

KAUNO TECHNOLOGIJOS UNIVERSITETAS
VYTAUTO DIDŽIOJO UNIVERSITETAS
VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

ANDRIUS LAURAITIS

**NEUROLOGINIŲ SUTRIKIMŲ APTIKIMO HIBRIDINIS
KLASIFIKAVIMO MODELIS**

Daktaro disertacijos santrauka
Gamtos mokslai, informatika (N 009)

2020, Kaunas

Disertacija rengta 2015–2019 m. Kauno technologijos universiteto Informatikos fakultete, Multimedijos inžinerijos katedroje.

Mokslinis vadovas:

Prof. dr. Rytis MASKELIŪNAS (Kauno technologijos universitetas, technologijos mokslai, informatikos inžinerija, T 007 ir gamtos mokslai, informatika, N 009).

Redagavo: Aurelija Gražina Rukšaitė (Leidykla „Technologija“) ir Brigita Brasienė (Leidykla „Technologija“)

Informatikos mokslo krypties disertacijos gynimo taryba:

Prof. habil. dr. Rimantas BARAUSKAS (Kauno technologijos universitetas, gamtos mokslai, informatika, N 009) – pirmininkas;

Prof. habil. dr. Gintautas DZEMYDA (Vilniaus universitetas, gamtos mokslai, informatika, N 009);

Prof. dr. Vacius JUSAS (Kauno technologijos universitetas, gamtos mokslai, informatika, N 009);

Dr. Zbigniew MARSZAŁEK (Lenkijos Silezijos regiono technikos universitetas, gamtos mokslai, informatika, N 009);

Prof. dr. Alfonsas MISEVIČIUS (Kauno technologijos universitetas, gamtos mokslai, informatika, N 009).

Disertacija bus ginama viešame Informatikos mokslo krypties disertacijos gynimo tarybos posėdyje 2020 m. balandžio 7 d. 13 val. Kauno technologijos universiteto Disertacijų gynimo salėje.

Adresas: K. Donelaičio g. 73-403, 44249 Kaunas, Lietuva.

Tel. (370) 37 300 042; faks. (370) 37 32 41 44; el. paštas doktorantura@ktu.lt.

Disertacijos santrauka išsiųsta 2020 m. kovo 6 d.

Su disertacija galima susipažinti internetinėje svetainėje <http://ktu.edu>, Kauno technologijos universiteto (K. Donelaičio g. 20, 44239 Kaunas), Vytauto Didžiojo universiteto (K. Donelaičio g. 52, Kaunas) ir Vilniaus Gedimino technikos universiteto (Saulėtekio al. 14, 10223 Vilnius) bibliotekose.

Turinys

1.	Įvadas	6
1.1.	Darbo aktualumas	6
1.2.	Darbo tikslas	7
1.3.	Darbo uždaviniai	7
1.4.	Mokslinis naujumas	8
1.5.	Darbo objektas	8
1.6.	Praktinė darbo reikšmė	8
1.7.	Darbo ginamieji teiginiai	9
1.8.	Darbo rezultatų aprobavimas	9
1.9.	Darbo struktūra	11
2.	Literatūros analizė	12
2.1.	CNSS pacientų simptomų apžvalga: Hantingtono ligos atvejis.....	13
2.2.	Susiję darbai	14
2.2.1.	CNSS pacientų sprendimų priėmimo klasifikavimo algoritmų apžvalga	14
2.2.2.	Naujausi informatikos mokslo srities darbai tremorų (motorikos), kognityviniams, kalbos ir energijos sąnaudų sutrikimams aptikti	15
2.3.	Analizės apibendrinimas.....	22
3.	Priemonės ir metodai.....	24
3.1.	Siūloma sveikatos būklės įvertinimo sistema	24
3.2.	Požymių atrinkimo algoritmų realizacija.....	25
3.2.1.	T1, T2, T3: Lietimai pirštai	26
3.2.2.	T4: Archimedo spiralė.....	27
3.2.3.	T5: Bendros įžvalgos.....	28
3.2.4.	T6: Orientacija (einamoji data)	28
3.2.5.	T7: Paveikslų įvardijimas	28
3.2.6.	T8: Panašumai, skaičiavimai	29

3.2.7. T9: 3-D figūros konstravimas	29
3.2.8. T10: Laikrodžio konstravimas	31
3.2.9. T11: Žodinis sklandumas.....	31
3.2.10. T12: Vizualinis loginis	32
3.2.11. T13: Problemų sprendimas	32
3.2.12. T0: Atmintis.....	33
3.2.13. T14: Balso įrašymas	33
3.2.14. T15: Dienos energijos sąnaudos	34
3.3. Priemonių ir metodų skyriaus apibendrinimas.....	35
4. Eksperimentiniai tyrimai	37
4.1. Duomenų rinkimo eksperimentas: dalyviai, procedūra ir duomenų rinkinys	37
4.2. Klasifikavimo metodai sprendimų priėmimo sistemose	38
4.2.1. Mokymo su mokytoju klasifikatoriai.....	38
4.2.2. Klasifikatorių „ansamblis“ (hibridinis modelis)	39
4.3. Atlikti eksperimentai, skirti sergančių ir sveikų asmenų klasifikavimo problemai išspręsti	41
4.3.1. E1: sergančių ir sveikų asmenų klasifikavimas paskirstant požymius atskiroms užduotims.....	41
4.3.2. E2: sergančių ir sveikų asmenų klasifikavimas naudojant visą požymių rinkinį	41
4.3.3. E3: balso sutrikimų klasifikavimas naudojant dvikryptį rekurentinį neuroninį tinklą su ilga trumpalaikė atmintimi	45
4.3.4. E4: balso sutrikimų klasifikavimas naudojant išsibarsčiusių (Gabor) bangelių transformaciją	45
4.4. Eksperimentų apibendrinimas	46
5. Išvados ir ateities darbai	47
5.1. Išvados	47
5.2. Ateities darbai	49

6.	Literatūros sąrašas	51
7.	Darbo sklaida	63
8.	Informacija apie autorių	65
9.	Summary	67
9.1.	Introduction	67
9.1.1.	Relevance of the research.....	67
9.1.2.	Aim of the research	68
9.1.3.	Tasks of the research	68
9.1.4.	Scientific novelty.....	69
9.1.5.	Research object	69
9.1.6.	Practical significance.....	69
9.1.7.	Dissertation statements.....	70
9.1.8.	Scientific approval.....	70
9.1.9.	Dissertation structure.....	72
9.2.	Literature Review	73
9.3.	Materials and methods	74
9.4.	Experimental Research	75
9.5.	Conclusions and future work	76
9.5.1.	Conclusions	76
9.5.2.	Future Work	79

1. Įvadas

1.1. Darbo aktualumas

Egzistuoja daug įvairių kompiuterinio intelekto modelių ir programų, skirtų asmenų sveikatos sutrikimams klasifikuoti. Kai kurie sprendimai yra sukurti medicininiam vaizdams apdoroti, o kiti naudoja klinikinių tyrimų duomenis sveikatos būklei numatyti. Vystantis technologijoms ir plečiantis medicininiam tyrimams, yra diegiamos išmaniosios diagnostikos priemonės, pvz., tyrimai mėgintuvėliuose ne žmogaus organizme [1]. Šiais laikais mobilieji išmanieji įrenginiai yra plačiai naudojami visame pasaulyje. Mobiliosios išmaniosios programėlės naudoja įrenginio palaikomas daugialypės terpės (angl. *multimedia*) galimybes, taip suteikdamos informacijos apie charakteringus žmogaus kūno bruožus ir požymius. Norint sukurti išmaniają sistemą, kuri galėtų būti naudojama medicininiam patikrinimams atlikti ar padėti paprastiems vartotojams kasdieniniame gyvenime, reikia klasifikavimo modelių, gebančių priimti sprendimus automatizuotuose medicininio tyrimų procesuose.

Mobiliosios sveikatos programos populiarėja, nes daugelis žmonių sėkmingai jas naudoja įvairiose medicinos praktikose. Programėlių taikymo sritys labai įvairios: nuo medicinos informacinių sistemų [2], fiziologinių sveikatos priežiūros sistemų [3], išpėjamųjų pranešimų (aliarmų) siuntimo sveikatos priežiūros paslaugų tiekėjams [4], reikalingų insulino dozių apskaičiavimo sistemų [5], kartu su kai kuriais e. sveikatos nešiojamaisiais prietaisais [6], iki labai specifinių programų priešlaikiniams skilvelių susitraukimams aptikti [7]. Diagnostikos programos turi didžiulį potencialą suteikiant prieigą diagnostikos terminui apibrėžti ir gali būti patrauklios tiek sveikatos priežiūros profesionalams, nustatant ligos ankstyvą stadiją, tiek patiems pacientams, pasiūlant papildomą diagnostikos priemonę [8]. Tokios sistemos gali būti adaptuotos gerinant Hantingtono ligos (HL) gydymo procesus, t. y. padedant daktarams stebėti pacientų sveikatos būklę nuotoliniu būdu [9].

Centrinės nervų sistemos sutrikimų (CNSS) turintiems asmenims dažniausiai pasireiškia kūno valdymo ar atminties sutrikimai. Senstant gyventojams, žmonių, kenčiančių nuo patologinio drebulio, skaičius nuolat didėja. 4 proc. pagyvenusių žmonių (> 65 metų) kenčia nuo drebjimų, o nuo Parkinsono ligos (PL) – 1 proc. vyresnių nei 50 metų

žmonių. Alzheimerio liga (AL) yra neurodegeneracinė liga, kurios simptomai yra atminties praradimas ir elgesio problemos, palaipsniui blogėjančios dėl kalbos sunkumų, dezorientacija ir elgesio problemos tvarkant savęs priežiūrą, kurios galiausiai lemia demenciją. Hantingtono liga (HL) daugiausia paveldima ir atsiskleidžia mirus smegenų ląstelėms. Ankstyvoje stadijoje pacientams kyla psichinių, suvokimo, koordinacijos stokos ar netolygaus eisenos problemų.

Vėlesnėse ligos stadijose pasireiškia nekoordinuoti judesiai ir blogėja sveikatos būklė, taip pat sutrinka kalba. Lengvi pažinimo sutrikimai (angl. *Mild Cognitive Impairment*) yra asmens mąstymo sugebėjimų pokyčiai (pablogėjimas), kuriuos pastebi aplinkiniai. Kai kurie lengvi pažinimo sutrikimų atvejai yra laikomi ankstyva AL stadija. Taigi atsiranda motyvacija aptikti išvardintas ligas dar ankstyvų stadijų, dar iki tol, kol simptomai pasimato vizualiai, o medicininis gydymas vis dar yra veiksmingas.

1.2. Darbo tikslas

Sukurti centrinės nervų sistemos sutrikimų (CNSS) turintiems pacientams adaptuotą kompiuterizuotą modelį, pagal kurį sudaromos prielaidos prognozuoti neurologinius sutrikimus ankstyvoje stadijoje.

1.3. Darbo uždaviniai

Suformuluoti penki uždaviniai:

1. Išanalizuoti ir palyginti naujausius duomenų gavybos ir mašininio mokymosi tyrimus, susijusius su klasifikavimo metodu taikymu pacientams, kuriems yra centrinės nervų sistemos sutrikimų (CNSS);
2. Sukurti išplėstinę kompiuterizuotą duomenų rinkimo metodologiją, naudojant CNSS pacientams adaptuotą išmaniąją sąsają;
3. Realizuoti požymių išgavimo sistemą, skirtą pacientams pasireiškiantiems motorikos (tremorų), kognityviniams, kalbos, energijos sąnaudų simptominiams sutrikimams kompiuterizuoti. Sukaupiti požymius duomenų rinkinyje;

4. Sukurti pasiūlytiems kompiuterizuotiems metodologiniams uždaviniams pritaikytus dvejetainius sveiko ir sergančio asmens klasifikavimo modelius;
5. Sujungti skirtingų tipų (funkcijos, Bajeso, sprendimų medžiai, taisyklės) mokymo su mokytoju klasifikatorius į hibridinį modelį, pagerinantį sveiko ir sergančio asmens klasifikatoriaus tikslumą ir sumažinantį klaidų skaičių.

1.4. Mokslinis naujumas

Išvardinami darbo naujumo aspektai:

1. Automatizuotas ankstyvų kognityvinių sutrikimų įvertinimas, naudojant centrinės nervų sistemos sutrikimų (CNSS) pacientams adaptuotą išmaniąją sąsają (mobilioji programėlė) bei kompiuterizuojant savarankiškai atliekamo pažintinio testavimo metodologiją (angl. *Self-administered cognitive testing, SAGE*).
2. *SAGE* kompiuterizuotos metodologijos praplėtimas papildomomis užduotimis, siekiant įvertinti CNSS pacientų tremorų (motorikos), kalbos ir energijos sąnaudų sutrikimus.
3. Inovatyvus CNSS pacientų sveikatos būklės stebėjimo metodas, skirtas įsivertinti ir palyginti su sveikų asmenų rezultatais.
4. Sukūrus hibridinį klasifikavimo modelį (derinami visi požymiai pagal siūlomą išplėstinę metodologiją) CNSS pacientų pablogėjusiai sveikatos būklei prognozuoti, gaunamas didesnis tikslumas, palyginti su atskirais modeliais.

1.5. Darbo objektas

Duomenų gavybos (angl. *Data Mining*) metodai, skirti sveikiems ir sergantiems asmenims klasifikuoti į dvi grupes (klases), ir mašininio mokymosi (angl. *Machine Learning*) modeliai būsimiems rezultatams prognozuoti.

1.6. Praktinė darbo reikšmė

Sukurta neurodegeneracinių procesų įvertinimo metodologija realizuota programinėje įrangoje NITS (angl. *Neural Impairment Test Suite*), kuri yra viešai prieinama vartotojams (*Google Play* portalas) [11]. Ši įranga gali būti efektyviai panaudota pirminiems

neurodegeneraciniams procesams žmogaus elgsenoje analizuoti ir vertinti. NITS programa išplatinta įvairiais žiniasklaidos kanalais tiek Lietuvoje, tiek ir visame pasaulyje bei pristatyta neurologų medicinos bendruomenei [12], [13], [14]. Disertacijos rašymo metu NITS programėle naudojasi daugiau nei 500 vartotojų.

1.7. Darbo ginamieji teiginiai

1. Pristatomas inovatyvus pacientų, turinčių centrinės nervų sistemos sutrikimų (pvz.: Hantingtono, Parkinsono, ankstyvos demencijos, cerebrinio paralyžiaus), sveikatos būklės stebėjimo metodas, pagrįstas išplėstine savarankiškai atliekamo pažintinio testavimo metodologija tremorams, kognityviniams, balso ir energijos sąnaudų sutrikimams įvertinti.
2. Sukurta sistema (NITS programėlė) gali būti adaptuota centrinės nervų sistemos sutrikimų (CNSS) turintiems pacientams namuose ar medicinos įstaigose, kaip sveikatos būklės sekimo ir grįžtamojo ryšio apie ankstyvą ligos būklės diagnozę suteikimo pagalbinė priemonė.
3. Sudarius atskirų klasifikatorių sujungimo mechanizmą visiems požymiams vertinti, gaunamas didesnę tikslumą ir mažesnes paklaidas turintis hibridinis modelis, sprendžiant dvejetainę (sveikas ir sergantis) klasifikavimo problemą.

1.8. Darbo rezultatų apibūdinimas

Web of Science ir Scopus duomenų bazių leidiniuose su citavimo rodikliu:

1. Lauraitis, A., Maskeliūnas, R., Damaševičius, R., Polap, D., Wozniak, M. (2019). A smartphone application for automated decision support in cognitive task based evaluation of central nervous system motor disorders. *IEEE Journal of Biomedical and Health Informatics*. Electronic ISSN: 2168-2208. DOI 10.1109/JBHI.2019.2891729. Autoriaus indėlis: 0.400.
2. Lauraitis, Andrius; Maskeliūnas, Rytis; Damaševičius, Robertas. ANN and fuzzy logic based model to evaluate Huntington disease symptoms // *Journal of healthcare engineering*. New York

: Hindawi. ISSN 2040-2295. eISSN 2040-2309. 2018, vol. 2018, art. no. 4581272, p. 1-10. DOI: 10.1155/2018/4581272. [DOAJ; Scopus; MEDLINE; Science Citation Index Expanded (Web of Science)]. Autoriaus indėlis: 0.334.

3. Ivanavičius, Arnas; Simonavičius, Henrikas; Gelšvartas, Julius; Lauraitis, Andrius; Maskeliūnas, Rytis; Cimpmperman, Piotras; Serafinavičius, Paulius. Real-time CUDA-based stereo matching using Cyclops2 algorithm : research // EURASIP journal on image and video processing. Cham : Springer Open. ISSN 1687-5176. eISSN 1687-5281. 2018, vol. 2018, art. no. 12, p. 1-15. DOI: 10.1186/s13640-018-0253-2. [DOAJ; Scopus; Science Citation Index Expanded (Web of Science)]. Autoriaus indėlis: 0.142.

Tarptautinėse ir nacionalinėse konferencijose:

1. Lauraitis, Andrius; Maskeliūnas, Rytis. Investigation of predicting functional capacity level for Huntington disease patients // Information and software technologies: 23rd international conference, ICIST 2017, Druskininkai, Lithuania, October 12–14, 2017: proceedings / Robertas Damaševičius, Vilma Mikašytė (Eds.). Cham : Springer, 2017. ISBN 9783319676418. eISBN 9783319676425. p. 142-149. (Communications in computer and information science, ISSN 1865-0929, eISSN 1865-0937 ; vol. 756). DOI: 10.1007/978-3-319-67642-5_12. Autoriaus indėlis: 0.500.
2. Gelšvartas, Julius; Simonavičius, Henrikas; Lauraitis, Andrius; Maskeliūnas, Rytis; Cimpmperman, Piotras; Serafinavičius, Paulius. Realtime stereo matching using maximally stable extremal regions // Transylvanian review. 2016, vol. 24, iss. 11, p. 1-11. Autoriaus indėlis: 0.166.
3. Gelšvartas, Julius; Lauraitis, Andrius; Simutis, Rimvydas; Maskeliūnas, Rytis. Review of assistive technologies for disabled people // Biomedical engineering 2016 : proceedings of international conference / Kaunas University of Technology, Lithuanian Society for Biomedical Engineering. Kaunas : Kauno

technologijos universitetas. ISSN 2029-3380. 2016, p. 142-146. Autoriaus indėlis: 0.250.

