00	0,00	82,26	0,00	0,00	–	–	–
$y_d$	0,00	75,00	0,00	0,00	83,56	0,00	0,00	69,62	0,00	–	–	–
$y_v$	0,00	0,00	69,77	1,47	2,74	75,36	0,00	0,00	60,00	–	–	–

**4.12 lentelė.** Trijų klasifikatorių sistemos testavimo rezultatai su teksto Nr. 4 duomenimis, kai naudojama papildomos vidutinės klavišų poros KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	86,71	0,00	0,00	99,37	21,88	14,01	88,29	0,00	0,00	89,42	0,00	0,00
$y_d$	0,00	89,42	0,00	0,63	87,50	0,00	0,33	93,69	0,00	0,53	86,57	0,00
$y_v$	0,00	0,00	87,79	0,00	0,00	90,88	0,00	0,00	89,86	0,53	0,00	95,00

Iš gautų duomenų matyti, jog dauguma  $X_k$  imties duomenis geriausiai klasifikuoja atitinkamos nuovargio klases  $svm_k$  klasifikatorius, t.y. padavus sistemai ryte rinkto teksto duomenų imtį  $X_r$ , išėjimo vektorių  $Y$  reikšmė  $y_r$ , lyginant ją su  $y_d$  ir  $y_v$ , turi didžiausią reikšmę. Atsižvelgus į gautus rezultatus ir slenkstinės funkcijos reikšmę  $sl_2$ , nustatius 50 %, fiksuojama viena neapibrėžtumo būseną, kuri gaunama klasifikuojant ketvirto dalyvio teksto Nr. 2 duomenis. Šiuo atveju  $svm_{r_2}$  ir  $svm_{v_2}$  teisingai klasifikavo tik 3,33 %  $X_r$  testavimo imties duomenų.

Taikant slenkstinę funkciją vektorių  $Y$  reikšmėms, galimos dvi duomenų klasifikavimo situacijos:

- slenkstinę reikšmę  $sl_2$  viršija vienas arba daugiau  $svm$  klasifikatorių;
- slenkstinės reikšmės  $sl_2$  neviršija nei vienas  $svm$  klasifikatorius.

Esant pirmai situacijai, nuovargio būseną priskiriama pagal maksimalią  $y_k$  reikšmę. Antroje situacijoje gaunamas neapibrėžtumas ir atsakymas nėra pateikiamas.

Tiksliausiai *svm* klasifikatoriai skyrė nuovargio klases, kai mokymo imtis buvo sudaroma iš visų dalyvio rinktų tekstų. Su šia duomenų imtimi apmokyti *svm* klasifikatoriai nuovargio klasių klasifikavimo tikslumas vidutiniškai buvo 21,2 % didesnis. Kadangi, *svm* klasifikatoriai tiksliausiai klasifikavimo, kai mokymo imčiai buvo naudojami visų tekstų duomenys, toliau tyrimas atliekamas neskirsiant duomenų pagal suvedamą tekstą.

4.13 ir 4.14 lentelėse pateikta bendro teksto (tekstas Nr. 4) sistemos klasifikavimo rezultatai, atitinkamai naudojant vidutinės dienos\būsenos ir vidutinės kiekvienos dienos klavišų poros\būsenos KD savybės.

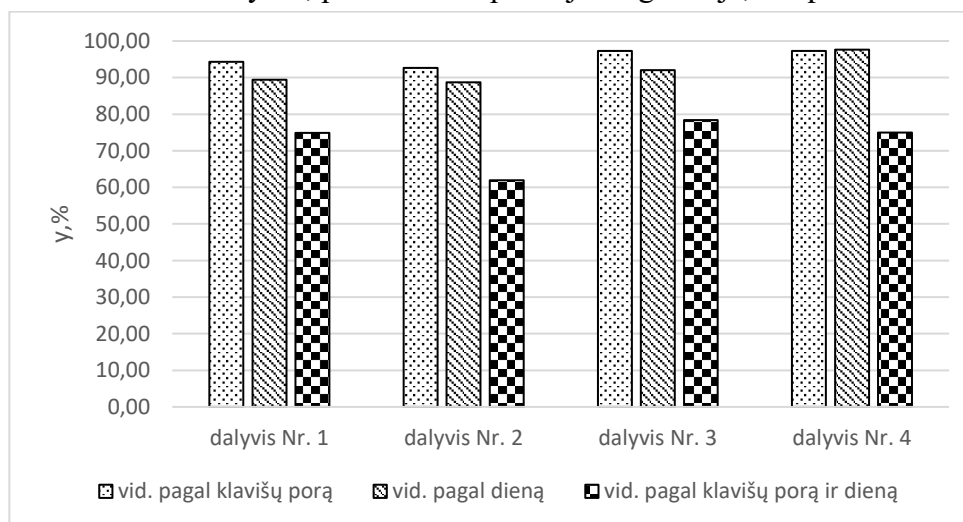
**4.13 lentelė.** Trijų klasifikatorių sistemos testavimo rezultatai, kai naudojama papildomos, vidutinės, dienos \būsenos KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	92,22	5,30	0,21	88,75	9,03	3,92	90,86	0,63	0,81	93,87	2,25	0
$y_d$	3,65	85,18	0,31	2,99	82,50	0,78	0,74	93,71	0,00	11,08	99,44	7,54
$y_v$	0,38	0,88	90,70	0,20	0,33	94,81	0,32	0	91,60	0	0	99,71

**4.14 lentelė.** Trijų klasifikatorių sistemos testavimo rezultatai, kai naudojama papildomos, vidutinės, kiekvienos dienos klavišų poros \būsenos KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	60,48	0,38	0,41	54,63	1,15	0,89	64,65	0,11	0,66	70,48	0,82	0,50
$y_d$	23,45	99,52	24,87	0,96	64,52	0,10	24,39	99,15	21,92	12,50	93,17	14,39
$y_v$	1,66	1,53	64,56	0,00	0,23	66,60	0,31	0,43	71,48	0,90	1,09	61,33

Trijų klasifikatorių sistemos būsenos nustatymo tikslumo palyginimas, naudojantis skirtingai apskaičiuotas vidutines KD savybes, pateiktas stulpelinėje diagramoje, 4.5 paveiksle.



**4.5 pav.** Klasifikatorių sistemos būsenų nustatymo tikslumo palyginimas, naudojant skirtingai apskaičiuotas vidutines KD savybes

Sistema nuovargio būsenas, naudojant vidutines dienos\būsenos KD savybių reikšmes, klasifikavo vidutiniškai 94,65 % tikslumu. Prasčiausias tikslumas buvo fiksuotas naudojant vidutinės kiekvienos dienos klavišų poros\būsenos KD savybės ir siekė 72,02 %.