4. Maskeliūnas, Rytis; Lauraitis, Andrius. Neural impairment test suite mobile app // Technorama 2019: from vision to innovation! : innovation catalogue. [Kaunas]: KTU. 2019, project no. 49, p. 59. Autoriaus indėlis: 0.500.

2018 m. liepos mėnesį dalyvauta vasaros mokykloje Talino universitete, išklaudytas kursas „Žmogaus ir kompiuterio sąveikos eksperimentiniai tyrimai ir projektavimas“, vertinamas 4 ECTS kreditais. Technorama 2019 konferencijoje autorius laimėjo R1 įmonių prizą.

1.9. Darbo struktūra

Disertaciją sudaro 6 skyriai. Antrame skyriuje apžvelgiami centrinės nervų sistemos sutrikimų simptomai, įskaitant tuos, kurie pasireiškia Hantingtono liga sergantiems pacientams, taip pat naujausi mokslo darbai, susiję su duomenų gavybos ir mašininio mokymo sritimis. Trečiame skyriuje aprašytos disertacijoje naudojamos priemonės ir metodai, t. y. kompiuterizuotos išplėstinės savarankiškai atliekamo pažintinio testavimo (angl. *SAGE*) metodologijos projektavimas ir kūrimas, požymių išgavimo metodai. Ketvirtame skyriuje aprašomi keturi atlikti eksperimentai, skirti sergančių ir sveikų asmenų klasifikavimo problemai išspręsti. Pateikiamas tirtų klasifikatorių sąrašas ir pasirinkimo motyvai. Pirmasis eksperimentas (E1) atskiroms užduotims taiko požymių paskirstymo metodą. Antrame eksperimente (E2) naudojamas visas požymių rinkinys ir, sujungiant atskirus klasifikatorius, sudaromas hibridinis modelis. Trečias (E3) ir ketvirtas eksperimentai (E4) apdoroja balso signalus galimiems kalbos sutrikimams įžvelgti. Penktame skyriuje pateikiamos darbo išvados ir diskusija ateičiai, šeštame – literatūros sąrašas. Septintame skyriuje nurodomi darbo sklaidos kanalai.

2. Literatūros analizė

Atliekama daug informatikos mokslo srities tyrimų, nagrinėjančių centrinės nervų sistemos sutrikimų (CNSS) turinčių pacientų sveikatos būklę. Daugiausia tokių tyrimų apima Hantingtono ligą (HL), Parkinsono ligą (PL), Alzheimerio ligą (AL), lengvus kognityvinius sutrikimus (LKS) (angl. *Mild Cognitive Impairment, MCI*) ar demenciją. Išvardintų ligų ir sutrikimų turintiems pacientams pasireiškia labai įvairūs simptomai (išskiriamos keturios grupės): tremorai (nevalingi kūno judesiai, pusiausvyros praradimas, raumenų sąstingiai), kognityviniai (sprendimų priėmimo sunkumai, dėmesio sutelkimo problemos, atminties stoka, nepakankamas apsirūpinimas savimi), balso (mažesnis ištartų žodžių skaičius, trumpesnių frazių naudojimas, pauzės) ir energijos sąnaudų disbalansai (pakitęs svoris).

Kompiuterių mokslininkų tikslai orientuojasi į neurologinių sutrikimų turinčių pacientų sveikatos būklės įvertinimo diagnostines priemones ar sprendimo priėmimo modelius [11] [69]. Didžioji dalis tokio pobūdžio tyrimų grindžiami medikų žiniomis patvirtintomis metodologijomis. Populiariausios yra trys metodologijos, skirtos kognityviniams sutrikimams išvelgti: 1) Savarankiškai atliekamas pažintinis testavimas (angl. *Self-Administered Gerocognitive Examination, SAGE*) [70]; 2) Monrealio pažintinio vertinimo metodika (angl. *Montreal Cognitive Assessment, MoCA*) [72]; 3) Mini-protinės veiklos sutrikimų tyrimas (angl. *Mini-Mental State Examination, MME*) [73]. Visos šios trys metodologijos turi užduočių vertinimo sistemą taškais. Skirtingose metodikose randama persidengiančių (pasikartojančių) užduočių. Kognityviniai testuojamų asmenų gebėjimai vertinami pagal šias užduotis: atminties patikra (frazės atkartojimas), orientacija (einamosios datos, laiko ir vietos nurodymas), duotos figūros nubraižymas, vizualiniai-erdviniai sugebėjimai, objektų identifikavimas, kalbos patikra ir pan. Sutrikusi sveikatos būklė diagnozuojama tada, kai nepasiekama nustatyta taškų riba. Literatūroje rasta kompiuterizuota elektroninė *SAGE* metodologijos versija (e*SAGE*) [71], tačiau *MoCA* ir *MME* testai vis dar atliekami naudojant popierinę formą su tiesiogine gydytojo priežiūra.

Skyriuje pirmiausia pateikiama Hantingtono ligos (HL) simptomų apžvalga. HL pasirinkta, nes apima visas keturias disertacijoje nagrinėjamų sutrikimų grupes. Kita priežastis – HL simptomai labiausiai

nenuspėjami ir mažiausiai prognozuojami, palyginti su kitomis neurologinėmis ligomis. Papildomas motyvas – didžioji dalis disertacijoje tirtų asmenų, kuriems pasireiškė neurologiniai sutrikimai, yra HL pacientai. Literatūros analizės skyriuje pateikiami panašaus pobūdžio susiję darbai, apimantys sprendimus kompiuterizuoti pacientų tremorų (motorikos), kognityvinius, balso ir energijos sąnaudų sutrikimus. Pabaigoje apibendrinama panašių darbų lyginamoji analizė, pateikiant egzistuojančių sprendimų trūkumus ir motyvus kurti disertacijoje aprašomus metodus.

2.1. CNSS pacientų simptomų apžvalga: Hantingtono ligos atvejis

Hantingtono liga (HL) – tai judėjimo, mąstymo ir psichikos sutrikimais apibūdinama neurodegeneracinė būseną. HL smarkiai pakeičia sergančių asmenų kasdieninį gyvenimą. HL yra nepagydoma, todėl didžioji dalis kompiuterizuotų tyrimų skirti ligos progresavimą mažinančioms prevencinėms priemonėms ieškoti ir taip laimėti daugiau sveiko gyvenimo metų [74] [75]. HL yra labai reta liga, t. y. statistiškai sergamumas Europoje yra 1 atvejis iš 20 000–30 000 asmenų. Lietuvoje, 2015 m. Santariškių klinikų HL koordinacinio centro duomenimis, yra oficialiai užregistruoti 177 ligos atvejai [76]. 2.1 lentelėje apibendrinami HL būdingi simptomai.

2.1. lentelė Hantingtono ligos simptomų apžvalga

Stadija	Sutrikimas	Simptomai
Ankstyva		Nerimas, nevalingi veido ir galūnių judesiai
Vidurinė		Chorėja, nenormalus raumenų tonusas
Vėlyva	Tremorai (motorika)	Nelankstumas, akinezija (nejudrumas dėl paralyžiaus), hipokinezija (raumenų sustingimas), dizartrija (kalbos artikuliacijos sutrikimas), disfagija (rijimo sutrikimai)
Ankstyva, vidurinė, vėlyva	Kognityviniai, elgesio (emociniai)	Sprendimo priėmimo sunkumai, problemos sutelkti dėmesį, atminties praradimas, nerimas, depresija, miego sutrikimai, dirglumas, agresija, dirglumas, įtarumas, pyktis, įmantrus ir ekscentriškas elgesys,

		netinkamas seksualinis elgesys, pavydas, apatija
-----	Kalbos, energijos sąnaudos	Tariamų žodžių trūkumas, trumpesnių frazių vartojimas, pauzės. Svorio metimas, neigiamas energijos balansas

2.2. Susiję darbai

2.2.1. CNSS pacientų sprendimų priėmimo klasifikavimo algoritmų apžvalga

Skyriuje pateikiamas egzistuojančių mašininio mokymo (MM) metodų palyginimas. Nagrinėti tiriamieji subjektai: Parkinsono liga (PL), Hantingtono liga (HL) ir Alzheimerio liga (AL), taip pat sveiki asmenys. 2.2 lentelėje literatūros šaltiniui (Lit.) nurodyti klasifikatoriai (jų sutrumpinimai) anglų kalba.

2.2. lentelė. Mašininio mokymosi (MM) metodų, pritaikytų sprendimų priėmimo sistemoms neurologiniams sutrikimams aptikti, palyginimas

Lit.	MM metodas	Algoritmo tipas	MM problema	Tiriamųjų Skaičius
[77]	ANN- MLP	Mokymas su mokytoju	Klasifikavimas	21
[78]	RBFNN		Regresija	-
[79]	DNN		Klasifikavimas	12
[80]	Sprendimų medis, ID3		Klasifikavimas	31
[81]	Neuro-neraiškios aibės	Hibridinis		-
[82]	Neuro-neraiškios aibės			-
[83]	Klasifikatorių sintezė (Bajeso, SVM, k-NN)		AL (13), PL (15), HL (16), sveiki (16)	
[84]	Markovo modeliai		72 (82 PL, 62 sveiki)	

[85]	Atsitiktiniai miškai (C-RT), ID3, Binarinė logistinė regresija, k-NN, (SVM)	Mokymas su mokytoju	Klasifikavimas	31 (23 PL, 8 sveiki)
[86]	FCM	Mokymas be mokytojo	Klasterizavimas	-

Dirbtinis neuroninis tinklas su daugiasluoksniu perceptronu (angl. *ANN-MLP*) pritaikytas PL ligai klasifikuoti naudojant akcelerometro signalus [77]. Radialinių bazių neuroninis tinklas (angl. *RBFFN*) [78] pritaikytas PL ligos progresavimui prognozuoti pagal elektromiografiją (angl. *EMG*). Dinaminiai neuroniniai tinklai (angl. *DNN*) naudoti PL nekontroliuojamiems judesiams identifikuoti naudojant EMG [79]. Sprendimų medis (angl. *Decision Tree ID3*) [80] naudotas HL ir PL progresavimui prognozuoti pagal surinktus tremorų sutrikimų duomenis (rankų ir kojų drebėjimai, sutrikusi kalbos artikuliacija). Neuro-neraiškių aibių sistemos (angl. *Neuro-fuzzy systems*) adaptuotos darbuose [81] [82], skirtuose PL ir AL pacientų sveikatos būklei įvertinti pagal spektrinius, su tremorais susijusius požymius. Bajeso (angl. *Bayes*), atraminių vektorių mašinų (angl. SVM) ir k-artimiausių kaimynų (angl. k-NN) klasifikatorių sintezė HL, PL ir AL ligų prevencijai naudota [83] darbe. Daugiapakopiai Markovo modeliai (angl. *Markov Model*) [84] pritaikyti adaptyviose sprendimo priėmimo sistemose, skirtose smegenų ir sutrikusios eisenos informacijai sujungti. Parkinsono ligai identifikuoti skirti klasterizavimo algoritmai (pvz.: angl. *Fuzzy Means Clustering, FCM*) aprašomi [85] ir [86] darbuose.

2.2.2. Naujaisi informatikos mokslo srities darbai tremorų (motorikos), kognityviniams, kalbos ir energijos sąnaudų sutrikimams aptikti

2.3 lentelėje pateikta piršto lietimų užduočių (PLU) kompiuterizuotų sprendimų lyginamoji analizė. Autorių C. Morimoto ir kt. darbe [87] PLU eksperimentams vykdyti reikalinga speciali aplinka ir aparatinė įranga (magnetinių jutiklių sistema, susieta su asmeniu kompiuteriu). Mokslininkų J. Payne ir kt. [88] sprendime PLU testas

pritaikomas Parkinsono (PL) pacientų sulėtintiems piršto judesiams aptikti, t. y. naudojamos specialios pirštinės liečiamam ekranui, įvertinant laiko intervalus ir pauzes tarp paspaudimų. Autoriaus J. Binder darbe [89] tiriamieji buvo Hantingtono pacientai (HL), o jų pirštų miklumas buvo matuojamas išmanojo įrenginio programėle dviem režimais: 1) pacientams liepta kuo greičiau spausti pirštu ekrane; 2) įrenginys laikomas testuojamoje rankoje ir 30 sekundžių pirštu spaudžiama ekrane. Autorių J. Milica ir kt. tyrime [90] eksperimentams vykdyti reikalinga speciali aplinka ir belaidis ryšis su asmeniniu kompiuteriu, įvertinant PLU užduočių atlikimo rezultatus PL pacientams. Mokslininko S. Suzumura [91] darbe pritaikomas specialus terminalas su integruotais magnetiniais jutikliais ritmiškai PLU atlikti. Į tyrimą įtraukti Alzheimerio (AL) ir lengvo kognityvinio sutrikimo (LKS) pacientai.

2.4 lentelėje pateikta spiralės kontūro sekimo (SKSU) užduočių kompiuterizuotų sprendimų lyginamoji analizė. Autoriaus H. N. Pham sprendime [92] PL identifikuoti naudojamas jau egzistuojantis duomenų rinkinys, kuriame išsaugoti pacientų ranka nupiešti Archimedo spiralės brėžiniai. C. Pereira ir kt. darbe [93] sveiko ir sergančio klasifikavimo uždaviniui pritaikomi ne tik spiralės paveikslai, bet ir išmaniaisiais rašikliais sveikų bei PL asmenų surinkti judesių motorikos požymiai. SKSU užduotis vykdyta įprastu rašikliu ant A4 formato popieriaus lapo, pritvirtinto prie grafinės planšetės [94]. PL pacientai instruktuoti braižyti spiralę trimis režimais: 1) spontaniškai ir kuo greičiau; 2) didelė ir maža spiralė; 3) kaire ir dešine rankomis. Tyrėjai L. Bernardo ir kt. [95] sprendime sukūrė specialią sąsają, adaptuotą asmeniniams kompiuteriams MS Windows aplinkoje. PL pacientų SKSU atlikta užduotis vertinama pagal atstumo metrikas (Euklido, Manhatan). Mokslininkų M. Moetesum ir kt. darbe [96] PL pacientai brėžė spiralę, naudoti įprasti rašikliai ir atsižvelgta į (x,y) kontūro koordinatas bei požymį, ar rašiklis liečia paviršių.

2.3. lentelė PLU panašių darbų palyginimas

Lit.	Tiriamųjų grupė	Specializuota įranga	Specializuota aplinka	Mobiliosios technologijos
[87]	Autizmas	Taip (magnetiniai jutikliai)	Taip	Ne
[88]	PL	Taip (specialios	Taip	Ne

		pirštines liečiamam ekranui)		
[89]	HL	Ne	Ne	Taip
[90]	PL	Taip (inerciniai, jėgos jutikliai)	Taip	Ne
[91]	LKS, AL	Taip (terminalas su magnetiniais jutikliais)	Ne	Taip

2.4. lentelė SKSU panašių darbų palyginimas

Lit.	Tiriamųjų grupė	Egzistuojantis duomenų rinkinys	Mobiliosios technologijos	Klasifikavimo tikslumas, % (metodas)
[92]	PL	Taip (UCI saugykla)	Taip (mobilioji aplikacija)	99.6 (k-NN)
[93]		Taip (HandPD) ir autoriaus indėlis	Dalinai (išmanieji parkeriai)	95 (CNN)
[94]		Ne	Dalinai (popieriaus lapai sujungti su planšete)	-
[95]		Ne (MS Windows sąsaja)	Ne (MS Windows sąsaja)	100 (SVM)
[96]		Taip (HandPD)	Ne	83 (SVM)

2.5 lentelėje pateikta kognityvinių sutrikimų kompiuterizuotų sprendimų lyginamoji analizė. Autorių M. Lunven ir kt. darbe [97] tirti HL pacientai, įvertinant jų sugebėjimų pokyčius atlikti matematinės užduotės. Ontologijomis grindžiama kompiuterizuota sistema pritaikyta PL pacientų kognityvinių sutrikimų reabilitacijai atliekant nurodytos kategorijos ir elementų paieškos, vizualinio stimulo skatinimo užduotės [98].

2.5. lentelė Kognityvinių sutrikimų panašių darbų palyginimas

Lit.	Tiriamųjų grupė	Metodologija	Duomenų rinkinys	Sistema
[97]	HL	-	34 HL ir 23 sveiki (autorius indėlis)	Nežinoma (psichometrijos prietaisas)
[98]		MME, MoCA	31 PL, dvi duomenų bazės (autorius indėlis)	Ontologija (PC)
[99]	PL	MoCA ir papildoma (SAGE)	37 PL, 47 sveiki (autorius indėlis)	Dviem mygtukais valdomas įrenginys, Matlab programinė įranga
[100]	Demencija	MME, MoCA ir papildoma (SAGE)	5 demencija ir 5 sveiki (autorius indėlis)	PC (Windows OS) ir Android programa žaidimizacijai
[101]	PD	-	118 PL pacientai (nėra autorius indėlio)	Prietaisas EEG signalams įrašyti, specializuota programinė įranga

Kognityvinių sutrikimų kompiuterizavimas PL pacientams tirtas [99] atliekant vizualių objektų (pvz.: rodomų priebalsių, raidžių ar paveikslų korteles) identifikavimo ekrane užduotis. Mobilioji programėlė

(žaidimas), skirta kognityviniams sutrikimams nustatyti demencija sergantiems pacientams, pristatoma [100] sprendime. Tiriamieji turėjo spręsti einamosios datos nurodymo, dėmesio sutelkimo, skirtingų elementų išvardinimo, kelių trajektorijų konstravimo, kompiuterizuotas užduotis. Autorių N. Betrouni ir kt. darbe [101] pagal surinktus elektroencefalogramos duomenis PL pacientai suskirstomi (klasifikuojami) į penkias atskiras grupes, atsižvelgiant į atliktų kognityvinių užduočių rezultatus.

2.6 lentelėje pateikta kalbos sutrikimų kompiuterizuotų sprendimų lyginamoji analizė. Tyrėjų A. Tsanas ir kt. darbe [102] tiriamieji klasifikuojami į sveikus ir sergančius (PL atvejais) pagal signalo-triukšmo ir harmoninį-triukšmo santykių požymius ir atraminių vektorių mašinų ir atsitiktinių miškų klasifikatorius. Mokslininkai Y. Hauptman ir kt. [103] savo darbe PL identifikuoti pritaiko šiuos požymius balso signalui: melo dažnio cepstraliniai koeficientai (angl. *MFCC*), tiesinės prognozės koeficientai (angl. *Linear Prediction Coefficients, LPC*), diskreti bangelių transformacija, Gauso maišos modeliai (angl. *Gaussian mixture model, GMM*), laiko srities entropija. 77,2 proc. tikslumas pasiektas naudojant atraminių vektorių mašinų algoritmą. Mokslininko K. Wu tyrime [104] balso įrašai surinkti specializuotame garsui nelaidžiame kambaryje, o PL nustatyti naudoti *MFCC* koeficientai ir atsitiktinių miškų klasifikatorius (96,37 proc. tikslumas). M. Perez ir kt. [105] tyrinėta Hantingtono ligos identifikacija pagal akustinius ir leksinius (*MFCC, GMM*, pauzes, kalbos dažnį, tarties tikslumo) balso požymius. Naudoti k-artimiausių kaimynų (*KNN*) ir rekurentinių neuroninių tinklų su ilga-trumpalaikė atmintimi (angl. *LSTM-RNN*) algoritmai (gauta 0,87 koreliacija). W. Q. Oung [106] darbe analizuotos Parkinsono ligos progresavimo stadijos pagal entropijos (bangelių, Shannon, Renyi, Tsallis, kombinacinė, neraiškių aibių logikos) požymius. Pasiiektas 91,11 proc. klasifikavimo tikslumas su ekstremalaus mokymosi mašinų metodu (angl. *Extreme learning machine, ELM*).

2.7 lentelėje pateikta energijos sąnaudų sutrikimų kompiuterizuotų sprendimų lyginamoji analizė. Autoriai M. Cersosimo ir kt. [107] savo darbe pasiūlo daugiamatės logistinės regresijos algoritmą, siekiant nustatyti ryšį tarp svorio netekimo (laikui bėgant) ir PL motorinių pasireiškimų. Tyrėjai B. Singh ir T. Hissam [108] pritaiko atraminių vektorių mašinų ir dirbtinio neuroninio tinklo mašininio mokymosi metodus, skirtus nutukimo (priaugusio svorio) rizikai paaugliams

prognozuoti pagal jų kūno masės indeksą (KMI). Mokslininkų V. Villasana ir kt. [109] tyrime energijos sąnaudų sutrikimams analizuoti sukurta Android mobilioji programėlė (žaidimas), skirta asmens sveikatos mitybos įpročiams ir fizinei veiklai kontroliuoti pagal antropometrinius (pakitusio gyvenimo būdo) parametrus. Autorių B. Tirasirichai ir kt. [110] darbe energijos sąnaudų sutrikimai nagrinėjami pasitelkiant Android mobiliąją aplikaciją, siūlančią asmeniui subalansuotą kalorijų planą pagal rekomenduojamas dienos normas (atsižvelgiama į bazinę medžiagų apykaitą). Tyrėjų K. Merchant ir kt. [111] darbe prognozuojamas asmens tinkamos mitybos planas pagal mobiliojoje programoje rodomus poveikius (konvoliucinis neuroninis tinklas).

2.6. lentelė Kalbos sutrikimų panašių darbų palyginimas

Lit.	Tiriamųjų grupė	Duomenų rinkinys	Aparatinė, programinė įranga	Statistika
[102]	PL	14 PL subjektų (autorius indėlis)	Ant galvos montuojamas mikrofonas, Audacity	90 proc. tikslumas (SVM)
[103]		12 PL ir 12 sveikų (autorius indėlis)	Mikrofonas, nepateikta	77,2 proc. tikslumas (SVM su tiesiniu branduoliu)
[104]		27 PL ir 446 sveiki (autorius indėlis)	Garsui nelaidus kambarys, Matlab (Voice Analysis paketas)	96,37 proc. (atsitiktinių miškų klasifikatorius)
[105]	HL	31 HL ir 31 sveikas (paimta iš kito tyrimo)	Nepateikta, kompiuterizuota kalbos analizė (angl. CLAN)	Koreliacijos analizė (0,87)

[106]	PL	65 PL (autorius indėlis)	Ausinės Sennheiser DW Pro2, Matlab ir Simulink	91,11 proc. (ekstremalus mokymasis)
-------	----	--------------------------------	---------------------------------------------------------	-------------------------------------------

2.7. lentelė Energijos sąnaudų sutrikimų panašių darbų palyginimas

Lit.	Tiriamųjų grupė	Duomenų rinkinys	Sistemos tipas	Statistika
[107]	PL	144 PL ir 120 sveikų (autorius indėlis)	SPSS programinė įranga statistinei analizei atlikti (PC)	Regresijos analizė ($p < 0,001$)
[108]	Jaunimas	46 jaunuoliai (autorius indėlis)	ANOVA programinė įranga statistinei analizei atlikti (PC)	Regresijos analizė ($p < 0,05$)
[109]	Studentai savanoriai	Tuščia (prototipas)		Netyrinėta
[110]	Asmenys, norintys pakeisti valgymo įpročius	Nepateikta	Mobilioji Android aplikacija	
[111]	Cukriniu diabetu sergantys asmenys	Paveikslai (ETHZ Food-101 duomenų rinkinys)		

2.3. Analizės apibendrinimas

Pateikta panašių darbų literatūros analizė įrodo disertacijoje nagrinėjamos temos populiarumą ir potencialą, siekiant aptikti neurologinius sutrikimus, pasireiškiančius asmenims, sergantiems Hantingtono liga (HL), Parkinsono liga (PL), Alzheimerio liga (AL) ar demencija. Tačiau egzistuojantys sprendimai turi trūkumų, kurie suteikia motyvacijos kurti siūlomą sistemą ir metodą, įvertinant šiuos aspektus:

1. Tremorų (motorikos) sutrikimų panašių darbų analizėje apžvelgti susiję pirštų lietimų (baksnojimų) užduočių (PLU) ir spiralės kontūro sekimo (braižymo) užduočių (SKSU) darbai. Nė viename PLU ir SKSU literatūros šaltinyje nėra tyrinėta Hantingtono liga. Didžiąjai daliai PLU darbų reikalinga papildoma specializuota programinė ir aparatinė įranga bei aplinka eksperto prižiūrėtiems eksperimentams vykdyti. Mažuma PLU užduočių pritaiko mobiliąsias technologijas. Dauguma SKSU panašių darbų tikrina Parkinsono liga sergančių asmenų rašysenos sutrikimus (pusė jų naudoja jau egzistuojančius duomenų rinkinius). Mažuma SKSU apžvelgtų darbų naudoja mobiliąsias Android programėles požymiams atrinkti, vietoj to yra pritaikomos skaitmeninės planšetės ar rašikliai duomenims rinkti. Siūlomi modeliai (PLU ir SKSU užduotims) yra statistiškai įvertinami klasifikavimo tikslumo metrika.
2. Kognityvinių sutrikimų panašių sprendimų analizėje apžvelgti darbai, susiję su atminties praradimo užduočių, grindžiamų *MoCA*, *MME* ar *SAGE* metodologijomis, kompiuterizavimu Parkinsono ir demencijos pacientams (tik į vieną tyrimą buvo įtraukti Hantingtono ligoniniai). Daugumoje darbų adaptuota specializuota aparatinė ir programinė įranga, skirta pacientų sveikatos būklei įvertinti asmeniniame kompiuteryje (eksperimento vykdymą prižiūri ekspertas). Siūlomi modeliai yra įvertinami taikant regresijos analizę ar klasifikavimo tikslumo statistines metrikas.
3. Kalbos sutrikimus labai intensyviai analizuoja kiti kompiuterių srities mokslininkai. Didžioji dalis susijusių darbų tiria Parkinsono ligą. Siūloma labai įvairių metodų eksperimentams

atlikti iš neurologinių pacientų surinktų balso įrašų: daugumai sprendimų reikalinga speciali aparatinė įranga (mikrofonai, stiprintuvai, ausinės) ir signalų apdorojimo programinė įranga (*MATLAB*, *Praat*, *Audacity* ar *SPSS*). Be to, kalbos sutrikimams aptikti naudojami labai įvairūs požymiai (ir jų deriniai), t. y. akustiniai (virpėjimas, impulsų skaičius, balso pertraukos ir t. t.), melo dažnio cepstraliniai koeficientai (*MFCC*), Gauso mišinių modeliai (*GMM*), spektriniai (plitimas, centroidas, entropija ir pan.), bangelių transformacija.

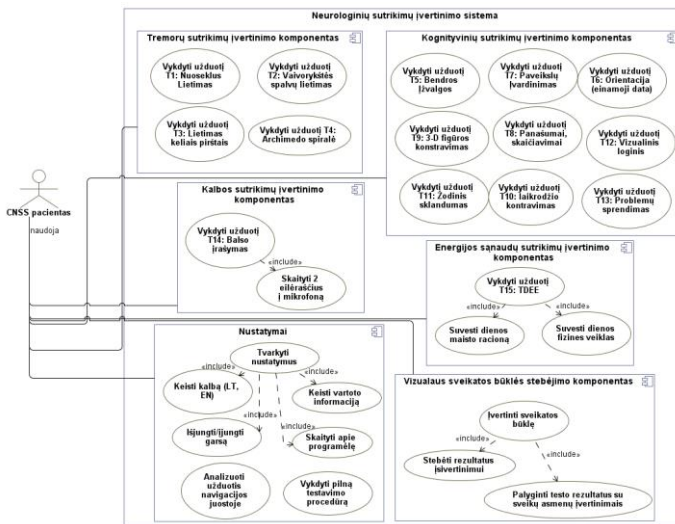
4. Panašių darbų analizė parodė, kad egzistuoja labai mažai neurologinių sutrikimų turintiems pacientams (rastas vienas Parkinsono ligonių svorį tiriantis darbas) adaptuotų kompiuterizuotų sprendimų, skirtų energijos sąnaudų sutrikimams aptikti. Nė viename iš nagrinėtų darbų nėra tirti Hantingtono pacientai. Daugumoje panašaus pobūdžio tyrimų bandoma aptikti pacientų svorio mažėjimą pagal kūno masę (KM) ir indeksą (KMI). Pusė aprašytų darbų naudoja pacientų mitybą ir dienos kalorijų balansus įvertinančias Android mobiliąsias aplikacijas, o kita pusė – statistinei analizei asmeniniuose kompiuteriuose pritaikytas *ANOVA*, *SPSS* programas. Kompiuterizuotos sprendimų priėmimo sistemos dienos energijos sąnaudų požymiams yra retai analizuojamos.

Taigi disertacijos autorius pristato naują CNSS pacientų automatizuoto sveikatos būklės stebėjimo metodą. Duomenims rinkti naudojama viena išmanioji sąsaja, t. y. mobilusis įrenginys (telefonas ar planšetė), nereikalaujantis papildomų jutiklių ir daviklių, pritvirtintų ant tiriamojo kūno. Pasiūlytas metodas yra inovatyvus dėl unikalios suformuotos požymių kombinacijos, naudojamos pacientų tremorų, kognityviniams, balso ir energijos sąnaudų sutrikimams aptikti. Panašių darbų analizė įrodo, kad kiekvienas iš šių sutrikimų (atskirai) yra reikšmingas pablogėjusios sveikatos būklės diagnozavimo rodiklis, tačiau tokių sutrikimų sintezė (vertinant pagal požymius) nebuvo nagrinėta. Darbe pasirinkta *SAGE* (lyginant su *MoCA*) metodologija dėl šių motyvų: 1) *SAGE* yra naujausia (vėliausiai sukurta) metodologija; 2) *SAGE* turi daugiausiai skirtingų variantų klausimynų; 3) egzistuoja elektroninė *SAGE* versija (*eSAGE*), taip tyrėjams suteikiama galimybė ne tik kompiuterizuoti, bet ir praplėsti mediciniškai patvirtintą priemonę.

3. Priemonės ir metodai

3.1. Siūloma sveikatos būklės įvertinimo sistema

Darbe pristatomas savarankiška pažintinio testavimo metodologija (angl. *Self-administered gerocognitive testing methodology, SAGE*) [15] [16] grindžiamas duomenų rinkimo įrankis, skirtas lengviems pažinimo sutrikimams (angl. Mild Cognitive Impairment, *MCI*) nustatyti. Metodika taikoma pacientams, kuriems pasireiškia ankstyvosios demencijos požymiai, taip pat sergantiems Hantingtono (HL), Parkinsono (PL) ar Alzheimerio liga (AL). Praktikoje gydytojai neurologai naudoja *SAGE*, pateikdami pacientams klausimynus ir patys įvertindami jų sveikatos būklę. Šiame skyriuje pateikiamas inovatyvus *SAGE* anketų kompiuterizavimo sprendimas, sujungiant visų 4 skirtingų *SAGE* variantų (grupių) 12 klausimų su papildomomis užduotimis HL, PL, AL pacientų tremorų (motorikos), kalbos ir energijos sąnaudų sutrikimams aptikti. Tokiu principu paciento sveikatos būklės stebėjimas tampa automatizuotas, t. y. be tiesioginės gydytojo priežiūros. 3.1 pav. iliustruotas siūlomos sistemos panaudos atvejų modelis. Išplėstinė metodologija realizuota mobiliojoje programėlėje NITS (angl. *Neural Impairment Test Suite*) [11].



3.1 pav. Siūloma sveikatos būklės įvertinimo sistemos schema (NITS)

Disertacijoje nagrinėjami požymių atrinkimo (angl. *Feature Extraction*) metodai: Euklido ir Frechet atstumai (trajektorijų panašumui įvertinti), Jaro algoritmas eilutėms palyginti po simbolių, kaimyninių viršūnių atitikmenų (angl. *Neighbour Matching*) algoritmas grafų panašumui įvertinti, kampo tarp lygiagrečių 3D figūros briaunų apskaičiavimas, piršto judesių nustatymas (kryptinis greitis, pagreitis ir kt.), dienos gautų ir sudegintų kalorijų sąnaudos pagal metabolizmo lygį BMR (angl. *Basal Metabolic Rate*). Kalbos analizės metodai: garso signalo kokybė pagal vibracijų dažnį (angl. *pitch*), melo dažnio cepstraliniai koeficientai (angl. *Mel-frequency cepstral coefficients, MFCC*), išsibarsčiusios Gabor bangelės (angl. *Wavelet Scattering Transform, WST*), ir spektriniai algoritmai (nuolydis, centroidas, sklaida, entropija ir kt.).

3.2. Požymių atrinkimo algoritmų realizacija

Sukurtoje sistemoje palaikomi treniravimo ir testavimo režimai. Treniruočių režime pacientas kiekvieną užduotį atskirai gali išbandyti navigacijos juostoje. Testavimo metu visos užduotys yra pateikiamos atsitiktiniu būdu, o pabaigoje peržiūrima sveikatos būklė. Visos užduotys (jų numeracija atitinkamai T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, T12, T13, T14, T15, T0) yra kompiuterizuotos Android platformoje. 3.1 lentelėje pateikiamas visas sąrašas.

3.1. lentelė Visų sukurtos sistemos užduočių sąrašas

Nr.	Užduoties pavadinimas	Sutrikimai
T1	Nuoseklus lietimasis	Tremorai, kognityviniai
T2	Vaivorykštės spalvų lietimasis	Tremorai, kognityviniai
T3	Lietimai keliais pirštais	Tremorai, kognityviniai
T4	Archimedo spiralė	Tremorai, kognityviniai
T5	Bendros įžvalgos	Kognityviniai
T6	Orientacija (einamoji data)	Kognityviniai
T7	Paveikslų įvardinimas	Kognityviniai
T8	Panašumai, skaičiavimai	Kognityviniai
T9	3-D figūros konstravimas	Kognityviniai, tremorai
T10	Laikrodžio konstravimas	Kognityviniai, tremorai
T11	Žodinis sklandumas	Kognityviniai

T12	Vykdomasis: vizualinis loginis	Kognityviniai, tremorai
T13	Vykdomasis: problemų sprendimas	Kognityviniai, tremorai
T14	Balso įrašymas	Kalba
T15	Dienos energijos sąnaudos	Energijos sąnaudos
T0	Atmintis	Kognityviniai

3.2.1. T1, T2, T3: Lietimai pirštais

Šiose trijose užduotyse pacientui nurodoma liesti apskritos formos objektus vienu (T1, T2) ir keliais (T3) pirštais. Užduočių tikslas: paliesti rodomus objektus kuo arčiau centro ir kuo greičiau. Daroma prielaida, kad asmenys, kuriems pasireiškia neurologiniai sutrikimai, užduotis atliks lėčiau ir ne taip tiksliai.

T1 veikimo principas: tos pačios spalvos apskritimai (po 2, po 3 ir po 5) yra sugeneruojami skirtingose mobiliojo įrenginio ekrano vietose, be persidengiančių zonų, atsitiktiniu būdu. Aktyvus apskritimas, kurį reikia paliesti, pažymėtas juodu kontūru, išskiriant jį nuo kitų objektų.

T2 veikimo principas: 7 skirtingų spalvų (pagal vaivorykštės spektrą) apskritimai sugeneruojami skirtingose mobiliojo įrenginio ekrano vietose, be persidengiančių zonų, atsitiktiniu būdu. Ekrano apačioje pateikiama etiketė su užrašu, kurią spalvą reikia paspausti. Testo metu tiriamajam reikiama spalvą reikia paspausti 5 kartus (tiek kartų perbraižomas ekranas). Kiekvieną kartą apskritimo pozicija ekrane ir spalva, kurią reikia paliesti, skiriasi (nustatoma atsitiktiniu būdu). Pagrindinis testo skirtumas, palyginti su T1, yra tas, kad T2 labiau provokuoja neurologinių sutrikimų turintį pacientą, t. y. didesnis reikiamo paliesti apskritimo neapibrėžtumo faktorius.

T3 veikimo principas: apskritimai (režimai po 2, po 3 ir po 6) yra sugeneruojami skirtingose mobiliojo įrenginio ekrano vietose, be persidengiančių zonų, atsitiktiniu būdu. Režime po 2 ir 3 apskritimus tiriamasis turi paspausti apskritimus keliais pirštais (naudoti vieną ar abi rankas). Pirštų paspaudimo tvarka (eiliškumas) nėra svarbus. T3 teste, norint pereiti į kitą režimą, pirštų, liečiančių ekraną, skaičius turi sutapti su atvaizduojamų apskritimų skaičiumi (išpildžius šią sąlygą, pirštai atitraukiami nuo ekrano). Režime po 6 apskritimus (3 iš jų pateikiami viena spalva, 3 kita) turi būti naudojamos abi rankos (rekomenduojama mobilųjį įrenginį laikyti ant fiksuoto paviršiaus). Kiekviena ranka turi liesti tik 1 spalvą.