#### 4.5. Klavišų porų klasifikatorių sistemos testavimas

Klavišų porų klasifikatorių sistema, modifikuoja 4.4 skyriuje tirtą sistemą. Šiuo atveju klasifikatoriai sudaromi kiekvienai klavišų porai atskirai su sąlyga jog klasifikatoriaus mokymo imtyje

yra 10 arba daugiau įrašų. Pagal duomenų imtyje esamas klavišų poras suformuojamas nuovargio klasę reprezentuojantis klasifikatorių masyvas  $T_k$ , iš kurio, atrenkami  $svm$  klasifikatoriai, kurių klasifikavimo tikslumas viršija slenkstinę  $sl_1 = 90$  vertę. 4.15 lentelėje pateikta, kiek klavišų porų klasifikatorių naudojama atitinkamam dalyvio rinktam tekstui prognozuoti nuovargio būsenas.

**4.15 lentelė.** Klavišų porų klasifikatorių skaičius prieš ir po KP įvertinimo

Dalyvis	1		2		3		4	
$Svm$ skaičius Imtis	Be $sl_1$	Su $sl_1$	Be $sl_1$	Su $sl_1$	Be $sl_1$	Su $sl_1$	Be $sl_1$	Su $sl_1$
$X_r$	333	208	313	214	305	195	287	250
$X_d$	333	187	313	199	305	184	287	184
$X_v$	333	193	313	209	305	188	287	176

Pradinis  $T_k$  masyvo dydis, dėl pradinės sąlygos, jog  $svm$  klasifikatorius sudaromas tik toms klavišų poroms, kurių mokymo imtyje yra 10 arba daugiau įrašų, dalyviams skiriasi. Pritaikius klasifikatoriams slenkstinę  $sl_1$  reikšmę, atrenkamos klavišų poros, kurių KD savybėms nuovargis turėjo didžiausią įtaką ir atmetamos tos klavišų poros, kurioms nuovargis neturėjo svarios įtakos. Atrinkus klasifikatorius masyvo  $T$  dydis vidutiniškai sumažėjo 110 klasifikatorių.

Sistemos efektyviu įvertinti įvedama antra slenkstinė vertė  $sl_2 = 60$ . Antra slenkstinė funkcija nustato ar  $svm$  klasifikatorius pakankamai tiksliai klasifikavo klavišų poras testavimo duomenis, kad būtų galima laikyti, jog ši klavišų pora priklauso konkrečiai nuovargio klasei. Sistemos klasifikavimo rezultatai, pagal atitinkamas papildomas KD vidutines savybių reikšmes, pateikti 4.16, 4.17 ir 4.18 lentelėse.

**4.16 lentelė.** Klavišų porų klasifikatorių sistemos testavimo rezultatai, kai naudojama papildomos, vidutinės, klavišų poros būsenos KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	96,75	0	0	94,78	0	0	95,18	0	0	93,17	0	0
$y_d$	0	92,57	0	0	96,91	0	0	93,28	0	0	95,87	0
$y_v$	0	0	95,59	0,57	0	97,28	0	0	98,16	0	0	98,21

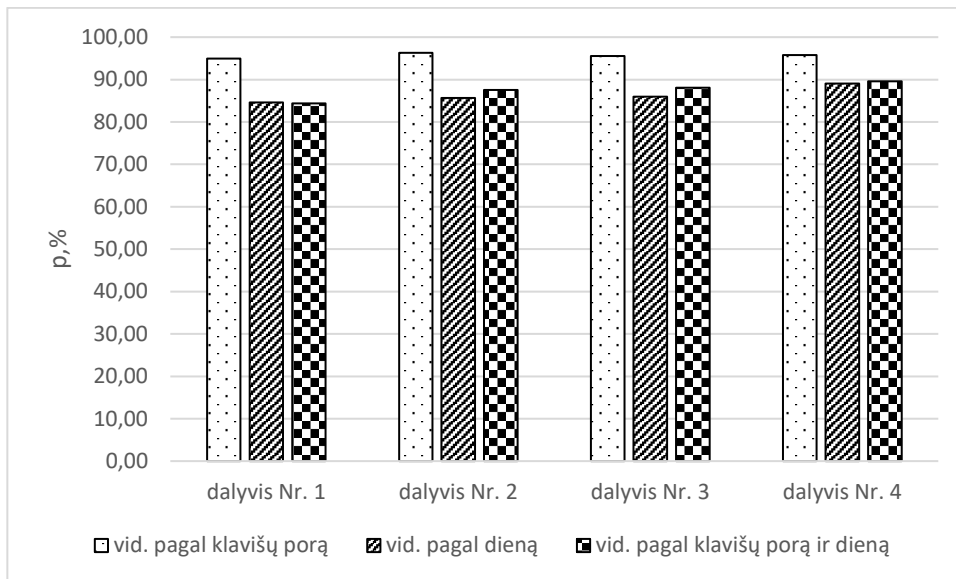
**4.17 lentelė.** Klavišų porų klasifikatorių sistemos testavimo rezultatai, kai naudojama papildomos vidutinės dienos būsenos KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	83,42	0,56	0,00	87,00	0,00	0,00	87,70	0,56	0,00	86,11	0,00	1,24
$y_d$	0,51	83,71	0,00	1,51	90,91	0,00	0,53	86,19	0,00	0,00	91,11	0,00
$y_v$	0,00	0,00	85,95	1,02	0,00	84,69	0,00	0,00	90,27	0,00	0,00	91,62

**4.18 lentelė.** Klavišų porų klasifikatorių sistemos testavimo rezultatai, kai naudojama papildomos, vidutinės, kiekvienos dienos klavišų poros būsenos KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$	$X_r$	$X_d$	$X_v$
$y_r$	85,65	0,00	0,00	85,79	0,00	0,00	85,15	0,00	1,09	92,58	1,24	0,00
$y_d$	0,89	81,67	0,00	1,55	85,98	0,00	2,50	83,16	0,00	0,81	86,47	0,00
$y_v$	1,35	0,00	86,59	0,58	0,00	85,14	2,45	0,00	89,56	0,87	0,00	88,24

Klavišų porų klasifikatorių sistemos būsenos nustatymo tikslumo palyginimas, naudojant skirtingai apskaičiuotas vidutines KD savybes, pateiktas stulpelinėje diagramoje, 4.6 paveiksle.



**4.6 pav.** Klavišų porų klasifikatorių sistemos būsenų nustatymo tikslumo palyginimas, naudojant skirtingai apskaičiuotas vidutines KD savybes

Geriausias nuovargio klasių klasifikavimo rezultatas pasiektas naudojant vidutinės klavišų porų būsenų KD savybes, ir vidutiniškai buvo 95,65%. klasifikavimo rezultatai, naudojant vidutines dienos būsenos ir vidutines kiekvienos dienos klavišų poros būsenos KD savybes atitinkami buvo 86,33% ir 87,39%.