T1, T2, T3 užduočių požymių atrinkimo metodai (*vid* nurodo reikšmių vidurkius):

$$\mathit{delta}_{vid} = \sum_{i=1}^n \mathit{delta}_i / n, \quad \mathit{delta}_i = |p_i - st_i|, \quad \begin{array}{l} n = 10 \text{ (T1)} \\ n = 5 \text{ (T2)} \\ n = 11 \text{ (T3)} \end{array} \quad (1)$$

$$\mathit{reakcijos_laikas_vid} = \sum_{i=1}^n \mathit{rt}_i / n, \quad \begin{array}{l} n = 10 \text{ (T1)} \\ n = 5 \text{ (T2)} \\ n = 11 \text{ (T3)} \end{array} \quad (2)$$

$$\mathit{testo_trukme} = \sum_{i=1}^n t_i, \quad (3)$$

čia n – bendras apskritimų skaičius, p_i – i -tojo objekto centro koordinatės (x, y),

t_i – i -tojo objekto palietimo laikas (sekundės),

st_i – i -tojo ekrano palietimo koordinatės (x, y),

rt_i – i -tojo objekto palietimo reakcijos laikas (sekundės),

$\mathit{testo_trukmė}$ – bendras testo atlikimo laikas (sekundės).

3.2.2. T4: Archimedo spiralė

T4 testas susideda iš dviejų režimų. Pirmame mobiliojo įrenginio ekrane pateikiamas Archimedo spiralės, nubraižytos pagal laikrodžio rodyklę, kontūras. Tiriamajam (pacientui) reikia vienu pirštu sekti duotos spiralės trajektoriją. Atleidus pirštą 3 kartus arba 3 sekundes esant neveiklumo būsenai, pereinama į kitą režimą. Antrame režime atvaizduojamas Archimedo spiralės kontūras prieš laikrodžio rodyklę. Duodama 10 sekundžių įsidėmėti spiralę (pasibaigus laikui, ekranas išsivalo), po to vienu pirštu reikia nubraižyti tą patį kontūrą. Testas baigiamas, kai antrame režime pirštas atitraukiamas 3 kartus ar pasitaikius 3 sekundėms neveiklumo.

T4 požymių atrinkimo metodai (4 formulė). Įvertinamas apvestų (nubraižytų) unikalių spiralės taškų procentinis atitikmuo (mp).

$$mp = \left(\sum_{i=1}^{msp} psp_i \right) / nsp \cdot 100,$$

$$psp = \begin{cases} 1, & \sqrt{|x_{min} - x_{st}|^2 + |y_{min} - y_{st}|^2} \leq rsp, \\ 0, & otherwise \end{cases} \quad (4)$$

čia nsp – spiralės taškų skaičius, msp – vartotojo (paciento) ekrano paspaudimų skaičius, x_{st}, y_{st} – ekrano palietimo koordinatės (x, y),

x_{min}, y_{min} – artimiausias spiralės taškas duotam x_{st}, y_{st} , rsp – taško spindulys.

3.2.3. T5: Bendros įžvalgos

Tai pirmoji *SAGE* metodologijos užduotis. Ji nėra vertinama taškais, t. y. pacientui pateikiamas klausimynas (9 klausimai), kuriame reikia atsakyti apie demografiją, šeimos istoriją, atminties problemas, motorikos sutrikimus, asmenybės pokyčius, insulto požymius (jei įvyko), funkcinį veiksnumą. Užduotis kompiuterizuojama ekrane atsitiktiniu būdu pateikiant po 1 klausimą ir navigacijos mygtukus (pirmyn bei atgal). Galimi klausimų tipai: atviras ir su pasirinkimo variantais.

3.2.4. T6: Orientacija (einamoji data)

Antroji *SAGE* užduotis ir pirmoji vertinama taškais (maksimaliai galima surinkti 4). Pacientas turi įrašyti einamąją datą (metus Y , mėnesį M ir dieną D) iš atminties, nesinaudodamas papildomais resursais. Kompiuterizuojant užduotį yra patikrinama, ar paciento įrašyta data yra galiojanti, ir tik tuomet priskiriami taškai (suma) pagal formulę (5):

$$tsage1 = Y + M + D, \quad (5)$$

čia Y – teisingai įvesti metai (1 taškas), M – teisingai įvestas mėnuo (1 taškas), D – teisingai įvesta diena (2 taškai).

3.2.5. T7: Paveikslų įvardijimas

Trečioji *SAGE* užduotis (vertinama taškais): pacientui nurodoma įvardinti 2 paveikslus, atvaizduotus ekrane. Kompiuterizuojant užduotį visi 8 paveiksai buvo parinkti pagal tuos, kurie oficialiai pateikiami metodologijoje. Papildomai įtraukti 2 paveiksai dėl didesnio netikėtumo faktoriaus. Paveiksai spalvoti, jų formatas *.jpg*. Kompiuterizuojant užduotį, paveikslų rinkinys yra susiejamas su atsitiktinai parinkta *SAGE* grupe (variantu) G [1, 4], o pacientui pateikiami du. Paveikslas įvardinamas atviro klausimo forma (įvedimo laukas, kuriame įvedamas ne daugiau kaip vienas žodis) ir navigacijos mygtukais (pirmyn ir atgal), jei norima pasikoreguoti įvestą atsakymą.

Atliekant T7 užduotį maksimaliai galima surinkti 2 taškus (po 1 už teisingai įvardintą paveikslą). Tuo atveju, jei paveikslas įvardinamas dalinai teisingai, pvz., įvedama rašybos klaida ar sumaišoma raidžių tvarka, tada taikomas Jaro algoritmas (JA), skirtas dviejų teksto eilučių simboliniam palyginimui:

$$JA = \begin{cases} 0, & mj = 0 \\ \frac{1}{3} \left(\frac{mj}{|s_1|} + \frac{mj}{|s_2|} + \frac{mj-t}{mj} \right), & \text{kitais atvejais} \end{cases} \quad (6)$$

$$\left\lfloor \frac{\max(|s_1|, |s_2|)}{2} \right\rfloor - 1, \quad (7)$$

čia mj – sutampančių simbolių skaičius $|s_i|$ – eilutės s_i ilgis;

t – transpozicijos: skirtinga tvarka sutampančių simbolių skaičius / 2.

3.2.6. T8: Panašumai, skaičiavimai

Ši užduotis apima trečią, ketvirtą ir penktą *SAGE* metodologijos klausimus.

Pirmame klausime (K1) vartotojui pateikiami du daiktai (pvz.: liniuotė ir laikrodis) ir reikia įvardinti jų panašumą. Vertinamos dvi atsakymų formos (abstrakti ir konkreti). Maksimalus K1 klausimo įvertinimas yra 2 taškai (atsakant abstrakčia forma) ir 1 taškas (konkrečia). Tuo atveju, jei terminas įvardinamas dalinai teisingai (1 taškas), pvz., įvedama rašybos klaida ar sumaišoma žodžių ar raidžių tvarka, tada taikomas Jaro algoritmas (JA). Kitais atvejais skiriama 0 taškų.

Antrame klausime (K2) tikrinamos paciento matematinės atimties žinios. Pateikiama suma (P), už kurią buvo pirktą prekių, ir kasininkui duodamų pinigų vertė (V). Reikia apskaičiuoti gaunamą grąžą G, t. y. kiek pinigų liks. K2 klausimas vertinamas 1 tašku (atsakius teisingai) ir 0 kitu atveju.

Trečiame klausime (K3) tikrinamos paciento matematinės dalybos žinios. Pateikiama pinigų suma eurais (realus skaičius) ir nominalo vertė. Reikia suskaičiuoti, kiek nurodyto nominalo monetų sudaro duotą sumą. Atsakant į K2 ir K3 klausimus negalima naudotis skaičiuotuvu. K3 klausimas vertinamas 1 tašku (atsakius teisingai) ir 0 kitu atveju.

3.2.7. T9: 3-D figūros konstravimas

Septintoji *SAGE* užduotis: pacientas turi sukonstruoti 3-D figūrą. Pagal metodologiją, galimi 4 konstravimo variantai: kubas, stačiakampis gretasienis, kubas be paviršiaus, stačiakampis gretasienis be paviršiaus. Kompiuterizuotoje užduotyje palaikomi 2 režimai: pasiruošimo ir atlikimo.

Pasiruošimo režime pacientui parodoma atsitiktiniu būdu parinkta 3-D figūra iš *SAGE* grupės G [1,4]. Ekranu apačioje pateikiamos

instrukcijos įsidėmėti atvaizduotą figūrą (laikas nėra ribojamas). Ekranu centre pateikta figūra yra interpretuojama kaip 8 viršūnių grafas. Viršūnė apibūdinama centro koordinatėmis (x, y) ir jungtimis (briaunomis) į kitas viršūnes (numeruojamos nuo 1 iki 8 imtinai). Grafas transformuojamas į gretimumo matricą, kurios elemento reikšmė lyg vienetui, jei egzistuoja briauna tarp 2 viršūnių, kitu atveju 0. Į kitą režimą pereinama paspaudus ekrane.

Užduoties atlikimo režime atvaizduojama ta pati (kaip ir pirmame režime) 3-D figūra, tik be briaunų. Paciento užduotis – tinkamai sudėlioti briaunas tarp dviejų viršūnių, naudojant laisvo braižymo ekrane formą vienu pirštu. Formuojant briaunas, gali pasitaikyti įvairių klaidų, pvz.: bendras netinkamai sujungtų viršūnių skaičius, per ilgą piršto laikymą ekrane laikas (> 3 sekundės), per mažas pasirinktų viršūnių skaičius, per didelis pasirinktų viršūnių skaičius, egzistuojančios jungties (briaunos) atkartojimas. Įvykus klaidai, pacientas ekrano apačioje informuojamas, nurodant klaidos paskirtį ir tolimesnius veiksmus.

T9 požymiais atrinkti naudojamas kaimyninių viršūnių atitikmenų algoritmas [28] (8) (9) (10) (11) (12) (13) (14) ir kampo tarp lygiagrečių briaunų apskaičiavimas [29]. Maksimaliai galima surinkti 2 taškus.

$$x_{ij}^{k+1} \leftarrow \frac{s_{in}^{k+1}(i,j) + s_{out}^{k+1}(i,j)}{2} \quad (8)$$

$$s_{in}^{k+1}(i,j) \leftarrow \frac{1}{m_{in}} \sum_{l=1}^{n_{in}} x_{f_{ij}^{in}(l)g_{ij}^{in}(l)}^k \quad (9)$$

$$s_{out}^{k+1}(i,j) \leftarrow \frac{1}{m_{out}} \sum_{l=1}^{n_{out}} x_{f_{ij}^{out}(l)g_{ij}^{out}(l)}^k \quad (10)$$

$$m_{in} = \max(id(i), id(j)) \quad (11)$$

$$n_{in} = \min(id(i), id(j)) \quad (12)$$

$$m_{out} = \max(od(i), od(j)) \quad (13)$$

$$n_{out} = \min(od(i), od(j)), \quad (14)$$

čia $S_{in}(i,j)$ – įėjimo viršūnių laipsnio panašumas (A grafo i-toji ir B grafo j-toji viršūnės), $S_{out}(i,j)$ – išėjimo viršūnių laipsnio panašumas, $id()$ – įėjimo viršūnės laipsnis, $od()$ – out – išėjimo viršūnės laipsnis, $f(i,j)$ ir $g(i,j)$ – išskaičiavimo funkcijos (funkcijos, kurios suteikia maksimalią kiekvienos viršūnės panašumo reikšmę nurodytame viršūnių sąrašė).

3.2.8. T10: Laikrodžio konstravimas

Aštuntasis *SAGE* testas: reikia nubraižyti analoginį laikrodį. Kompiuterizuotoje užduotyje (pusiau automatinis įvertinimas) palaikomi du režimai: pasiruošimo (pirmasis) ir atlikimo (antrasis).

Pirmajame režime ekrano centre yra atvaizduojamas analoginis laikrodis, rodantis H [1, 12] valandų ir M [0, 60] minučių. H ir M reikšmės parenkamos atsitiktinai. Analoginį laikrodį sudaro šie 4 komponentai: laikrodžio kontūras ir skaičiai (ciferblatas), 12 skaičių, išdėstytų tinkama tvarka, rodyklių pozicijos ir jų žymėjimas (valandinė, minutinė). Ekrano apačioje pacientui yra pateikiamos instrukcijos įsidėmėti laikrodį (laikas nėra ribojamas). Į kitą režimą pereinama ekrane paspaudus pirštu.

Antrame (užduoties atlikimo) režime pacientui pateikiama aktyvi zona, kurioje reikia atkartoti įsidėmėtą laikrodį. Braižyti reikia pirštu arba išmaniuoju rašikliu. Galimos klaidos braižant laikrodį: išeinama už aktyvios zonos ribų, per ilgai sulaikomas pirštas ekrano vienoje vietoje. Atliekant užduotį maksimaliai galima surinkti 2 taškus (visi 4 komponentai teisingi), 1 taškas (3 komponentai teisingi) ir 0 kitu atveju.

3.2.9. T11: Žodinis sklandumas

Devintas *SAGE* testas: pacientui reikia užrašyti nurodytos kategorijos 12 skirtingų daiktų (elementų). Pagal metodologiją yra 4 skirtingos kategorijos: vaisiai ir daržovės, virtuvės daiktai, gyvūnai, šalys. Mobiliojoje programoje elementams išvardinti yra skiriamas įvedimo laukas (atvira forma). Viename negali būti įvedami daugiau nei 3 žodžiai. Laukas gali likti tuščias. Atliekant testą yra fiksuojama bendra testo atlikimo trukmė (sekundėmis).

Įvertinant pacientą, kiekvienai kategorijai yra surenkami teisingų atsakymo žodynai. Visi žodynai surinkti pasinaudojant atviraisiais švietimo ištekliais (AŠI) internete. Tuo atveju, jei terminas įvardinamas dalinai teisingai, pvz., įvedama rašybos klaida ar sumaišoma raidžių tvarka, taikomas Jaro algoritmas (JA). Atliekant užduotį maksimaliai galima surinkti 2 taškus (visi 12 elementų teisingi), 1 taškas (11 elementų teisingi) ir 0 kitu atveju.

3.2.10. T12: Vizualinis loginis

Dešimtas *SAGE* testas: pacientui reikia atkartoti schemą, pradedant trajektoriją nuo viršūnės 1 ir pakaitomis keičiant į raidę užbaigti viršūne F. Pagal metodologiją yra 4 skirtingi schemų variantai. Kompiuterizuotoje užduotyje yra du režimai: pasiruošimo (pirmasis) ir užduoties atlikimo (antrasis).

Pirmame režime pacientui pateikiama pavyzdinė schema (grafo viršūnės sujungtos pagal principą: 1(pradžia)→A→2→B→3→C (pabaiga)). Ekranu apačioje pacientui pateikiamos instrukcijos išsidėmėti sudarytos schemas jungimo principą. Į kitą režimą pereinama paspaudus ekrane.

Antrame režime pateikiama kitokia schemas konfigūracija, t. y. viršūnių-skaičių aibė {1, 2, 3, 4, 5, 6}, o viršūnių-raidžių aibė {A, B, C, D, E, F} be jungčių (briaunų). Paciento užduotis – tinkamai suformuoti reikiamas jungtis. Paciento būklei įvertinti adaptuojami T9 užduoties metodai: klaidų metrikos formuojant jungtis, kaimyninių viršūnių grafų palyginimo algoritmas (pavyzdinei ir paciento suformuotai schemoms), tiesinė interpoliacija. Atliekant užduotį maksimaliai galima surinkti 2 taškus.

3.2.11. T13: Problemų sprendimas

Vienuoliktas *SAGE* testas: pacientui pagal duotą pradinę taškų, sujungtų linijomis, konfigūraciją reikia suformuoti naują geometrinę figūrą. Užduotyje yra įvestas nurodytų operacijų limitas. Pagal metodologiją yra keturi 2-D figūrų transformavimo variantai. Pirmame ir antrame variantuose reikia sudėti 4 kvadratus iš 2 kvadratų ir 2 trikampių. Trečiame reikia išbraukti 3 linijas ir sudaryti 4 kvadratus iš turimų 5. Ketvirtame reikia išbraukti 2 linijas ir sudaryti 3 trikampus iš turimų 4. Atliekant bet kurį iš variantų, sugeneruotame taškų tinklelyje negali būti palikta jokių papildomų linijų.

Kompiuterizuojant užduotį, pacientui pirmiausia nurodoma mozaika (taškų tinklėlis) su pradine 2-D figūra. Priklausomai nuo atsitiktinai pateikto varianto (galimi keli teisingi dėlionės variantai), atliekamos linijos įterpimo ar pašalinimo operacijos. Įterpimo operacija atliekama du kartus nuosekliai vienu pirštu paspaudus taškus, jau sujungtus linija. Šalinimo operacija atliekama du kartus nuosekliai vienu pirštu paspaudus taškus, dar nesujungtus linija. Operacija atliekama

sėkmingai (tokiu atveju sumažinamas liekančių vykdymo operacijų skaičius), jei abu nuoseklūs piršto paspaudimai yra tikslūs, t. y. atitinkamo taško aplinkoje (išmatuojama pagal Euklido atstumą). Taškų tinkelis yra interpretuojamas kaip grafas, sudarytas iš viršūnių ir briaunų. Grafas yra transformuojamas į gretimumo matricą.

Paciento būklei įvertinti adaptuojami T9 ir T12 užduočių metodai: klaidų metrikos formuojant jungtis, kaimyninių viršūnių grafų palyginimo algoritmas (pavyzdinei ir paciento suformuotai 2-D figūroms). Atliekant užduotį maksimaliai galima surinkti 2 taškus (teisingai sukonstruota figūra per nurodyta operacijų skaičių). Dalinis įvertinimas (1 taškas) skiriamas, jei atliktas tinkamas operacijų skaičius, tačiau netinkamai suformuota figūra, ir 0 taškų kitu atveju.