Sistemos testavimo metu buvo stebima, kurie KP SVM klasifikatoriai tiksliausiai identifiko nuovargio klases. Priklausomai nuo prognozuojamos nuovargio klasės ir dalyvio, KP klasifikatorių masyvo dydis, kito. Visuose testavimo rezultatuose pasikartojusios klavišų poros pateiktos 4.19 lentelėje.

**4.19 lentelė.** Klavišų poros, kurių SMV klasifikatoriai tenkino  $sl_2$  slenkstinės funkcijos reikšmę.

,-Tarpas; -Tarpas; 7-Tarpas; A-= +; A-D; A-I; A-L; A-M; A-N; A-R; A-S; A-Tarpas; A-U; B-I; D-A; E-I; E-K; E-R; E-S; E-T; E-Tarpas; G-A; G-E; G-U; I-6; I-A; I-E; I-K; I-M; I-N; I-R; I-S; I-Tarpas; J-A; J-E; J-O; J-U; K-A; K-I; L-T; M-A; M-E; M-I; M-O; N-E; N-U; O-J; O-M; O-S; O-Tarpas; P-A; P-I; R-A; R-I; R-Tarpas; S-I; S-T; S-Tarpas; S-U; Tarpas-= +; Tarpas-A; Tarpas-D; Tarpas-G; Tarpas-I; Tarpas-J; Tarpas-K; Tarpas-L; Tarpas-M; Tarpas-N; Tarpas-P; Tarpas-S; Tarpas-Shift; Tarpas-T; Tarpas-V; T-I; T-O; T-U; U-O; U-S; U-Tarpas; U-V; V-A; V-E;
--

Iš viso rastos 83 KP, kurios pasikartoto visuose testavimo rezultatuose. KP lentelėje atskirtos kabliataškiais, o klavišų pora pateikiama per brūkšnį. Kadangi visų šių porų klasifikatorių klasifikavimo tikslumas, visais atvejais viršijo  $sl_2$  reikšmę, galima teigti, jog šių porų KD geriausiai reprezentavo nuovargio daromą poveikį teksto rinkimo, su kompiuterio klaviatūra, stiliui.

#### 4.6. Trijų klasifikatorių sistemos, su klavišų porų svarbumo įvertinimu, testavimas

Iš 4.5 skyriuje gautų rezultatų galima teigti, kad ne visų klavišų porų *svm* klasifikatoriai tiksliai klasifikuoja nuovargio būsenas. Klavišų poras, kurių klasifikatorių klasifikavimo tikslumas netenkina slenkstinės reikšmės  $sl_2$ , galime priskirti prie sistemos triukšmų, kurie mažina sistemos tikslumą.

Norit įvertinti, kokią įtaką trijų klasifikatorių sistemai daro triukšmai, panaudojama jungtinė sistema aprašyta 3.9.3 skyriuje. Ši sistema, prieš klasifikatorių mokymo imties sudarymą, įvertina, kaip tiksliai galima atpažinti nuovargio klases su kiekviena klavišų pora. Klavišų pora, pagal kurią apmokyto atraminių vektorių mašinos klasifikavimo tikslumas neviršijo  $sl_2$  slenkstinės reikšmės, laikomas sistemos triukšmu. Sudarant mokymo imtį visos klavišų poros priskirtos triukšmams neįtraukiamos į mokymo ir testavimo imtis. Sistemos klasifikavimo rezultatai, pagal papildomas, vidutines, KD savybių reikšmes, pateikti 4.20, 4.21 ir 4.22 lentelėse.

**4.20 lentelė.** Trijų klasifikatorių sistemos, su klavišų porų svarbumo įvertinimu, testavimo rezultatai, kai naudojama papildomos, vidutinės, klavišų poros\būsenos KD savybės.

Dalyvis	1			2			3			4		
Imtis SVM	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>
$y_r$	99,63	0	0	93,91	1,40	0,10	99,78	0	0	98,75	0	0
$y_d$	0	89,90	0	4,91	97,79	5,54	0	99,89	0,00	0	99,45	0
$y_v$	0	0	99,02	0	0	99,69	0	0,11	99,52	0	0,28	100

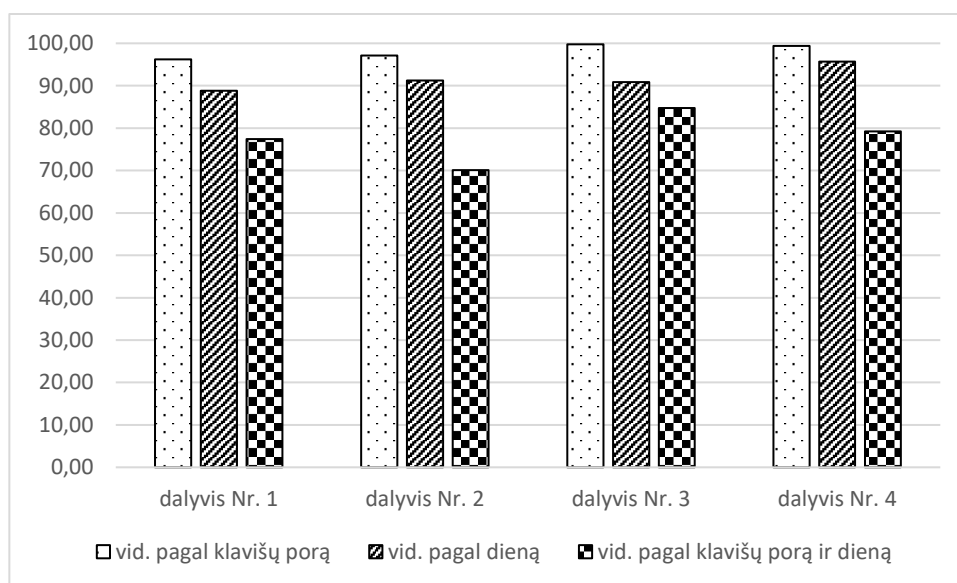
**4.21 lentelė.** Trijų klasifikatorių sistemos su klavišų porų svarbumo įvertinimu testavimo rezultatai, kai naudojama papildomos, vidutinės, dienos\būsenos KD savybės

Dalyvis	1			2			3			4		
Imtis SVM	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>
$y_r$	92,41	6,41	0,38	94,31	14,01	7,21	93,64	1,77	0	93,68	1,49	0
$y_d$	2,74	86,22	2,63	6,58	85,99	1,32	2,93	87,45	0,13	12,17	99,67	5,73
$y_v$	0,95	1,45	87,88	0,11	0,28	93,51	0	0,25	91,49	0	0,00	93,69

**4.22 lentelė.** Trijų klasifikatorių sistemos su klavišų porų svarbumo įvertinimu testavimo rezultatai, kai naudojama papildomos, vidutinės, kiekvienos dienos klavišų poros\būsenos KD savybės.

Dalyvis	1			2			3			4		
Imtis SVM	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>	X <sub>r</sub>	X <sub>d</sub>	X <sub>v</sub>
$y_r$	56,99	0	0	59,93	0,39	0,24	71,73	0,26	0,28	73,40	0,16	0
$y_d$	0,42	75,80	0	10,92	98,83	10,02	41,23	99,62	38,73	33,84	92,22	34,57
$y_v$	14,93	14,20	99,39	0,25	0,13	51,43	0,25	0,26	82,99	0,17	0,16	72,14

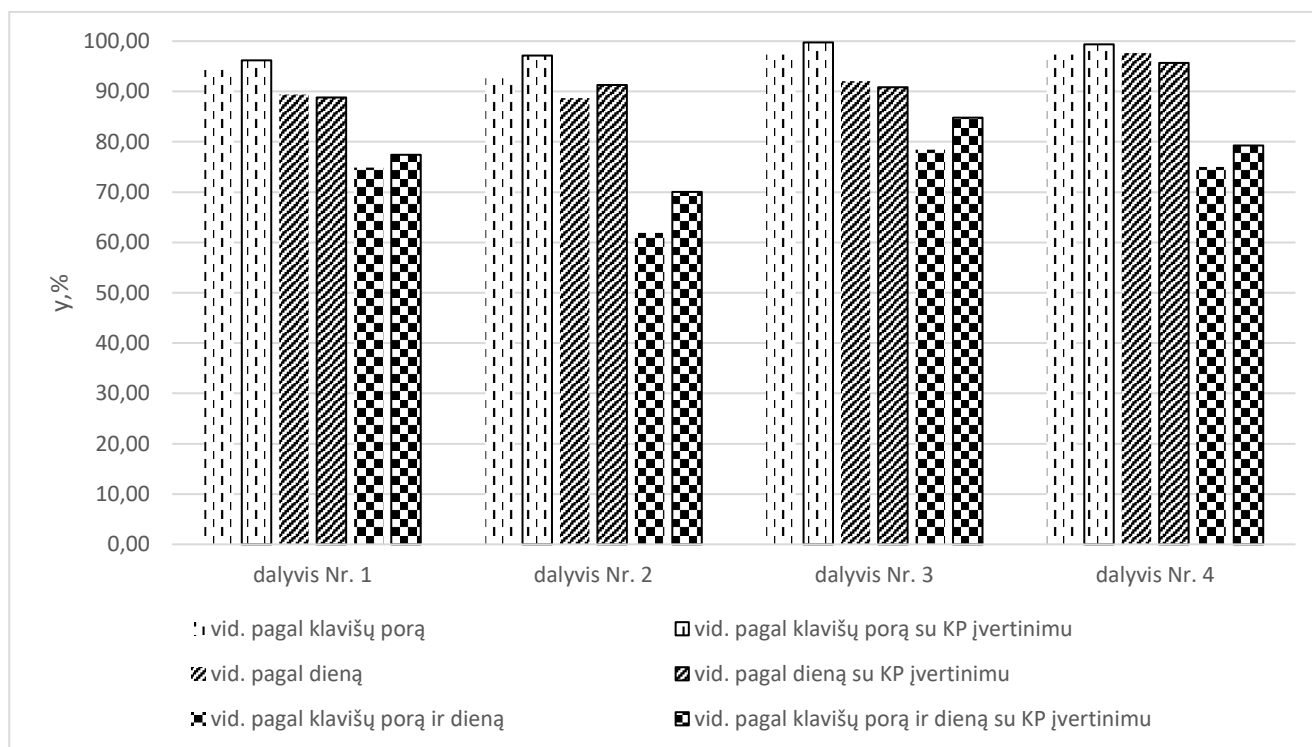
Apibendrintas, visų nuovargio būsenų, kiekvieno dalyvio sistemos klasifikavimo tikslumas, pagal papildomas KD savybių tipus, pateiktas stulpeline diagrama, 4.7 paveiksle. Geriausi rezultatai pasiekti naudojant papildomas vidutines klavišų porų KD savybes ir vidutiniškai buvo 98,11%, mažiausias tikslumas 77,87% pasiektas naudojant papildomas klavišų porų – dienų vidutines KD savybes



**4.7 pav.** Trijų klasifikatorių sistemos, su klavišų porų svarbumo įvertinimu, būsenų nustatymo tikslumo palyginimas, naudojant skirtingai apskaičiuotas vidutines KD savybes.

4.8. paveiksle pateikta grafinis trijų klasifikatorių sistemos su klavišų porų svarbumo įvertinimu be palyginimas. Sistema, kuri duomenų klasifikavime naudoja papildomas klavišų porų arba kiekvienos dienos klavišų porų vidutines KD savybes, pašalinus nesvarbias klavišų poras, nuovargio

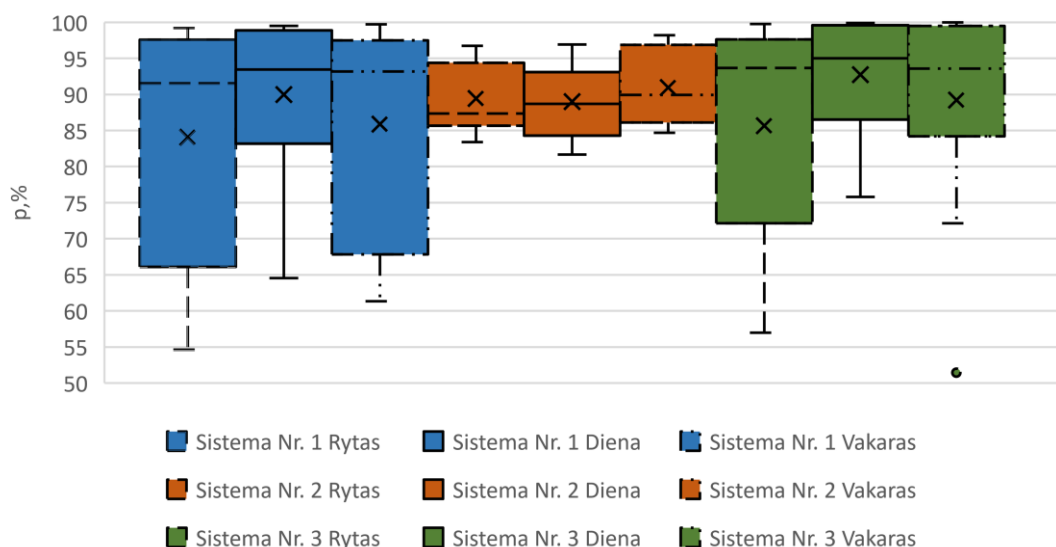
būsenas klasifikavo, atitinkamai 2,73% ir 5,32%, tiksliau. Naudojant vidutines dienų KD savybes matomas 0,28% tikslumo sumažėjimas.