3.2.12. T0: Atmintis

Paskutinis, dvilyktas *SAGE* testas siejasi su šeštuoju metodologijos klausimu. Šioje užduotyje pacientui pačioje testavimo pabaigoje reikia atsiminti duotą frazę ir ją įvesti ekrane. Frazė, sudaryta iš dviejų žodžių (*Testas atliktas, Pabaigtas testas*), pateikiama bet kuriuo testavimo momentu (šeštasis klausimas). Atliekant užduotį maksimaliai galima surinkti 2 taškus, t. y. pacientui teisingai atsiminus frazę (žodžiai gali būti apkeisti vietomis). 1 taškas skiriamas, jei įvestoje frazėje aptinkamas bent vienas reikiamas žodis, 0 taškų kitu atveju. Atliekant visą testavimo procedūrą T0 užduotis visą laiką pateikiama paskutinė, po jos apskaičiuojama bendra visų *SAGE* metodologijos užduočių taškų suma (jungia užduotis T6, T7, T8, T9, T10, T11, T12, T13 ir T0).

3.2.13. T14: Balso įrašymas

T14 veikimo principas: pacientas į mobiliojo įrenginio mikrofoną turi perskaityti duotą eilėrašį (pasirinktai anglų arba lietuvių kalba). Procedūra kartojama du kartus, deklamuojamas skirtingas eilėraštis. Kompiuterizuotoje užduotyje yra du mygtukai: *Įrašyti Balsą* ir *Sustabdyti įrašymą*. Užduoties laikas nėra ribojamas. Atlikus užduotį, išsaugojami du *.mpeg4* formato garso failai. T14 požymių atrinkimo metodai: garso signalo kokybė pagal vibracijų dažnį (angl. *pitch*) pagal liekamųjų harmonikų sumas (angl. *Summation of Residual Harmonics*) [30] (15):

$$SRH(f) = E(f) + \sum_{k=2}^N \left[E(k \cdot f) - E\left(\left(k - \frac{1}{2}\right) \cdot f\right) \right], \quad (15)$$

čia $E(f)$ – signalo spektro amplitudė (f – dažnis intervale $[F_{min}, F_{max}]$), apskaičiuotas kiekvienai Hanning tipo funkcijai k -toje harmonikoje, apimančioje kelis gauto liekamojo signalo ciklus, N – analizuojamų harmonikų skaičius.

Melo dažnio cepstraliniai koeficientai (angl. *Mel-frequency cepstral coefficients*) [31] (16):

$$mfcc_t[m] = \frac{1}{R} \sum_{r=1}^R \log(MF_t[r]) \cdot \cos \left[\frac{2\pi}{R} \left(r + \frac{1}{2} \right) m \right], \quad (16)$$

čia $MF_t[r]$ – Mel-dažnių skalės spektrinė analizė m -toje signalo dalyje t laiko momentu, kai $r = 1, 2, \dots, R = 22$.

Išsibarsčiusios Gabor bangelės $\varphi_\omega(t)$ (angl. *Wavelet Scattering Transform*) [32], [33] (17):

$$\varphi_\omega(t) = \omega \cdot \varphi(\omega t), \quad \check{\varphi}_\omega(s) = \check{\varphi} \left(\frac{s}{\omega} \right), \quad (17)$$

čia $\check{\varphi}$ – centrinis dažnis normalizuotas iki 1, $\omega = 2^{k/Q}$, Q – bangų skaičius vienoje oktavoje, $k \in \mathbb{Z}$. $\check{\varphi}$ eilė yra Q^{-1} , skirtas visai dažnių skalei su bangų filtrais analizuoti.

Spektrinė entropija, [34] (18):

$$entropija = \left(- \sum_{k=b_1}^{b_2} s_k \log(s_k) \right) \log(b_2 - b_1), \quad (18)$$

čia s_k spektrinė reikšmė k -tame signalo segmente, b_1, b_2 – analizuojamos signalo spektro juostos ribinės vertės (pradžia ir pabaiga).

3.2.14. T15: Dienos energijos sąnaudos

Šiame teste vertinamos paciento dienos kalorijų vertės. Pirmiausia apskaičiuojama pacientui rekomenduojama energijos sąnaudų norma *TDEE* (angl. *Total Daily Energy Expenditure*) pagal (19) ir (20) formules [35]:

$$BMR = \begin{cases} (6,25 \cdot H_{tdee} \cdot 100) + (10 \cdot W_{tdee}) - (5 \cdot A_{tdee}) + 5, & moterys \\ (6,25 \cdot H_{tdee} \cdot 100) + (10 \cdot W_{tdee}) - (5 \cdot A_{tdee}) - 161, & vyrai \end{cases} \quad (19)$$

$$TDEE = BMR \cdot FAL, \quad (20)$$

čia BMR – bazinė medžiagų apykaitos norma (Mifflin St Jeor formulė), FAL – fizinis aktyvumo lygis (sėdimas, lengvas (1–3 dienos per savaitę), vidutinis (3–5 dienos), didelis (6–7 dienos), atletas (2 kartai per dieną), H_{tdee} – ūgis (metrais), W_{tdee} – svoris (kg), A_{tdee} – amžius (metais).

Antrame etape surenkamas paciento dienos suvartoto maisto sąrašas. Pagal įvestus maisto produktus gautos kalorijos apskaičiuojamos (21) naudojant AŠI [36], kuriame pateikiamos maisto produkto kalorijų normos 100 g:

$$p_{d_gautos} = \sum_{i=1}^{n_{tdee}} \frac{m_{pacientas} \cdot Kal_{100}}{m_{norma}}, \quad (21)$$

čia p_{d_gautos} – paciento dienos gautos kalorijos (su maistu), n_{tdee} – paciento įvestų produktų skaičius, $m_{pacientas}$ – įvesto produkto kiekis (gramais), Kal_{100} – kalorijų norma (100 g.), m_{norma} – reikšmė, lygi 100.

Trečiame etape surenkamas paciento dienos atliktų fizinių veiklų sąrašas. Pagal įvestas veiklas sudegintos kalorijos apskaičiuojamos (22) naudojant AŠI [37], kuriame pateikiamos fizinių veiklų MET koeficientų reikšmės:

$$p_{d_sudegintos} = \sum_{i=1}^{m_{tdee}} \frac{MET_i \cdot W_{tdee} \cdot D_i}{60}, \quad (22)$$

čia m_{tdee} – paciento įvestų dienos fizinių veiklų skaičius, MET_i – fizinės veiklos MET koeficientas (angl. *Metabolic equivalent of task*), D_i – i -tos veiklos trukmė minutėmis, p_{d_burned} – paciento dienos sudegintos kalorijos, W_{tdee} – paciento svoris (kilogramais).

Paskutiniame etape apskaičiuojamas paciento dienos kalorijų balansas (23). Siekiant palaikyti vienodą kūno masę, $p_{d_balansas}$ turėtų būti artimas 0, o p_{d_gautos} – artima apskaičiuotai TDEE normai reikšmė.

$$p_{d_balansas} = p_{d_gautos} - p_{d_sudegintos}. \quad (23)$$

3.3. Priemonių ir metodų skyriaus apibendrinimas

Trečiame darbo skyriuje pateikiamas detalus požymių atrinkimo algoritmų ir jų įgyvendinimo mechanikų aprašas. Sukurti algoritmai yra integruoti Android OS mobiliojoje programėlėje NITS (angl. *Neural Impairment Test Suite*). Iš viso kompiuterizuota 16 užduočių, generuojančių 238 skirtingus požymius. Šios užduotys realizuotos siekiant aptikti tremorų (motorikos), kognityvinius, kalbos ir energijos sąnaudų sutrikimus (4 grupės) pacientams, kenčiantiems nuo centrinės nervų sistemos sutrikimų (CNSS). Užduotys projektuotos pagal savarankiškai atliekamo pažintinio testavimo (angl. *SAGE*) mediciniškai patvirtintą metodologinę priemonę, skirtą kognityviniams sutrikimams (pvz., atminties praradimas) nustatyti. T5, T6, T7, T8, T9, T10, T11, T12

T13 ir T0 užduotys atitinka *SAGE* metodologinius klausimus. Disertacijoje *SAGE* metodologija praplėsta užduotimis tremorų (T1, T2, T3, T4), kalbos (T14) ir energijos sąnaudų (T15) sutrikimams aptikti, remiantis kitų autorių susijusių darbų moksliniais tyrimais ir išvadomis.

Kompiuterizavus užduotis, CNSS pacientams suteikiama inovatyvi galimybė automatiškai būdu stebėti savo sveikatos būklę namuose ar medicinos įstaigoje. Tam reikalinga tik viena išmanioji sąsaja – mobilioji programėlė, o sveikatos būklės įvertinimas atliekamas elektroniniu (ne tradiciniu popieriniu) būdu. Sukurtoje sistemoje atskira užduotis gali atpažinti kelis neurologinius sutrikimus vienu metu, pvz., konstruojant 3-D figūrą tikrinami paciento rankų tremorai ir vizualinės erdvinės pažintinės reakcijos (skirtingai nei tradiciniame *SAGE* teste).

Išskirti požymiai duomenų gavybos (angl. *Data Mining*) procese lemia ypatingą vaidmenį sprendimų priėmimo sistemose. Toks procesas aprašomas disertacijos eksperimentinių tyrimų dalyje (4 skyrius), t. y. duomenų rinkimas naudojant NITS įrankį ir tiriamųjų klasifikavimas į sveikų ir sergančių individų grupes.

4. Eksperimentiniai tyrimai

Šiame disertacijos skyriuje pirmiausia aprašomas duomenų rinkimo eksperimentas, naudojant sukurtą mobiliąją programėlę NITS: tiriamųjų skaičius, duomenų formalizavimo procedūros ir sukauptas duomenų rinkinys. Tuomet aprašomi mokymo su mokytoju (angl. *Supervised Learning*) klasifikavimo metodai. Tyrime naudota 17 skirtingų klasifikatorių, adaptuojamų sprendimų priėmimo sistemose. Darbe aprašomas klasifikatorių jungimo į „ansamblį“ hibridinis modelis, pagerinantis klasifikavimo rezultato tikslumą, palyginti su atskirais metodais. Skyriaus pabaigoje analizuojami keturi eksperimentai, skirti suprojektuotiems modeliams validuoti ir klasifikuoti tiriamuosius į sveikų ir sergančių grupes pagal tremorų, kognityvinius, kalbos ir energijos sąnaudų neurologinius sutrikimus.

4.1. Duomenų rinkimo eksperimentas: dalyviai, procedūra ir duomenų rinkinys

Tyrime dalyvavo 15 asmenų iš Lietuvos: 7 turintys neurologinių sutrikimų (3 Hantingtono pacientai, 2 iš jų vyresni nei 40 metų, 1 juvenilinės formos (18 metų) atvejis, 1 Parkinsono liga sergantis 74 metų pacientas, 1 po insulto (60 metų), 1 sergantis demencija (40 metų), 1 sergantis cerebriniu paralyžiumi (20 metų)) ir 8 sveiki (amžius nuo 20 iki 78 metų). Visi dalyviai atliko visą šešiolikos užduočių rinkinį. Tyrimo pacientams pasireiškė neurologiniai sutrikimai, apibūdinami simptomais, kurie atsiranda ankstyvoje stadijoje, t. y. rankų drebėjimas, koordinacijos sutrikimai, dėmesio sutelkimo problemos, raumenų sąstingiai, atminties praradimas, sprendimų priėmimo sunkumai ir pan. [25], [26].

Duomenų rinkinyje sukaupta 150 įrašų (kiekvieną eilutę sudaro 238 agreguotų požymių), 2019 m. vasario–rugpjūčio mėnesiais. Iš viso atliktos 5 rinkimo iteracijos (pakartotinės pacientų vizitacijos). Disertacijos autorius kiekvieno paciento akivaizdinio lankymo metu prižiūrėdavo testo atlikimo eigą. Testo, atliekamo su mobiliuoju telefonu ar planšete, trukmė ~ 20 minučių. Atskira duomenų rinkimo iteracija būdavo atkartojama įvertinant laiko skalę, t. y. po paskutinio testo bandymo praeidavo apytiksliai 1 mėnuo. Surinkti duomenys sužymėti naudojant sveiko (skaičius 0) ir sergančio (skaičius 1) asmens objektyvų sveikatos būklės įvertinimo kriterijų (nurodomas mobiliojoje programėlėje prieš atliekant testą). Tyrime naudojami tik Kauno

technologijos universiteto (KTU) Informatikos fakulteto leidimu patvirtinti duomenys, tiriamiesiems pasirašant informuoto sutikimo formą, leidžiančią surinktus duomenis naudoti mokslo tikslams.

4.2. Klasifikavimo metodai sprendimų priėmimo sistemose

4.2.1. Mokymo su mokytoju klasifikatoriai

Mokymo su mokytoju (angl. *Supervised Learning*) algoritmai yra pritaikomi disertacijoje nagrinėjamam klasifikavimo uždaviniui spręsti. Toks pasirinkimo motyvas grindžiamas tuo, kad, testuojant tiriamuosius, duomenys buvo sužymėti pagal sveiko arba sergančio sveikatos būklės įvertinimo kriterijų. Darbe analizuojama netiesinė sistema (modelio įėjimai nėra proporcingi išėjimams). Klasifikavimo problema yra anksčiau nenagrinėta ir inovatyvi (modelio įėjime naudojamas unikalus požymių rinkinys). Dėl šių priežasčių nagrinėjamos penkios metodų grupės: 1) funkciniai klasifikatoriai (aprašomi matematinėmis išraiškomis); 2) Bajeso (angl. *Bayesian*); 3) sprendimų medžiai (angl. *Decision Trees*); 4) taisyklės (angl. *Rules*); 5) *Lazy* tipo klasifikatoriai.

Metodai: 1) atraminės vektorių mašinos (angl. *Support Vector Machines, SVM*) [38] [39]; 2) dirbtiniai neuroniniai tinklai su daugiasluoksniu perceptronu (angl. *Artificial Neural Networks Multilayer Perceptron, ANN-MLP*) [41]; 3) K-artimiausių kaimynų metodas (angl. *K-Nearest Neighbours, KNN*) [40]; 4) nuosekliai minimalus optimizavimas atraminei vektorių mašinai (angl. *Sequential minimal optimization, SMO*) [42] [43] [44]; 5) tiesinė (angl. *Linear discriminant analysis, LDA*) [45] ir Fišerio (angl. *Fisher's Linear discriminant analysis (FLDA)*) [46] diskriminanto analizės; 6) gilus mokymo tinklai (angl. *Deep Learning Networks DNN-LSTM,)* [47]; 7) atsitiktiniai miškai (angl. *Random Forests, RF*) [48]; 8) Bajeso tinklai (angl. *Bayes Nets*) [49]; 9) naivus Bajeso algoritmas (angl. *Naive Bayes*) [50]; 10) sprendimų medžiai (angl. *Decision Tree J48* ir *Decision Stump*) [51] [55]; 11) stochastinio gradiento minimumo algoritmas (angl. *Stochastic gradient descent, SGD*) [52]; 12) logistiniai medžiai (angl. *Logistic Model Trees, LMT*) [53] [54]; 13) balsavimo perceptronas (angl. *Voted Perceptron*) [56]; 14) logistinė regresija (angl. *Logistic Regression*) [57]; 15) sprendimų lentelės (angl. *Decision Tables*) [58]; 16) lokalus svorinis mokymasis (angl. *Local Weighted Learning, LWL*) [59] [60].

Klasifikatorių tikslumui pagerinti naudojamas adaptyvusis spartinimo algoritmas AdaBoostM1 [61], atributams parinkti – principinė komponentių analizė (angl. *Principal Component Analysis, PCA*).

Kituose skyriuose rezultatai pristatomi naudojant pateiktą klasifikavimo algoritmą angliškus trumpinius.

4.2.2. Klasifikatorių „ansamblis“ (hibridinis modelis)

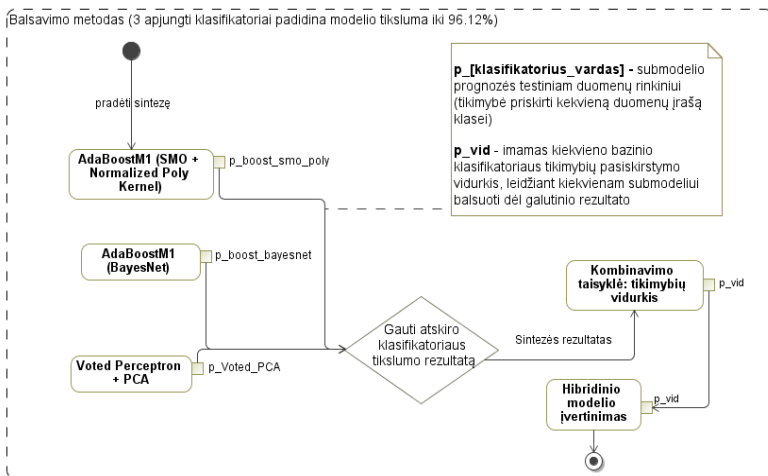
Hibridinis klasifikavimo modelis sujungia skirtingų žinių atvaizdavimo tipų stipriąsias puses, pvz., taikant funkcijų sintezės techniką ir Bajeso metodus. Motyvas kurti hibridinį klasifikavimo modelį grindžiamas keliomis prielaidomis. Pirmasis aspektas nurodo, kad sujungiant klasifikatorius (sintezė) gaunamas didesnis tikslumas priimant galutinį sprendimą [63]. Antrasis aspektas pažymi, kad, ieškant tikslesnio klasifikavimo modelio, prioritetas yra kombinuoti atskirus klasifikatorius, o ne ieškoti požymių poaibių rinkinių [64].

Mokymo su mokytoju algoritmų hibridizacijai reikia pritaikyti metodus, skirtus tikslinių klasių (angl. *target class*) žymių (angl. *label*) sujungtai reikšmei nustatyti. Disertacijoje nagrinėjami tikimybių vidurkių (angl. *average of probability*) kombinavimo taisyklė ir balsavimo (angl. *Vote*) algoritmai [65]. Kiekvienu atveju galutinis sprendimas $H(X)$ yra i klasė, kuriai apskaičiuojama didžiausia tikimybė $P_{MAX_i}(X)$ (24) (25).

$$H(X) = \max_{i=1..n}(P_{MAX_i}(X)) , \quad (24)$$

$$P_{MAX_i}(X) = \frac{1}{m} \sum_{j=1}^m P_j(w_i | x) , \quad (25)$$

čia n – tikslinių klasių skaičius, m – sujungime naudojamų klasifikatorių skaičius, $P_j(w_i | x)$ – i -tojo duomenų įrašo priskyrimo klasei x tikimybė, įvertinama j -tuoju klasifikatoriumi.