4.8 pav. Trijų klasifikatorių sistemos, su klavišų porų vertinimu ir be, tikslumo palyginimas

#### 4.7. Tyrimo rezultatų apibendrinimas ir apribojimai

Tyrimo rezultatus galima analizuoti iš dviejų perspektyvų: nuovargio nustatymo sistemos ir asmens. Analizuojant sistemos rezultatus, koku tikslumu buvo identifikuojamos konkrečios nuovargio klasės, nevertinamas skirtingų dalyvių faktorius. Tokiu būdu galime matyti, sistemos universalumą pritaikant ją skirtingiems vartotojams. 4.9 paveiksle pateikta tirtų sistemų statistiniai būsenos nustatymo įverčiai.



4.9 pav. Statistiniai būsenų nustatymo įverčiai.

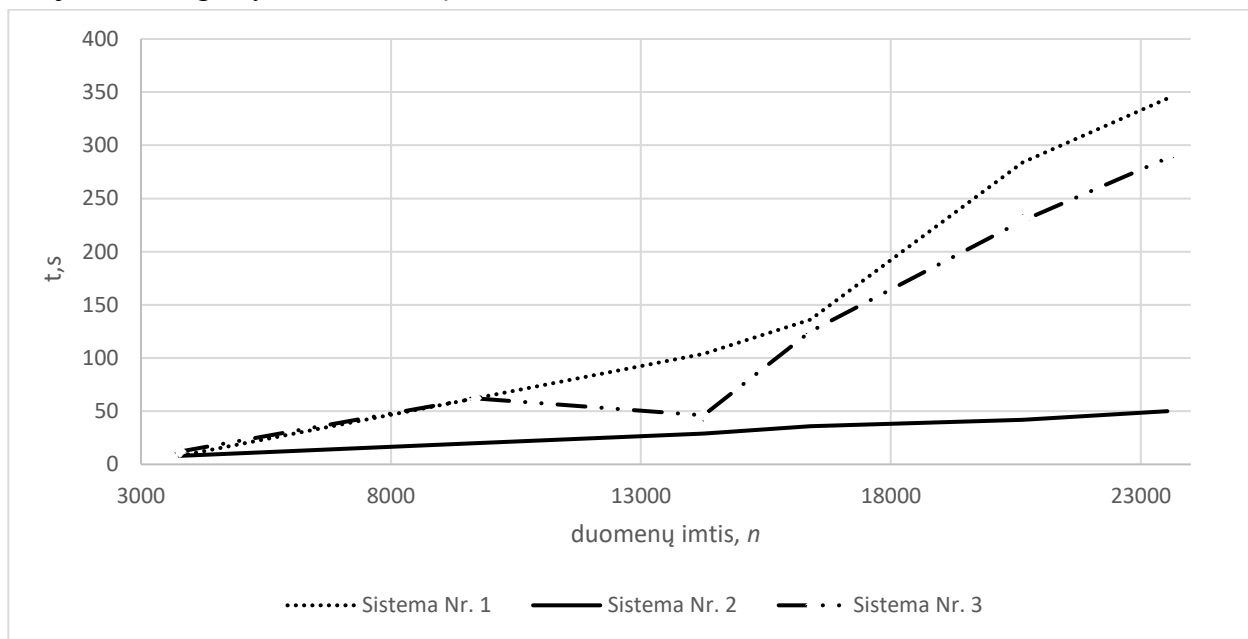
Sistemą Nr.1 atitinka trijų klasifikatorių sistema, sistemą Nr. 2 – klavišų porų klasifikatorių sistema, sistemą Nr. 3 – trijų klasifikatorių sistema su KP svarbumo įvertinimu. Taip pat, 4.22 lentelėje pateikta sistemų atitinkamos nuovargio būsenos klasifikavimo standartinis nuokrypis, mediana ir vidurkis.

**4.23 lentelė.** Sistemų atitinkamos nuovargio būsenos klasifikavimo standartinis nuokrypis, mediana ir vidurkis

	Sistema Nr. 1			Sistema Nr. 2			Sistema Nr. 3		
Standartinis nuokrypis, %	16,56	10,46	15,07	<b>4,68</b>	<b>5,12</b>	<b>5,16</b>	28,52	4,70	11,76
Mediana, %	91,54	93,44	93,20	87,35	88,69	85,54	90,86	94,45	92,66
Vidurkis, %	84,06	90,00	85,87	89,44	88,98	90,94	79,45	93,57	88,47

Kaip matyti iš pateiktų rezultatų, Sistemai Nr.2 fiksuoti mažiausi standartiniai nuokrypiai, kurie rytui, dienai ir vakarui atitinkamai buvo 4,68 %; 5,12 %; 5,16 %. Taip pat medianos ir vidurkio reikšmės yra artimos viena kitai. Visa tai lemia sistemos nuovargio identifikavimo tikslumo pastovumą ir nuovargio būsenos identifikavimo patikimumą, kuris pagal mūsų gautus rezultatus yra 89,79 % su galimu 4,99 % svyravimu.

Svarbus sistemos parametras yra jos greitimeika. Sistemų testavimui buvo naudotas *Lenovo Ideapad-700* su Intel i7 -6700HQ procesoriumi, 8GB operatyvinės atminties ir 1TB, 5400aps/min, kietuoju disku. Mūsų tirtuose sistemose didžiausi laiko resursai buvo reikalingi sistemų apmokymo metu. Kadangi, kiekvienu atveju reikėjo apmokyti tris SVM klasifikatorius, laikas, reikalingas apmokymui, didėjo tris kartus. 4.10 paveiksle pavaizduota sistemų greitimeikos apmokymo laikai, naudojant skirtingo dydžio duomenų imtis.