4.1 pav. Siūlomas hibridinis klasifikavimo modelis (3 sujungti klasifikatoriai)

Išskiriami du autoriaus sukurtu hibridinio modelio naujumo aspektai: 1) pasiekiamas didesnis unikaliam suformuotam požymių rinkinio klasifikavimo tikslumas, palyginti su atskirais modeliais; 2) hibridinis modelis sukomponuojamas naudojant kitas hibridines komponentes, pvz., adaptyvusis spartinimo algoritmas AdaBoostM1, sujungtas su Bajeso tinklu.

4.1 pav. iliustruotas siūlomas hibridinis modelis. Naudojama trijų klasifikatorių kombinacija (schemoje pateikiami angliški metodų trumpiniai):

[AdaBoostM1 (SMO)] + [Voted Perceptron with PCA] + [AdaBoostM1 (Bayes Net)]. Disertacijoje atliktas įvadinis (pilotinis) kombinacijų sudarymo tyrimas, atsitiktinai parenkant $N = 50$ skirtingų kombinacijų. Buvo rastos kelios klasifikatorių kombinacijos (potencialiai lokalūs optimumai), pagerinančios modelio tikslumą (palyginti su atskirais klasifikatoriais). Pirmenybė teikiama kombinacijoms, generuojančioms didžiausią tikslumą (šiam kontekste tai 96,12 proc.) su mažiausiomis modelio mokymo laiko sąnaudomis (0,59 sekundės).

4.3. Atlikti eksperimentai, skirti sergančių ir sveikų asmenų klasifikavimo problemai išspręsti

Disertacijoje pateikimas požymių sujungimo (sintezės) metodas, t. y. naudojamas visas rinkinys (daugiau nei 230 požymių) ir autoriaus sukurtas duomenų rinkinys analizuojant sveikų ir sergančių individų grupes. Iš viso atlikti keturi eksperimentai (du iš jų skirti galimiems balso sutrikimams aptikti). Klasifikavimas pagal tremorų, kognityvinių ir energijos sąnaudų požymių sintezės metodą atliktas *Weka* įrankiu [62]. Balso sutrikimams klasifikuoti naudota MATLAB Audio Toolbox R2019a programinė įranga.

1. 1 eksperimentas (E1): požymių paskirstymas atskiroms užduotims;
2. 2 eksperimentas (E2): visų (>230) savybių sujungimas ir hibridinio modelio sudarymas (klasifikatoriaus lygmeniu);
3. 3 eksperimentas (E3): balso įrašų (sukauptų T14 užduotyje) klasifikavimas naudojant rekurentinius neuroninius tinklus (RNN) su ilga trumpalaikė atmintimi (angl. *LSTM*);
4. 4 eksperimentas (E4): balso įrašų (sukauptų T14 užduotyje) klasifikavimas naudojant išsibarsčiusių Gabor bangelių transformaciją (angl. *Wavelet Scattering Transform*).

4.3.1. E1: sergančių ir sveikų asmenų klasifikavimas paskirstant požymius atskiroms užduotims

E1 eksperimente buvo sudaryta 14 skirtingų klasifikatorių, pritaikytų atskiroms (pavienėms) užduotims. Tokio eksperimento tikslas – prognozuoti perspėjimus apie pablogėjusią sveikatos būklę, atsižvelgiant tik į vienos konkrečios užduoties rezultatus. Geriausias pasiektas E1 rezultatas užfiksuotas T9: 3-D figūros konstravimo užduočiai (lokalaus svorinio mokymo klasifikatorius 91,50 proc.). Įvertinus situaciją, kad E1 klasifikavimo tikslumą dar reikia gerinti, lygiagrečiai buvo suprojektuotas eksperimentas (E2) naudojant visą 238 požymių rinkinį.

4.3.2. E2: sergančių ir sveikų asmenų klasifikavimas naudojant visą požymių rinkinį

4.1 lentelėje pateikiami gauti klasifikavimo rezultatai naudojant visą požymių rinkinį (tremorų, kognityviniams ir energijos sąnaudų

sutrikimams aptikti). Modelių tikslumas įvertintas pritaikius 10 dalių kryžminio validavimo procedūrą (angl. *Cross validation*). Statistiniams modelių įvertinimui atlikti naudojamos *TPR* (angl. *True Positive Rate*) ir *FPR* (angl. *False Positive Rate*) metrikos (26), (27):

$$TPR = \frac{TP}{(TP + FN)} \tag{26}$$

$$FPR = \frac{FP}{(FP + TN)} \tag{27}$$

TP – sergančių tiriamųjų, kuriems diagnozuojamas neurologinis sutrikimas, skaičius, *FP* – nesergančių tiriamųjų, kuriems diagnozuojamas neurologinis sutrikimas, skaičius, *FN* – sergančių tiriamųjų, kuriems nėra diagnozuojamas neurologinis sutrikimas, skaičius, *TN* – nesergančių tiriamųjų, kuriems nėra diagnozuojamas neurologinis sutrikimas, skaičius.

4.1. lentelė Sveiko ir sergančio asmens visų požymių klasifikavimo rezultatai (modelių tikslumo statistika, N=10 kryžminio validavimo procedūra)

Klasifikatorius	Tipas	Tikslumas, %	TPR	FPR
AdaBoostM1 (Decision Stump)	Spartinimas ir <i>Lazy</i>	93,023	0,930	0,081
AdaBoostM1 (RandomForest)	Spartinimas ir medis	94,573	0,946	0,078
AdaBoostM1 (J48)	Spartinimas ir medis	86,046	0,860	0,169
Logistic Regression	Funkcija	89,922	0,899	0,100
AdaBoostM1 (LMT)	Spartinimas ir medis	91,472	0,915	0,083
SGD	Funkcija	91,472	0,915	0,091
AdaBoostM1 (NaiveBayes)	Spartinimas ir Bajeso	88,372	0,884	0,103
AdaBoostM1 (ANN-MLP)	Spartinimas ir funkcija	92,480	0,922	0,086
AdaBoostM1 (SMO)	Spartinimas ir funkcija	92,248	0,922	0,078
AdaBoostM1 (k-NN)	Spartinimas ir <i>Lazy</i>	94,573	0,946	0,071
AdaBoostM1 (LWL)	Spartinimas ir <i>Lazy</i>	91,472	0,915	0,091

AdaBoostM1 (BayesNet)	Spartinimas ir Bajeso	93,798	0,938	0,083
SVM (sigmoid) + PCA	Funkcija	91,472	0,915	0,091
SVM (linear) + PCA	Funkcija	92,248	0,922	0,086
(LDA), FLDA	Funkcija	92,248	0,922	0,100
Decision Table	Taisyklės	86,821	0,868	0,172
DNN (LSTM)	Funkcija	94,573	0,946	0,056
Voted Perceptron + PCA	Funkcija	93,023	0,930	0,081
Hibridinis modelis (autoriaus sukurtas)	Klasifikatorių sintezė	96,124	0,961	0,054

4.2 lentelėje pateikiama tyrimo klasifikatorių klaidų statistika. Naudotos metrikos: Kappa koeficientas (KP) (28), vidutinė absoliutinė paklaida (angl. *Mean Average Error, MAE*) (29), vidutinė kvadratinė absoliutinė paklaida (angl. *Root Mean Squared Error, RMSE*) (30), santykinė absoliuti paklaida (angl. *Relative absolute error RAE*) (31), kvadratinė santykinė paklaida (angl. *Root relative squared error, RRSE*) (32).

$$KP = \frac{p_0 - p_e}{1 - p_e} = 1 - \frac{1 - p_0}{1 - p_e}, \quad (28)$$

čia p_0 – santykinis vertintojų susitarimas, p_e – hipotetinė susitarimo tikimybė, taikoma kiekvienam stebėtojuj turimai duomenų imčiai.

$$MAE_i = \frac{1}{n} \sum_{j=1}^n |P_{(i,j)} - T_j| \quad (29)$$

$$RMSE_i = \sqrt{\frac{1}{n} \sum_{j=1}^n (P_{(i,j)} - T_j)^2} \quad (30)$$

$$RAE_i = \left(\sum_{j=1}^n |P_{i,j} - T_j| \right) / \left(\sum_{j=1}^n |T_j - \bar{T}_j| \right), \quad (31)$$

$$RRSE_i = \sqrt{\left(\left(\sum_{j=1}^n (P_{(i,j)} - T_j)^2 \right) / \left(\sum_{j=1}^n (T_j - \bar{T}_j)^2 \right) \right)} \quad (32)$$

$$\bar{T}_j = \frac{1}{n} \sum_{j=1}^n T_j \quad (33)$$

čia $P_{i,j}$ - atskiro modelio i sprognozuota reikšmė j pavyzdžiui (iš n stebėjimų), T_j – tikslinė (etaloninė) reikšmė j duomenų pavyzdžiui.

4.2. lentelė Sveiko ir sergančio asmens visų požymių klasifikavimo rezultatai (modelių klaidų statistika)

Klasifikatorius	KP	MAE	RMSE	RAE %	RRSE %
AdaBoostM1 (Decision Stump)	0,852	0,07	0,225	15,695	46,204
AdaBoostM1 (RandomForest)	0,883	0,243	0,284	51,290	58,373
AdaBoostM1 (J48)	0,701	0,145	0,369	29,570	75,910
Logistic Regression	0,790	0,099	0,3053	21,013	62,664
AdaBoostM1 (LMT)	0,822	0,080	0,263	17,008	54,160
SGD	0,821	0,081	0,292	17,945	59,934
AdaBoostM1 (NaiveBayes)	0,762	0,127	0,339	26,906	69,602
AdaBoostM1 (ANN-MLP)	0,836	0,086	0,255	18,131	52,356
AdaBoostM1 (SMO)	0,771	0,108	0,329	22,839	67,614
AdaBoostM1 (k-NN)	0,884	0,061	0,231	13,008	47,436
AdaBoostM1 (LWL)	0,821	0,096	0,271	20,323	55,733
AdaBoostM1 (BayesNet)	0,867	0,063	0,244	13,283	50,185
SVM (sigmoid) + PCA	0,821	0,085	0,292	17,945	59,933
SVM (linear) + PCA	0,836	0,077	0,278	16,313	57,147
(LDA), FLDA	0,834	0,484	0,484	102,024	99,526
Decision Table	0,715	0,243	0,333	51,233	68,445
DNN (LSTM)	0,886	0,068	0,204	14,485	41,936
Voted Perceptron + PCA	0,852	0,071	0,264	15,077	54,297
Hibridinis modelis (autoriaus sukurtas)	0,917	0,134	0,207	28,274	42,163

4.3.3. E3: balso sutrikimų klasifikavimas naudojant dvikryptį rekurentinį neuroninį tinklą su ilga trumpalaikė atmintimi

Disertacijoje atlikti du eksperimentai, skirti kalbos sutrikimams aptikti. Abiejų eksperimentų tikslas – analizuoti T14 užduotyje surinktus garso įrašus ir suklasifikuoti tiriamuosius į sveikų ir sergančių grupes (klases). Taip sudaromi galimo kalbos sutrikimo prognozavimo modeliai.

E3 eksperimente balso signalams (64 kbps .mp3 formatas) pritaikomas gilaus mokymo dvikryptis rekurentinis neuroninis tinklas su ilga trumpalaikė atmintimi (angl. *Bidirectional Long Short-Term Memory, BiLSTM*) [66], [67]. Norint klasifikuoti garso įrašus su BiLSTM, pirmiausia reikia atlikti signalo apdorojimo ir požymių atrinkimo veiksmus. Balso signalas apdorojamas eliminuojant tylos segmentus ir išskiriant aktyvias kalbėjimo sritis (aktyviSritis). Taikomas slenkstinių (angl. *threshold*) reikšmių parinkimo metodas signalo energijai (T_E) ir centrui (T_C) (34):

$$\text{aktyviSritis} = \text{signalEnergija} \geq T_E, \text{centras.} \leq T_C \quad (34)$$

E3 eksperimente naudojami požymiai (*SRH, MFCC*, entropija), aprašyti 3.2.13 poskyryje (gali būti kombinuojami įvairiais būdais, pvz.: įdedamas naujas, pašalinamas egzistuojantis ar keičiama jų tvarka). Sudarytos sujungtos požymių sekos (angl. *feature sequence*) yra normalizuojamos ir naudojamos BiLSTM tinklui mokytis. Klasifikatoriui reikia apdoroti požymių rinkinius, kuriuose sugeneruojami milijonai įrašų, todėl tinklo mokymas vykdomas grafiniame procesoriuje (angl. *GPU*), taikant lygiagrečius skaičiavimus. Tyrimo metu pasiektas 86,63 proc. testavimo rinkinio klasifikavimo tikslumas ir 100 proc. treniravimo rinkinio tikslumas (konverguojantis BiLSTM tinklas).

4.3.4. E4: balso sutrikimų klasifikavimas naudojant išsibarsčiusių (Gabor) bangelių transformaciją

E4 eksperimente balso signalams (1411 kbps .wav formatas) pritaikomas išsibarsčiusių (Gabor) bangelių transformacijų (angl. *Wavelet scattering transform, WST*) metodas [68]. Parenkamas 2^{19} natūrinio logaritmo signalo ilgis. Bangelių konfigūracijai reikia nurodyti tris parametrus (požymius): 1) laiko trukmės invariantiškumas (visiška ar dalinė nepriklausomybė nuo aplinkos veiksnių); 2) bangelių filtrų (atskiria įėjimo pradinį signalą į komponentus, kiekvienas iš jų su

skirtingo dažnio antrinėmis juostomis) skaičius; 3) bangelių skaičius oktavoje. Naudojami du filtrai: 8 oktavų (pirmas) ir 1 oktavos (antras) bangelėje. Nustatoma 0,5 sekundės laiko invariantiškumo reikšmė. Panašiu principu, kaip ir pirmame eksperimente, yra sudaromos sujungtos požymių sekos (matricos), o modelio mokymas atliekamas grafiniame procesoriuje, taikant lygiagrečius skaičiavimus. Tyrime pasiektas 100 proc. tikslumas naudojant atraminių vektorių mašinų (angl. *Support Vector Machines, SVM*) klasifikatorių su trečios eilės daugianarių branduoliu (angl. *polynomial kernel*).

4.4. Eksperimentų apibendrinimas

Visi atlikti eksperimentiniai tyrimai yra skirti tiriamiesiems suskirstyti į sveikų ir sergančių (su neurologiniais sutrikimais) asmenų klases, t. y. sprendžiama dvejetainė klasifikavimo problema. Rezultatams pristatyti yra naudojamas autoriaus sukurtas duomenų rinkinys. Duomenys (7 pacientai ir 8 sveiki asmenys) surinkti Android (6.0 ir aukštesnė versija) mobiliąja programėle, t. y. planšete, kurios matmenys: 10.1” ekrano dydis, 1280 x 800 pikselių rezoliucija.

Tremorų, kognityviniams ir energijos sąnaudų sutrikimams aptikti buvo pritaikyti naujausi ir populiariausi mokymo su mokytojų algoritmai, adaptuoti visam požymių (> 230) rinkiniui. Pastebėta, kad, suprojektavus hibridinį modelį pagal atskirų klasifikatorių sintezės metodą, gaunamas didesnio tikslumo, t. y. 96,12 proc. ir mažesnių paklaidų (Kappa koeficientas išauga iki 0,917) modelis.

Autorius sukūrė du atskirus eksperimentus sutrikimams pagal balso signalą nustatyti. Pirmame naudojamas dvikryptis rekurentinis neuroninis tinklas (BiLSTM) su ilga trumpalaikė atmintimi, o antrame – išsibarsčiusių (Gabor) bangelių transformacijos (WST). Gauti klasifikavimo rezultatai (BiLSTM 86,63 proc., WST 100 proc.) suteikia optimistinius lūkesčius pacientų kalbos sutrikimams prognozuoti.

Sukurti klasifikavimo modeliai gali būti papildomai validuojami, praplečiant duomenų rinkinį nenagrinėtomis ligomis, pvz., Alzheimerio liga sergančių pacientų stebėjimo įrašais.

5. Išvados ir ateities darbai

5.1. Išvados

1. Naujausių informatikos mokslo srities darbų analizė pateikia išsamų duomenų gavybos ir mašininio mokymosi metodų, skirtų centrinės nervų sistemos sutrikimų (CNSS) turinčių asmenų sveikatos būklei įvertinti, taikymo galimybes. Mažuma egzistuojančių tyrimų nagrinėja Hantingtono ligą (HL). CNSS pacientų sveikatos būklei stebėti pasitelkiami įvairūs būdai, turintys apribojimų: vieni nenaudoja mobiliųjų įrenginių, kiti pritaiko papildomus trečiųjų šalių nešiojamuosius prietaisus, kurie paprastai padidina streso faktorių pacientams ir yra nepatogūs. Nė viename iš aprašytų susijusių darbų nėra nagrinėjama išplėstinė metodologija, jungianti požymių atrinkimo algoritmus, skirtus CNSS pacientų tremorų (motorikos), kognityviniams, kalbos ir energijos sąnaudų simptomams įvertinti.
2. Disertacijoje egzistuojančių sprendimų trūkumai sprendžiami surenkant apriorines žinias iš medikų ir kompiuterizuojant savarankiškai atliekamo pažintinio testavimo *SAGE* (angl. *Self-administered cognitive testing*) metodologiją. Pasiūlyta sistema – domenų rinkimo karkasas, jai reikalinga tik viena išmanioji neinvazinė sąsaja, t. y. mobilusis prietaisas ar planšetinis kompiuteris, skirtas CNSS pacientų sveikatos būklei tikrinti. Sistemoje realizuota 16 užduočių, 12 iš jų atkartojant *SAGE* testus, kitose pritaikoma pirštų lietimui, spiralės kontūro sekimo (braižymo), balso įrašymo ir energijos sąnaudų apskaičiavimo išplėstinė metodologija. Karkase generuojama daugiau nei 230 požymių (piršto judesių sekimas, testų trukmės, atstumo tarp geometrinių figūrų 2-D erdvėje įvertinimas, dviejų grafų panašumo nustatymas, laikrodžio paveikslų rinkimas, *SAGE* taškai, balso įrašai, dienos kalorijų balansai ir t. t.). Kitas inovatyvus disertacijos aspektas yra naujai suprojektuotas duomenų rinkinys, suteikiant galimybę kitiems informatikos mokslininkams juo naudotis mašininio mokymo tyrimų srityje (pvz., perkeliant į saugyklą UCI).