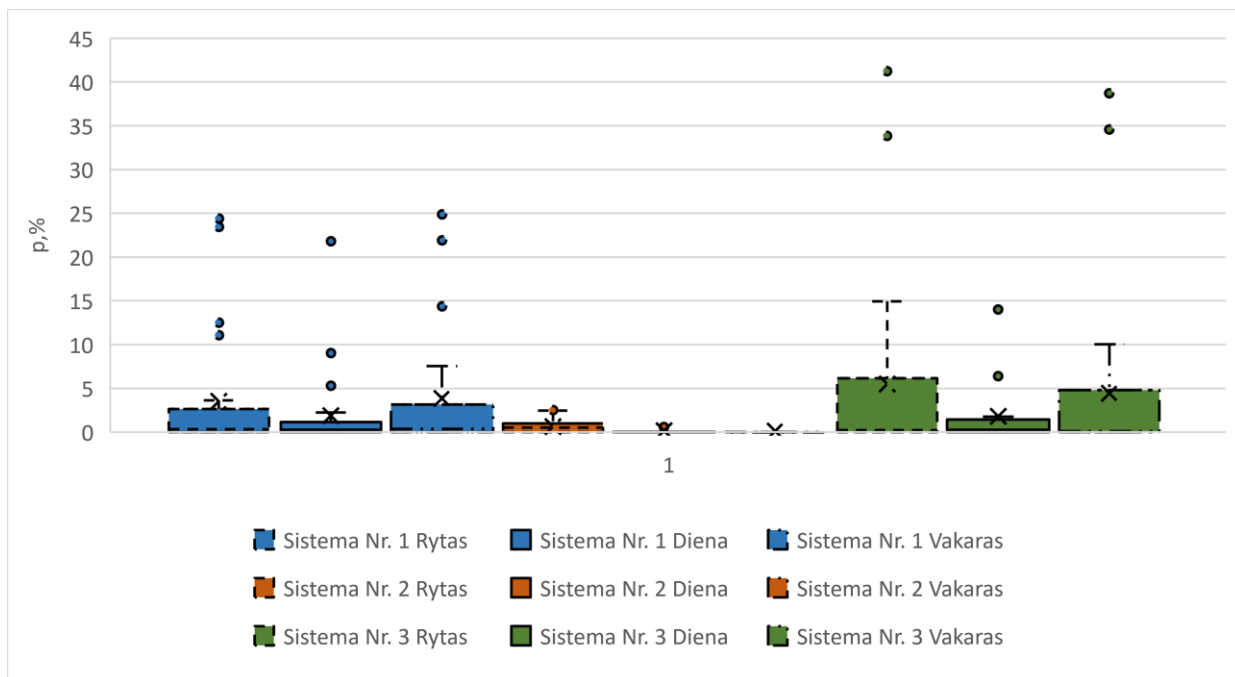


4.10 pav. Sistemų greitimeikos apmokymo laikai, naudojant skirtingo dydžio duomenų imtis

Kaip matyti iš aukščiau pateikto grafiko laikas, reikalingas apmokyti pirmą ir trečią sistemą, didėja eksponentiškai ir norint sistemą apmokyti su dviejų savaičių ( $n = 23532$ ) duomenų imtimi, mokymas atitinkamai užtrunka 344 s ir 288 s. Apmokymo laikas Sistemai Nr. 2 didėja tiesiškai. Sistemos apmokymas su dviejų savaičių imtimi užtrunka 50 s.

Svarbu aptarti kokie klaidų nuokrypiai buvo fiksuojami klasifikuojant atitinkamas klases. 4.11 paveiksle pateikta tirtų sistemų klaidų nuokrypio įverčiai.





4.11 pav. Sistemų klaidos nuokrypio įverčiai

4.23 lentelėje pateikta sistemų nuovargio būsenų klasifikavimo klaidos nuokrypių standartinis nuokrypis, mediana ir vidurkis.

4.24 lentelė. Sistemų būsenos klasifikavimo klaidos nuokrypio standartinis nuokrypis, mediana ir vidurkis

	Sistema Nr. 1			Sistema Nr. 2			Sistema Nr. 3		
Standartinė nuokrypa, %	7,08	4,70	7,38	<b>0,77</b>	<b>0,29</b>	<b>0,33</b>	10,81	4,02	10,31
Mediana, %	0,35	0,36	0,36	0,52	0,00	0,00	0,25	0,25	0,12
Vidurkis, %	3,53	1,92	3,85	0,63	0,10	0,10	5,52	1,79	4,45

Mažiausios medianų, standartinių nuokrypų ir vidurkių reikšmės visų nuovargio klasių nustatymui fiksuotos sistemai Nr. 2. Apskaičiuota standartinis nuokrypis rytui, dienai ir vakarui atitinkamai buvo 0,77 %; 0,29 %; 0,33 %, o vidurkiai – 0,63 %; 0,10 %; 0,10 %. Atsižvelgiant į tai, jog sistema Nr. 2 tikslines nuovargio klases identifiko 89 % tikslumu su 0,28 % klaidos nuokrypiu.

Pašalinus nesvarbias klavišų poras, iš mokymo ir testavimo imčių, dviejų nuovargio klasių („diena“, „vakaras“) nustatymo tikslumo padidėjimas, tačiau tuo pat metu fiksuotas ir netinkamo nuovargio klasių identifikavimo didėjimas. Kaip matyti iš 4.10 paveikslo keturis sykius buvo fiksuotas 30% viršijantis netinkamų klasių identifikavimas. Tai lemia, kad sistema Nr. 3 sunkiau skiria nuovargio klases viena nuo kitos.

Individualių eksperimento dalyvių nuovargio klasių nustatymo tikslumo rezultatų standartinis nuokrypis, medianos ir vidurkiai pateikti 4.24 lentelėje. 4.25 lentelėje pateikta netinkamo nuovargio klasių klasifikavimo klaidos nuokrypio standartinės nuokrypos, medianos ir vidurkiai.

4.25 lentelė. Eksperimento dalyvių nuovargio klasių nustatymo tikslumo rezultatų standartinis nuokrypis, medianos ir vidurkiai

Dalyvis	1			2			3			4		
Standartinis. Nuokrypis, %	15,85	7,20	10,77	15,92	11,41	16,45	11,86	6,84	8,74	10,38	4,43	13,55
Mediana, %	88,93	85,70	89,29	87,87	88,45	89,33	89,28	93,50	90,88	92,87	94,52	92,65
Vidurkis, %	83,44	86,98	89,62	82,39	87,13	85,68	86,09	93,38	90,30	87,76	94,81	89,13

**4.26 lentelė.** Eksperimento dalyvių standartinio nuokrypio, medianos ir vidurkių klaidos nuokrypiai.

Dalyvis	1			2			3			4		
Standartinis. Nuokrypis, %	6,21	3,81	5,84	2,90	3,97	4,31	10,84	0,45	10,17	8,69	0,70	8,63
Mediana, %	0,46	0,19	0,00	0,70	0,18	0,17	0,28	0,11	0,00	0,08	0,08	0,00
Vidurkis, %	2,83	1,92	1,60	1,83	1,68	2,51	4,21	0,27	3,53	4,03	0,47	3,55

Dalyvių nuovargio klasių identifikavimo vidurkio standartinis nuokrypis priklausomai nuo dalyvių svyruoja tarp 4,43 % ir 15,92 %. Remiantis gautais rezultatai, galima teigti, kad nuovargio klasė „diena“, visiems eksperimento dalyviams, buvo atpažįstama tiksliausiai. Tikslinė nuovargio klasė buvo nustatoma 90,57 % tikslumo, o netinkamų klasių identifikavimo tikslumas buvo 1,09 %.

Atliktas tyrimas turi tris apribojimus, kurie galėjo daryti įtaką, galutiniams tyrimo rezultatams. Eksperimento dalyvių skaičius (keturi dalyviai), jo atlikimo laikas (10 darbo dienų) ir priimta nuovargio nustatymo prielaida. Mažas dalyvių skaičius apribojo sistemų elgsenos tyrimą. Trumpas teksto rinkimo laikas apribojo SVM klasifikatorių apmokymo ir testavimo tyrimą.

Didesnė duomenų imtis leistų tirti imties įtaką nuovargio klasifikavimui, taip pat, smulkinti, dabar naudojamus tris, nuovargio intervalus. Tai leistų rasti optimalų nuovargio nustatymo intervalą ir duomenų imtį, kurios reikia identifikuoti nuovargio klases.

Taip pat nuovargio nustatymo rezultatus gali iškreipti priimta prielaida, kad laikas koreliuoja su nuovargiu. Idealiomis aplinkybėmis ši prielaida yra teisinga, tačiau dalyvis, ryte ar dieną, gali jaustis taip pat pavargęs, kaip ir vakare. Dalyvio savijautos apklausa, prieš pradėdant rinkti tekstą, leistų realiau įvertinti, kokią nuovargį asmuo jaučia.

## 5. IŠVADOS

1. Atlikus metodų apžvalgą nustatyta, jog esami nuovargio nustatymo metodai, naudoja specifines užduotis, kurios dažniausiai neturi nieko bendro su tiesiogine asmens veikla. Tokie nustatymo metodai neleidžia realiu laiku sekti ir identifikuoti asmens nuovargio.
2. Klavišo paspaudimo charakteristikos yra skirstomos į dvi grupes – klaviatūros klavišų ir išorinių prietaisų. Klaviatūros charakteristikos matuoja klavišų paspaudimo laiką, tuo tarpu išoriniai prietaisai gali fiksuoti klavišo spaudimo slėgį, pirštą kuriuo spaudžiamas klavišas ir pirštu liečiamą klavišo dalį. Esminiai požymiai, kurie geriausiai charakterizuoja subjekto rašymo šabloną yra klavišų paspaudimo, atleidimo – paspaudimo, paspaudimo – paspaudimo, atleidimo – atleidimo ir paspaudimo – atleidimo laikai.
3. Atlikta klasifikavimo metodų analizė parodė, kad atraminių vektorių mašinos, apmokomos su ta pačia imtimi, duomenis klasifikuoja vienodai. SVM klasifikavimo tikslumas priklauso nuo  $C$  ir  $\gamma$  reikšmių. Pasiiekti maksimalų klasifikavimo tikslumą, skirtingoms duomenų imtims, reikalinga parametrų poros paieška. Tyrimo metu buvo nustatyta, jog SVM prognozavimo tikslumas, intervaluose  $C = [C^4; C^{12}]$  ir  $\gamma = [\gamma^{13,5}; \gamma^{15}]$ , patenka į lokalius maksimumus. Globaliam prognozavimo tikslumui ir parametrų  $C$  ir  $\gamma$  porai nustatymui, kiekvienai mokymo duomenų imčiai, reikalinga tinklo paieška su  $y_0^{\ln(x_0)+0,5n}$  parametrų kitimo intervalu.
4. Tyrime testuoti trys nuovargio būsenos nustatymo metodai. Geriausi rezultatai fiksuoti sistemai tada, kai SVM klasifikatoriai sudaromi kiekvienai klavišų porai atskirai. Klasifikatorių sudarymas klavišų poroms leidžia įvertinti ar konkreti KP pora yra svarbi nustatant nuovargio klasę. Tyrimo metu vidutiniškai 110 klavišų porų netenkino  $sl_1 = 90$  % duomenų klasifikavimo tikslumo slenkstinės vertės ir buvo išimamos iš klasifikatorių masyvo. Visais sistemos testavimo atvejais fiksuotos 83 klavišų poros, kurių klasifikatoriai, nuovargio klases, klasifikavo didesniu nei 60 % tikslumu.
5. Pagal atliktą sistemų greitaveikos tyrimą nustatyta, jog duomenų imties dydis lėtina sistemos greitaveiką. Mažiausiai įtakos, didėjanti duomenų imtis, turėjo klavišų porų klasifikatorių sistemai. Ši sistema, lyginant ją su kitomis dvejomis, naudojanti dvejų sravaičių duomenimis, apsimokė apie šešis kartus greičiau ir užtruko 50s.
6. SVM Klasifikatorius sudarant klavišų poroms vidutinis nuovargio būsenos atpažinimo tikslumas: ryte – 89,44 %, dieną – 88,98 %, vakare – 90,94 % su atitinkamu 4,68 % 5,12 % ir 5,16 % standartiniu nuokrypiu.
7. Gauti rezultatai leidžia įvertinti žmogaus nuovargį dirbant su kompiuteriu susijusį darbą dienos eigoje ir gali būti panaudojami keliant darbo biure kokybę.