3. Praktinė darbo vertė įvertinta realizuotu įrankiu, t. y. neurologinių sutrikimų aptikimo mobiliąja programėle „Neural Impairment Test Suite“ (NITS), skirta paciento sveikatos būklei stebėti grafine forma ir įsivertinti bei palyginti su sveikų asmenų atliktų testų rezultatais. NITS nemokamai pasiekama Google Play portale. Disertacijos rašymo metu daugiau kaip 500 vartotojų jau yra įsidiegę NITS. Autorius pažymi, kad programėlė skirta paciento smalsumui patitikrinti ir nepakeičia medicinos specialisto profesionalių įvertinimų. Darbo tyrimuose naudojami tik KTU etikos leidimo patvirtinti duomenys, tiriamiesiems pasirašant informuoto sutikimo formą, leidžiančią surinktus duomenis naudoti mokslo tikslams. Kiti, nežinomi šaltiniai yra ignoruojami.
4. Darbe atlikti 4 eksperimentai, skirti suskirstyti tiriamuosius į sveikų ir sergančių asmenų grupes (sprendžiama binarinė klasifikavimo problema). Sukurtas duomenų rinkinys buvo naudojamas eksperimentams validuoti. Pirmame (E1) ir antrame (E2) eksperimentuose nagrinėjami tremorų, kognityvinius ir energijos sąnaudų sutrikimus įvertinantys požymiai. E1 eksperimente taikomas inovatyvus požymių paskirstymo metodas, t. y. kiekvienai metodinei priemonei yra atskirai sugrupuojami požymiai. Be to, klasifikavimo modeliams buvo sukurti trys papildomi unikalių savybių rinkiniai: tinkama rašyba, *SAGE* taškai ir užduoties trukmė sekundėmis (kombinacijos). E1 aukščiausi pasiekti rezultatai pastebėti vertinant T9 užduotį (lokalaus svorinio mokymo klasifikatorius 91,50 proc.), o žemiausi T11 (atsitiktinių miškų klasifikatorius = 74,52 proc.) ir rašybos (logistiniai medžiai su atsitiktiniu mišku = 74,52 proc.) užduotis. E2 eksperimente, suintegravus visus 238 požymius (kaip apibrėžta išplėstinėje metodologijoje) ir pritaikius adaptyviuosius spartinimo algoritmus (AdaBoostM1) su atsitiktinio miško klasifikatoriumi, pasiekiamas 94,57 proc. tikslumas, t. y. 3 proc. didesnis, palyginti su geriausiaisiais E1 modelių rezultatais. Visi E2 eksperimente suprojektuoti klasifikatorių tikslumo rezultatai nėra žemesni nei 91,47 proc. E3 ir E4 eksperimentai nagrinėja pacientų kalbos sutrikimus. E3 balso signalams pritaikius gilaus mokymo dvikryptį rekurentinį

neuroninį tinklą su ilga trumpalaikė atmintimi, gaunamas 86,63 proc. klasifikavimo modelio tikslumas testavimo duomenų rinkiniui ir 100 proc. tikslumas treniravimo rinkiniui. E4 pritaikomas išsibarsčiusių (Gabor) bangelių transformacijų metodas ir pasiekiamas 100 proc. tikslumas testavimo rinkiniui. Bangelių metodas pranašesnis už gilaus mokymo rekurentinį neuroninį tinklą, įvertinant faktą, kad sukauptų balso įrašų trukmė buvo iki 47 sekundžių.

5. Sujungiant 3 klasifikatorius – adaptyvų spartinimo algoritmą AdaBoostM1 (minimaliai nuoseklus optimizavimas), AdaBoostM1 (Bajeso tinklas) ir balsavimo perceptroną (principinė komponentų analizė) su pritaikytu daugumos balsavimo metodu (tikimybių vidurkių komponavimo algoritmas), gaunamas hibridinis modelis su ~ 2 proc. didesniu tikslumu, t. y. 96,12 proc., ir geresniais modelio statistiniais įverčiais (mažesniu klaidų skaičiumi) bei mažiausiomis modelio mokymo laiko sąnaudomis (0,59 sekundės). Hibridizacijos metodas nagrinėja tremorų (motorikos), kognityvinius ir energijos sąnaudų kombinuotus požymių rinkinius. E3 ir E4 yra atskiri eksperimentai, kuriems reikalinga specializuota programinė įranga balso signalams apdoroti ir papildomi kompiuteriniai skaičiavimo ištekliai.

5.2. Ateities darbai

Realizuoti modeliai apima labai įvairius sutrikimus, kurie gali pasireikšti pacientams, sergantiems Hantingtono, Parkinsono ligomis, ankstyva demencija: vieniems dreba rankos ar kūnas, kitiems sutrinka kognityvinis aparatas, taip pat atsiranda balso problemų ar krenta kūno svoris. Toks faktorius suteikia didelę paciento, kuriam pasireiškia neurologiniai sutrikimai, pablogėjusios sveikatos būklės nustatymo galimybę. Eksperimentų (E1, E2, E3 ir E4) pagrindu sukurti teoriniai modeliai gali būti integruojami į NITS programėlę išpėjimams (aliarmams) apie ligos progresavimą prognozuoti iki tol, kol simptomai pasimato vizualiai.

Padėka

Autorius norėtų padėkoti Lietuvos Hantingtono ligos asociacijos prezidentei dr. Živilei Navikienei už pagalbą kontaktuojant su sergančiais asmenimis ir jų šeimos nariais, taip pat už vykdomų tyrimų sklaidą feisbuke ir asociacijos puslapyje bei praktinius patarimus.

Žmogiškieji ištekliai

Disertacijoje aprašytus tyrimus patvirtino Kauno technologijos universiteto Informatikos fakulteto taryba.

6. Literatūros sąrašas

1. F. Patou, M. Dimaki, W. E. Svendsen, K. Kjaegaard, and J. Madsen, "A smart mobile lab-on-chip-based medical diagnostics system architecture designed for evolvability," in 2015 Euromicro Conference on Digital, System Design. IEEE, aug 2015.
2. C. Doukas, T. Pliakas, and I. Maglogiannis, "Mobile healthcare information management utilizing cloud computing and android os," in Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE, 2010, pp. 1037–1040.
3. W.-T. Sung and Y.-C. Chiang, "Improved particle swarm optimization algorithm for android medical care iot using modified parameters," *Journal of medical systems*, vol. 36, no. 6, pp. 3755–3763, 2012.
4. A. Kailas, C.-C. Chong, and F. Watanabe, "From mobile phones to personal wellness dashboards," *IEEE pulse*, vol. 1, no. 1, pp. 57–63, 2010.
5. K. Huckvale, S. Adomaviciute, J. T. Prieto, M. K.-S. Leow, and J. Car, "Smartphone apps for calculating insulin dose: a systematic assessment," *BMC medicine*, vol. 13, no. 1, p. 106, 2015.
6. Y. E. Gelogo and H.-K. Kim, "Integration of wearable monitoring device and android smartphone apps for u-healthcare monitoring system," *International Journal of Software Engineering & Its Applications*, vol. 9, no. 4, 2015.
7. A. Nugroho, N. Nuryani, I. Yahya, A. D. Sutomo, B. Haijito, and A. Lestari, "Premature ventricular contraction detection using artificial neural network developed in android application," in Electric Vehicular Technology and Industrial, Mechanical, Electrical and Chemical Engineering (ICEVT & IMECE), 2015 Joint International Conference. IEEE, 2015, pp. 212–214.
8. A. Jutel and D. Lupton, "Digitizing diagnosis: a review of mobile applications in the diagnostic process," *Diagnosis*, vol. 2, no. 2, pp. 89–96, 2015.

9. E. Rovini, C. Maremmani, and F. Cavallo, “Automated systems based on wearable sensors for the management of parkinson’s disease at home: A systematic review,” *Telemedicine and e-Health*, jul 2018. [Online]. Available: <https://doi.org/10.1089/tmj.2018.0035>
10. G. K. Wenning, S. Kiechl, K. Seppi, J. Müller, B. Högl, M. Saletu, G. Rungger, A. Gasperi, J. Willeit, and W. Poewe, “Prevalence of movement disorders in men and women aged 50–89 years (brunec study cohort): a population-based study,” *The Lancet Neurology*, vol. 4, no. 12, pp. 815–820, dec 2005. [Online]. Available: [https://doi.org/10.1016/s1474-4422\(05\)70226-x](https://doi.org/10.1016/s1474-4422(05)70226-x).
11. Lauraitis, A. (2018). *Neural Impairment Test Suite*. Google Play. [Online]. https://play.google.com/store/apps/details?id=com.alauraitis.test_suite&hl=en
12. Amanda Hollenbeck (2019). *Smartphone App Can Help Detect Early Symptoms of Rare Disease*. In *Laboratory Equipment*. [Online]: <https://www.laboratoryequipment.com/article/2019/06/smartphone-app-can-help-detect-early-symptoms-rare-disease>.
13. Iqra Mumal (2019). *New Smartphone App Able to Recognize Early Symptoms of Huntington’s Disease*. In *Huntington’s Disease News*. [Online]: <https://huntingtonsdiseasenews.com/2019/05/28/new-smartphone-app-recognizes-early-huntingtons-symptoms/>
14. LRT (2019). *Rytdienos pranešimas, padėsiantis gyventi ilgiau: atpažinti retą ligą gali padėti ir mobilioji programėlė*. In *Lithuanian Television and Radio portal* [Online]: <https://www.lrt.lt/naujienos/mokslas-ir-it/11/1060638/rytdienos-pranesimas-padesiantis-gyventi-ilgiau-atpazinti-reta-liga-gali-padeti-ir-mobilioji-programele>
15. D. W. Scharre, S.-I. Chang, R. A. Murden, J. Lamb, D. Q. Beversdorf, M. Kataki, H. N. Nagaraja, and R. A. Bornstein, “Self-administered gerocognitive examination (SAGE),” *Alzheimer Disease & Associated Disorders*, vol. 24, no. 1, pp. 64–71, jan 2010. [Online]. Available: <https://doi.org/10.1097/wad.0b013e3181b03277>

16. Scharre, Douglas W., et al. "Digitally translated Self-Administered Gerocognitive Examination (eSAGE): relationship with its validated paper version, neuropsychological evaluations, and clinical assessments." *Alzheimer's research & therapy* 9.1 (2017): 44.
17. Nasreddine, Ziad S., et al. "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment." *Journal of the American Geriatrics Society* 53.4 (2005): 695-699.
18. Huisinigh, Carrie, et al. "Relationship between Areas of Cognitive Functioning on the Mini-Mental State Examination and Crash Risk." *Geriatrics* 3.1 (2018): 10.
19. Tsiouris, Kostas M., et al. "PD_Manager: an mHealth platform for Parkinson's disease patient management." *Healthcare technology letters* 4.3 (2017): 102-108.
20. Fumo, David (2017). Types of Machine Learning Algorithms You Should Know. In magazine *Towards Data Science Sharing concepts, ideas, and codes*. [Online]: <https://towardsdatascience.com/types-of-machine-learning-algorithms-you-should-know-953a08248861>.
21. Brownlee, Jason (2019). A Tour of The Most Popular Machine Learning Algorithms. In magazine *Machine Learning Algorithms*. Tour of The Most Popular Machine Learning Algorithms. [Online] : <https://machinelearningmastery.com/a-tour-of-machine-learning-algorithm>
22. I. Witten, E. Frank, L. Trigg, M. Hall, G. Holmes, and S. Cunningham. *Weka: Practical machine learning tools and techniques with java implementations*. 1999
23. Ian H. Witten and Eibe Frank. *Data mining: Practical machine learning tools and techniques*.
24. Florens, Jean-Pierre. *Elements of Bayesian statistics*. Routledge, 2019.
25. Gaba AM, Zhang K, Marder K, Moskowitz CB, Werner P, Boozer CN (2005). Energy balance in early-stage Huntington disease. *Am J Clin Nutr.* 2005;81(6):1335-41.

26. Marleen R. van Walsem et al. Assistive Technology for Cognition and Health-related Quality of Life in Huntington's Disease (2016). *J Huntingtons Dis.* 2016; 5(3): 261–270.
27. Winkler W. E. (1990). String Comparator Metrics and Enhanced Decision Rules in the Fellegi-Sunter Model of Record Linkage, Proceedings of the Section on Survey Research Methods, American Statistical Association, 354–359.
28. Nikolic, M. (2010). Measuring Similarity of Graphs and their Nodes by Neighbor Matching. *Computer Science, Artificial Intelligence.* arXiv:1009.5290. [Online]. Available: <https://arxiv.org/pdf/1009.5290.pdf>.
29. Math Planet (2019). Angles and parallel lines. [Online]. Available: <https://www.mathplanet.com/education/pre-algebra/introducing-geometry/angles-and-parallel-lines>
30. Drugman, Thomas, and Abeer Alwan. "Joint Robust Voicing Detection and Pitch Estimation Based on Residual Harmonics." Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. 2011, pp. 1973–1976.
31. Rabiner, Lawrence R., and Ronald W. Schafer. Theory and Applications of Digital Speech Processing. Upper Saddle River, NJ: Pearson, 2010.
32. Andén, J., and S. Mallat. "Deep Scattering Spectrum." IEEE Transactions on Signal Processing. Vol. 62, Number 16, 2014, pp. 4114–4128.
33. Mallat, S. "Group Invariant Scattering." Communications in Pure and Applied Mathematics. Vol. 65, Number 10, 2012, pp. 1331–1398
34. Misra, H., S. Iqbal, H. Bourlard, and H. Hermansky. "Spectral Entropy Based Feature for Robust ASR." 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing.
35. Amirkalali, B., Hosseini, S., Heshmat, R., Larijani B. (2008). Comparison of Harris Benedict and Mifflin-ST Jeor equations with indirect calorimetry in evaluating resting energy expenditure. *Indian J Med Sci.* 62(7):283-90. PMID: 18688113.

36. Weight Loss Tips - Diet, Nutrition, & Exercise Advice and Plans (2019). WeightLossForAll [Online]. Available: <http://www.weightlossforall.com/>
37. Metabolic Equivalent of Task Values for Activities in American Time Use Survey and 2002 Census Occupational Classification System (2019). *National Cancer Institute*.
38. Yasser EL-Manzalawy (2005). WLSVM. <http://www.cs.iastate.edu/~yasser/wlsvm/>
39. Chih-Chung Chang, Chih-Jen Lin (2001). LIBSVM - A Library for Support Vector Machines. URL <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
40. D. Aha, D. Kibler (1991). Instance-based learning algorithms. *Machine Learning*. 6:37-66.
41. Ware, M. 2000. WEKA Documentation. University of Waikoto.
42. J. Platt: Fast Training of Support Vector Machines using Sequential Minimal Optimization. In B. Schoelkopf and C. Burges and A. Smola, editors, *Advances in Kernel Methods - Support Vector Learning*, 1998.
43. S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, K.R.K. Murthy (2001). Improvements to Platt's SMO Algorithm for SVM Classifier Design. *Neural Computation*. 13(3):637-649.
44. Trevor Hastie, Robert Tibshirani: Classification by Pairwise Coupling. In: *Advances in Neural Information Processing Systems*, 1998.
45. Fisher, R. A. (1936). "The Use of Multiple Measurements in Taxonomic Problems". *Annals of Eugenics*. 7 (2): 179–188. doi:10.1111/j.1469-1809.1936.tb02137.x. hdl:2440/15227.
46. Mika, S; Rätsch, G.; Weston, J.; Schölkopf, B.; Müller, KR (1999). Fisher discriminant analysis with kernels. *Neural Networks for Signal Processing*. IX. pp. 41–48. CiteSeerX 10.1.1.35.9904. doi:10.1109/NNSP.1999.788121. ISBN 978-0-7803-5673-3
47. Deeplearning4j Development Team. (2017) Deeplearning4j: Open-source distributed deep learning for the JVM, Apache Software Foundation License 2.0. <http://deeplearning4j.org>

48. Leo Breiman (2001). Random Forests. *Machine Learning*. 45(1):5-32.
49. Ben Gal I (2007). "Bayesian Networks" (PDF). In Ruggeri F, Kennett RS, Faltin FW (eds.). *Encyclopedia of Statistics in Quality and Reliability*. John Wiley & Sons. doi:10.1002/9780470061572.eqr089. ISBN 978-0-470-01861-3.
50. George H. John, Pat Langley: Estimating Continuous Distributions in Bayesian Classifiers. In: *Eleventh Conference on Uncertainty in Artificial Intelligence*, San Mateo, 338-345, 1995.
51. Ross Quinlan (1993). *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, San Mateo, CA.
52. Bottou L. (2012) Stochastic Gradient Descent Tricks. In: Montavon G., Orr G.B., Müller KR. (eds) *Neural Networks: Tricks of the Trade*. *Lecture Notes in Computer Science*, vol 7700. Springer, Berlin, Heidelberg.
53. Niels Landwehr, Mark Hall, Eibe Frank (2005). Logistic Model Trees. *Machine Learning*. 95(1-2):161-205.
54. Marc Sumner, Eibe Frank, Mark Hall: Speeding up Logistic Model Tree Induction. In: *9th European Conference on Principles and Practice of Knowledge Discovery in Databases*, 675-683, 2005.
55. Iba, Wayne; and Langley, Pat (1992); Induction of One-Level Decision Trees, in *ML92: Proceedings of the Ninth International Conference on Machine Learning*, Aberdeen, Scotland, 1-3 July 1992, San Francisco, CA: Morgan Kaufmann, pp. 233-240.
56. Y. Freund, R. E. Schapire: Large margin classification using the perceptron algorithm. In: *11th Annual Conference on Computational Learning Theory*, New York, NY, 209-217, 1998.
57. le Cessie, S., van Houwelingen, J.C. (1992). Ridge Estimators in Logistic Regression. *Applied Statistics*. 41(1):191-201.
58. Ron Kohavi: The Power of Decision Tables. In: *8th European Conference on Machine Learning*, 174-189, 1995.
59. Eibe Frank, Mark Hall, Bernhard Pfahringer: Locally Weighted Naive Bayes. In: *19th Conference in Uncertainty in Artificial Intelligence*, 249-256, 2003.