## 6. LITERATŪROS ŠALTINIAI

1. MITLER, M., Carskadon, M.A. ir kt. *Catastrophes, sleep, and public policy: Consensus report*, HHS Author Manuscripts, 1988, 11(1), p. 100–109.
2. LEGER, D. *The cost of sleep-related accidents: A report for the national commission on sleep disorders research*, Sleep., 1994 17(1), p. 84–93.
3. KIM, S., Cranor, B.D. and Ryu, Y.S. *FATIGUE: WORKING UNDER THE INFLUENCE*, Occupational Ergonomics and Safety Conference, 2009 ISBN # 97809652558-4-4, p. 317–322.
4. FISHER, S. Monk H. *Hours of work: Temporal factors in work-scheduling*, New York, Journal of Organizational Behavior, 1986 7(4), p. 345–347. doi: 10.1002/job.4030070408.
5. SIEVERTSEN, H.H., Gino, F. *Cognitive fatigue influences students performance on standardized tests*, Proceedings of the National Academy of Sciences, 2016, 113(10), p. 2621–2624. doi: 10.1073/pnas.1516947113.
6. WILLIAMSON, A.M. ir Feyer, A.-M. *Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance*, Accident Analysis & Prevention, 2001, 33(3), p. 313–326. doi: 10.1016/S0001-4575(00)00045-2.
7. DAWSON, D. ir Reid, K. *Fatigue, alcohol and performance impairment: Abstract: Nature*, Nature, 1997, 388(6639), p. 235. doi: 10.1038/40775.
8. BRANDT, T. and Stemmer, R. *Affordable visual driver monitorina svstem for fatigue and monotony*, 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583), 2004, 7, p. 6451–6456. doi: 10.1109/icsmc.2004.1401415.
9. THIFFAULT, P. and Bergeron, J. *Monotony of road environment and driver fatigue: A simulator study*, Accident; analysis and prevention., 2003, 35(3), p. 381–91.
10. Lietuvos Respublikos administracinių teisės pažeidimų kodeksas, 126 straipsnis. *Transporto priemonės vairavimas, esant vairuotojams neblaiviems arba apsvaigusiems nuo narkotinių, psichotropinių ar kitokių psichiką veikiančių medžiagų* [interaktyvus]. 2015 [žiūrėta 2016-12-29] Prieiga per: <https://www.infolex.lt/ta/103787:str126>
11. Lietuvos Respublikos darbo kodeksas, 123 straipsnis. *Nušalinimas nuo darbo* [interaktyvus]. 2016 [žiūrėta 2016-12-29]. Prieiga per: [http://www.vdi.lt/Tekstai/Tekstai1.aspx?Tekstai\\_ID=180](http://www.vdi.lt/Tekstai/Tekstai1.aspx?Tekstai_ID=180)
12. MASCORD, D.J. ir Heath, R.A. *Behavioral and physiological indices of fatigue in a visual tracking task*, Journal of Safety Research, 1992, 23(1), pp. 19–25. doi: 10.1016/0022-4375(92)90036-9.
13. MASCORD, D.J., Walls, J., ir kt.. *Fatigue and Driving: Driver Impairment, Driver Fatigue and Driving Simulation*. Fatigue and alcohol: interactive effects on human performance in driving-related tasks. Londonas, 1995, p. 189–205, ISBN: 0748402624
14. WYLIE, C.D., Shultz, T., ir kt., *Driver Fatigue and Alertness Study: Technical Summary*, Commercial Motor Vehicle. Federal Highway Administration, 1996, p. 10 - 31
15. BADDELEY, A.D. *A 3 min reasoning test based on grammatical transformation* Psychonomic Science, 10(10), 1968, p. 341–342. doi: 10.3758/bf03331551.
16. ANUSAS-AMORNKUL, T. ir Wangsuk, K. *A comparison of keystroke dynamics techniques for user authentication*, 2015 International Computer Science and Engineering Conference (ICSEC), 2015, ISBN: 978-1-4673-7825-3, doi: 10.1109/icsec.2015.7401401.
17. VIZER, L.M., Zhou, L. ir Sears, A. *Automated stress detection using keystroke and linguistic features: An exploratory study*, International Journal of Human-Computer Studies, 2016, 67(10), p. 870–886. doi: 10.1016/j.ijhcs.2009.07.005.
18. MONROSE, F. ir Rubin, A.D. *Keystroke dynamics as a biometric for authentication*, Future Generation Computer Systems, 2000, 16(4), p. 351–359. doi: 10.1016/s0167-739x(99)00059-x.
19. BARGOUTH, H. *Keystroke dynamics how typing characteristics differ from one application to another*, Department of Computer Science and Media Technology Gjøvik University College, 2009
20. MORALES, A., Falanga, M., ir kt.. *Keystroke dynamics recognition based on personal data: A comparative experimental evaluation implementing reproducible research*, 2015 IEEE 7th International Conference on

- Biometrics Theory, Applications and Systems (BTAS), 2015, ISBN: 978-1-4799-8776-4, doi: 10.1109/btas.2015.7358772.
21. CHANG, T.-Y. *Dynamically generate a long-lived private key based on password keystroke features and neural network*, Information Sciences, 2012, 211, p. 36–47. doi: 10.1016/j.ins.2012.04.009.
22. DARABSEH, A. and Namin, A.S. *Effective user authentications using keystroke dynamics based on feature selections*, 2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), 2015, ISBN: 978-1-5090-0287-0. doi: 10.1109/icmla.2015.90.
23. GHODSI, A. *Dimensionality Reduction A Short Tutorial*. [interaktyvus] 2006 [žiūrėta 2017-01-22]. Prieiga per: [http://www.stat.washington.edu/courses/stat539/spring14/Resources/tutorial\\_nonlin-dim-red.pdf](http://www.stat.washington.edu/courses/stat539/spring14/Resources/tutorial_nonlin-dim-red.pdf)
24. Hotelling, H. *Analysis of a complex of statistical variables into principal components*, Journal of Educational Psychology, 1933, 24(6), p. 417. doi: 10.1037/h0071325.
25. ULINSKAS M., *Išmaniojo termostato sukūrimas*. Bakalauro baigiamasis darbas, Kaunas 2015
26. DZEMYDA G., *Daugiamatčių duomenų vizualizavimo metodai*, p. 108 – 118, Vilnius 2008
27. CORTES, C. ir Vapnik, V. *Support-vector networks*, Machine Learning, 1995, 20(3), pp. 273–297. doi: 10.1007/bf00994018.
28. FLETCHER T. *Support Vector Machines Explained*. [interaktyvus] 2009 [žiūrėta 2017.01.08] Prieiga per: <http://www.tristanfletcher.co.uk/SVM%20Explained.pdf>.
29. CHIH-CHUNG Chang, Chih-Jen Lin. LIBSVM – A Library for Support Vector Machines. [interaktyvus] 2016 [žiūrėta 2017.03.21] Prieiga per: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>.
30. JOUTSIJOKI, H., Haponen, M. ir kt.. *Error-correcting output codes in classification of human induced Pluripotent stem cell colony images*, BioMed Research International, 2016 ID: 3025057. doi: 10.1155/2016/3025057.
31. Weka Programinė įranga [interaktyvus]. 2016 [žiūrėta 2017-01-21] Prieiga per: <http://www.cs.waikato.ac.nz/ml/weka/>
32. LANDSBERGIS V. V. „Partizanas“. [interaktyvus]. 2009 [žiūrėta 2017-02-15] Prieiga per: <http://www.vlkk.lt/naujienos/kitos-naujienos/vytautas-v-landsbergis-partizanas-diktanto-tekstas-su-skyrybos-variantais-ir-paaiskinimais>
33. MURAKAMI H. *Dramblys pradingsta*. Baltos lankos, 2015, pp. 359. ISBN 9789955237969
34. UTKA A. *Dažninis rašytinės lietuvių kalbos žodynas 1 milijono žodžių morfologiškai anotuoto teksto pagrindu*. Kaunas: Vytauto Didžiojo universitetas, 2009, ISBN 978-9955-12-546-4

## 7. PRIEDAI

Prie projekto pridedamas kompaktinis diskas, kuriame įrašyta:

- šiame tyrime naudotos PĮ *Microsoft Visual Studio 2015* projektas;
- *Microsoft SQL Server 2016* duomenų bazė .bak formatu;
- tyrimo duomenų failai.
- šio tyrimo tema parengtas *Analysis of Keystroke Dynamics for Fatigue Recognition* straipsnis, kuris yra priimtas į 2017m. liepos 3 – 6 dienomis, Trieste, Italijoje vyksiančią **17th International Conference on Computational Science and Applications (ICCSA 2017)** konferenciją. Straipsnį parengusių autorių grupė: Mindaugas Ulinskas, Marcin Woźniak, Robertas Damaševičius.