60. C. Atkeson, A. Moore, S. Schaal (1996). Locally weighted learning. *AI Review*.
61. Yoav Freund, Robert E. Schapire: Experiments with a new boosting algorithm. In: Thirteenth International Conference on Machine Learning, San Francisco, 148-156, 1996.
62. David, S. K., Saeb, A. T., Rafiullah, M., & Rubeaan, K. (2019). Classification Techniques and Data Mining Tools Used in Medical Bioinformatics. In S. Strydom, & M. Strydom (Eds.), *Big Data Governance and Perspectives in Knowledge Management*(pp. 105-126). Hershey, PA: IGI Global. doi:10.4018/978-1-5225-7077-6.ch005.
63. Ludmila I. Kuncheva (2004). *Combining Pattern Classifiers: Methods and Algorithms*. John Wiley and Sons, Inc.
64. T. K. Ho. Multiple classifier combination: Lessons and the next steps. In A. Kandel and H. Bunke, editors, *Hybrid Methods in Pattern Recognition*. World Scientific Publishing, 2002, pp. 171– 19.
65. Dehzangi, Abdollah, Somnuk Phon-Amnuaisuk, and Omid Dehzangi. "Enhancing protein fold prediction accuracy by using ensemble of different classifiers." *Australian Journal of Intelligent Information Processing Systems* 26.4 (2010): 32-4.
66. Schuster, Mike, and Kuldip K. Paliwal. "Bidirectional recurrent neural networks." *Signal Processing, IEEE Transactions on* 45.11 (1997): 2673-2681.2. Awni Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan.
67. Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". *Neural Computation*. 9 (8): 1735–1780. doi:10.1162/neco.1997.9.8.1735. PMID 9377276.
68. Lee, Tai S. (October 1996). "Image Representation Using 2D Gabor wavelets" (PDF). *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 18 (10): 959–971. doi:10.1109/34.541406.
69. Tsiouris, Kostas M., et al. "PD_Manager: an mHealth platform for Parkinson's disease patient management." *Healthcare technology letters* 4.3 (2017): 102-108.

70. D. W. Scharre, S.-I. Chang, R. A. Murden, J. Lamb, D. Q. Beversdorf, M. Kataki, H. N. Nagaraja, and R. A. Bornstein, "Self-administered gerocognitive examination (SAGE)," *Alzheimer Disease & Associated Disorders*, vol. 24, no. 1, pp. 64–71, jan 2010.
71. Scharre, Douglas W., et al. "Digitally translated Self-Administered Gerocognitive Examination (eSAGE): relationship with its validated paper version, neuropsychological evaluations, and clinical assessments." *Alzheimer's research & therapy* 9.1 (2017): 44.
72. Nasreddine, Ziad S., et al. "The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment." *Journal of the American Geriatrics Society* 53.4 (2005): 695-699.
73. Huisinigh, Carrie, et al. "Relationship between Areas of Cognitive Functioning on the Mini-Mental State Examination and Crash Risk." *Geriatrics* 3.1 (2018): 10.
74. Bennisar, M. et al. (2016). Huntington's Disease Assessment Using Tri Axis Accelerometers. In *Procedia Computer Science*, vol. 96, pages 1193-1201. <https://doi.org/10.1016/j.procs.2016.08.163>.
75. Wiecki, T. M. et al. (2016). A Computational Cognitive Biomarker for Early-Stage Huntington's Disease. *PLOS journal*. [Online]. <https://doi.org/10.1371/journal.pone.0148409>.
76. Lietuvos Vilniaus Santariškių klinikų Hantingtono ligos koordinacijos centras. Prieiga internete:
http://www.santa.lt/index.php?option=com_content&view=article&id=2041&catid=178&Itemid=129, 2015.
77. Engin, M., Demirag, S., Engin, Z., E., Celebi, G., Ersan, F., Asena, F., Colakoglu, Z. (2007). The classification of human tremor signals using artificial neural network (2007). *Expert Systems with Applications*, 33(3), pp. 754-761.
78. Wu D., Warwick K., Ma Z., Gasson M.N., Burgess J.G., Pan S., Aziz T.Z. (2010). Prediction of Parkinson's disease tremor onset using a radial basis function neural network based on particle swarm optimization. *Int. J. Neur. Syst.*, 20(02), 109.

79. Cole B.T., Roy S.H., De Luca C.J., Nawab S. H. (2010). Dynamic neural network detection of tremor and dyskinesia from wearable sensor data. In: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, 2010:6062-5.
80. Chandrashekhar A., Jain S., Kumar Jha V. (2014). Design and Analysis of Data Mining Based Prediction Model for Parkinson's disease. International Journal of Computer Science Engineering (IJCSE), Vol. 3 No.03.
81. Geman O. (2013). Parkinson's disease Assessment using Fuzzy Expert System and Nonlinear Dynamic. In Advances in Electrical and Computer Engineering, 13(1), pp. 41-46.
82. Obi J.C., Imainvan A.A. (2011). Decision Support System for the Intelligent Identification of Alzheimer using Neuro Fuzzy logic. International Journal on Soft Computing (IJSC), Vol.2, No.2, DOI: 10.5121/ijsc.2011.2203.
83. Iram S., Fergus P., Al-Jumeily D., Hussain A., Randles M. (2015). A classifier fusion strategy to improve the early detection of neurodegenerative diseases. International Journal of Artificial Intelligence and Soft Computing (IJASIS), 5(1), 23-44.
84. Geman O. (2014). Towards an inclusive Parkinson's screening system. In 18th International Conference on System Theory, Control and Computing (ICSTCC), 470-475.
85. Ramani, G.R, Sivagami, G. (2011). Parkinson Disease Classification using Data Mining Algorithms. International Journal of Computer Applications 32(9):17-22.
86. Rustempasic I., Can M. (2013). Diagnosis of Parkinson's disease using Fuzzy C-Means Clustering and Pattern Recognition. In SouthEast Europe Journal of Soft Computing, 2, 42-49.

87. Morimoto, Chie, et al. "Temporal processing instability with millisecond accuracy is a cardinal feature of sensorimotor impairments in autism spectrum disorder: analysis using the synchronized finger-tapping task." *Journal of autism and developmental disorders* 48.2 (2018): 351-360.
88. Payne, Joshua S., et al. "Study protocol for a randomized pilot study of a computer-based, non-pharmacological cognitive intervention for motor slowing and motor fatigue in Parkinson's disease." *Pilot and feasibility studies* 4.1 (2018): 190
89. Binder, Julius. Implementation and evaluation of two distinct electronic motion detection devices for the assessment of abnormal movements in Huntington's disease. Diss. Universität Ulm, 2019.
90. Đurić-Jovičić, Milica, et al. "Finger and foot tapping sensor system for objective motor assessment." *Vojnosanitetski pregled* 75.1 (2018): 68-77.
91. Suzumura, Shota, et al. "Differences among patients with Alzheimer's disease, older adults with mild cognitive impairment and healthy older adults in finger dexterity." *Geriatrics & gerontology international* 18.6 (2018): 907-914.
92. Pham, Hung N., et al. "Multimodal Detection of Parkinson Disease based on Vocal and Improved Spiral Test." 2019 International Conference on System Science and Engineering (ICSSE). IEEE, 2019.
93. Pereira, Clayton R., et al. "Handwritten dynamics assessment through convolutional neural networks: An application to Parkinson's disease identification." *Artificial intelligence in medicine* 87 (2018): 67-77.
94. Danna, Jérémy, et al. "Digitalized spiral drawing in Parkinson's disease: A tool for evaluating beyond the written trace." *Human movement science* 65 (2019): 80-88.
95. Bernardo, Lucas S., et al. "Handwritten pattern recognition for early Parkinson's disease diagnosis." *Pattern Recognition Letters* 125 (2019): 78-84.

96. Moetesum, Momina, et al. "Assessing visual attributes of handwriting for prediction of neurological disorders—A case study on Parkinson's disease." *Pattern Recognition Letters* 121 (2019): 19-27.
97. Lunven, Marine, et al. "F32 Cognitive decline in Huntington's disease (HD) in computerized arithmetic task." (2018): A51-A51.
98. Alloni, Anna, et al. "Evaluation of an ontology-based system for computerized cognitive rehabilitation." *International journal of medical informatics* 115 (2018): 64-72.
99. Flannery, Samuel L., et al. "Computerized testing in Parkinson's disease: Performance deficits in relation to standard clinical measures." *Journal of clinical and experimental neuropsychology*. 40.10 (2018): 1062-1073.
100. Wu, Jia-Yun, et al. "Computerized Cognitive Assessment System for Dementia Screening Application." *World Congress on Medical Physics and Biomedical Engineering 2018*. Springer, Singapore, 2019.
101. Betrouni, Nacim, et al. "Electroencephalography-based machine learning for cognitive profiling in Parkinson's disease: Preliminary results." *Movement Disorders*, 34.2 (2019): 210-217.
102. Tsanas A, Little MA, Fox C, Ramig LO (2014). Objective Automatic Assessment of Rehabilitative Speech Treatment in Parkinson's Disease. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*. 2014;22(1):181–190. doi: 10.1109/TNSRE.2013.2293575.
103. Hauptman, Yermiyahu et al. Identifying distinctive acoustic and spectral features in Parkinson's disease. *INTERSPEECH* 2019.
104. Wu, Kebin, et al. "Learning acoustic features to detect Parkinson's disease." *Neurocomputing* 318 (2018): 102-108.
105. Perez, Matthew, et al. "Classification of Huntington Disease Using Acoustic and Lexical Features." *Interspeech*. 2018.
106. Oung, Qi Wei, et al. "Empirical Wavelet Transform Based Features for Classification of Parkinson's Disease Severity." *Journal of medical systems* 42.2 (2018): 29.

107. Cersosimo, Maria G., et al. "Weight Loss in Parkinson's Disease: The Relationship with Motor Symptoms and Disease Progression." *BioMed research international* 2018 (2018).
108. Singh, Balbir, and Hissam Tawfik. "A Machine Learning Approach for Predicting Weight Gain Risks in Young Adults." *2019 10th International Conference on Dependable Systems, Services and Technologies (DESSERT)*. IEEE, 2019.
109. Villasana, María Vanessa, et al. "CoviHealth: Novel approach of a mobile application for nutrition and physical activity management for teenagers." *Proceedings of the 5th EAI International Conference on Smart Objects and Technologies for Social Good*. ACM, 2019.
110. Tirasirichai, Benjarat, et al. "Bloom Balance: Calorie Balancing Application With Scientific Validation." *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE, 2018.
111. Merchant, Kaiz, and Yash Pande. "ConvFood: A CNN-Based Food Recognition Mobile Application for Obese and Diabetic Patients." *Emerging Research in Computing, Information, Communication and Applications*. Springer, Singapore, 2019. 493-502.

7. Darbo sklaida

1. Video reportažas (interviu) LNK televizijai. Mobilios programėlės pristatymas (2019-05). Nurodyta KTU Informatikos fakulteto Facebook paskyroje: <https://www.facebook.com/ktuif/posts/10157213111908194/>.
2. Dalyvavimas ir filmavimasis žinių radijo laidoje "Mokslo Mygtukas" (2019-05). [Internete]: <https://www.ziniuradijas.lt/laidos/mokslo-mygtukas/tai-pades-nustatyti-itin-reta-liga-jau-ankstyvoje-stadijoje?video=1>
3. Sklaida Lietuvos spaudoje. Kanalas 1. [Internete]: <http://www.elektronika.lt/produktai/programos/68153/atpazinti-reta-liga-nuo-siol-gali-padeti-ir-mobilioji-programele>. Data: 2019 m. gegužės mėn..
4. Sklaida Lietuvos spaudoje. Kanalas 2. [Internete]: <https://www.delfi.lt/mokslas/technologijos/atpazinti-sunkia-liga-nuo-siol-gali-padeti-ir-mobilioji-programele.d?id=81222371>. Data: 2019 m. gegužės mėn..
5. Sklaida Lietuvos spaudoje. Kanalas 3. [Internete]: <https://www.etaplus.lt/atpazinti-reta-liga-nuo-siol-gali-padeti-ir-mobilioji-programele>. Data: 2019 m. gegužės mėn..
6. Sklaida Lietuvos spaudoje. Kanalas 4. Data: 2019 m. gegužės mėn.. [Internete]: <https://www.lrt.lt/naujienos/mokslas-ir-it/11/1060638/rytdienos-pranesimas-padesiantis-gyventi-ilgiau-atpazinti-reta-liga-gali-padeti-ir-mobilioji-programele>
7. Sklaida Lietuvos spaudoje. Kanalas 5. [Internete]: <https://kauno.diena.lt/naujienos/ivairenybes/mokslas-ir-it/atpazinti-reta-liga-nuo-siol-gali-padeti-ir-mobilioji-programele-914677>. Data: 2019 m. gegužės mėn..
8. Sklaida Lietuvos spaudoje. Kanalas 6. Data: 2019 m. gegužės mėn.. [Internete]: <https://klaipeda.diena.lt/naujienos/ivairenybes/mokslas-ir-it/atpazinti-reta-liga-nuo-siol-gali-padeti-ir-mobilioji-programele-914677>

9. Sklaida Lietuvos spaudoje. Kanalas 7. [Internete]:
<https://www.diena.lt/naujienos/ivairenybes/mokslas-ir-it/atpazinti-reta-liga-nuo-siol-gali-padeti-ir-mobilioji-programele-914677>. Data: 2019 m. gegužės mėn..
10. Sklaida pasaulinėje spaudoje. Kanalas 1. Data: 2019 m. birželio mėn. [Internete]:
<https://www.laboratoryequipment.com/article/2019/06/smartphone-app-can-help-detect-early-symptoms-rare-disease>
11. Sklaida pasaulinėje spaudoje.. Kanalas 2. Data: 2019 m. gegužės mėn. [Internete]: https://www.eurekalert.org/pub_releases/2019-05/kuot-sac052119.php
12. Sklaida pasaulinėje spaudoje. Kanalas 3. Data: 2019 m. gegužės mėn. [Internete]: https://www.business-standard.com/article/pti-stories/new-smartphone-app-may-help-diagnose-rare-genetic-disease-119052200565_1.html
13. Sklaida pasaulinėje spaudoje. Kanalas 4. Data: 2019 m. gegužės mėn. [Internete]: <https://globalgenes.org/2019/05/21/smartphone-app-can-help-diagnose-huntingtons-disease/>
14. Sklaida pasaulinėje spaudoje. Kanalas 5. Data: 2019 m. gegužės mėn. [Internete]: <https://medicalxpress.com/news/2019-05-smartphone-app-rare-disease.html>.

8. Informacija apie autorių

Asmeninė informacija

Andrius Lauraitis

Gimimo vieta: Kaunas, Lietuva

Gimimo data: 1986 m. sausio 20 d.

El. pašto adresas: andrius.lauraitis@ktu.lt

Institucija

Kauno technologijos universitetas

Informatikos fakultetas

Multimedijos inžinerijos katedra

K. Baršausko g. 59, A336, LT-51423, Kaunas, Lietuva

Išsilavinimas

2015-2019	Informatikos mokslo krypties doktorantūros studijos Kauno technologijos universiteto, Informatikos fakulteto Multimedijos inžinerijos katedroje, Kaunas, Lietuva.
2009-2011	Igytas informatikos inžinerijos magistro kvalifikacinis laipsnis (informacinių sistemų inžinerijos studijų programa) Kauno technologijos universiteto, Informatikos fakulteto Informatikos sistemų katedroje, Kaunas, Lietuva.
2005-2009	Igytas informatikos bakalauro kvalifikacinis laipsnis (specializacija – informacinių sistemų projektavimas) Kauno technologijos universiteto, Informatikos fakulteto Informatikos sistemų katedroje, Kaunas, Lietuva.

Moksliniai interesai

Žmogaus ir kompiuterio sąveika (angl. *Human Computer Interaction*),

Požymių atrinkimo metodai (angl. *Feature Extraction*),

Duomenų tyryba (angl. *Data Mining*),

Mašininis mokymasis (angl. *Machine Learning*),

Android mobiliosios programėlės.

Darbo patirtis

2009-2010	Informacinių technologijų mokytojas, Kauno „Santaros“ gimnazija, 8-12 klasių moksleiviams.
2011	Informacinių technologijų mokytojas, Kauno „Santaros“ gimnazija, 9-12 klasių moksleiviams.
2010-2012	Kompiuterinių sistemų analitikas „UAB HLTECH“. Veiklos: Duomenų analizavimas ir semantizavimas, ontologijų kūrimas, užduočių ruošimas ir dokumentavimas; darbas su reliacinėmis duomenų bazėmis.
2012-2013	Projektinių veiklų analitikas KTU E. mokymosi technologijų centre.
2013	Jaunesnysis mokslo darbuotojas KTU E. mokymosi technologijų centre. Veiklos: publikacijų rašymas e.mokymosi tematika, projektinių veiklų vykdymas.
2014-2015	Jaunesnysis mokslo darbuotojas KTU Informatikos fakulteto Multimedijos inžinerijos katedroje. Veiklos: užsakomųjų (mokslo) darbų vykdymas, algoritmų projektavimas ir kūrimas.
2015-2019	Asistentas KTU Informatikos fakulteto Multimedijos inžinerijos katedroje. Veiklos: objektinio programavimo laboratorinių darbų vedimas.
2019-dabar	Lektorius KTU Informatikos fakulteto Multimedijos inžinerijos katedroje. Veiklos: objektinio programavimo laboratorinių darbų vedimas, modulio metodinės medžiagos atnaujinimas, kontrolinių ir egzamino užduočių ruošimas bei vertinimas (komisijos narys). ICIST konferencijų publikacijų recenzavimas. Vadovavimas bakalauro baigiamiesiems darbas.

9. Summary

9.1. Introduction

9.1.1. Relevance of the research

The computational intelligence provides many models with devoted applications for the classification of health disorders. Some solutions are developed for medical image processing, while others use the data from the clinical examinations to predict the state of health. With the development of technology, new forms of smart diagnostic systems such as for in-vitro medical testing [1] are introduced. Nowadays, the mobile smart devices are widely used around the globe. The applications on smartphones make use of multimedia capabilities, which can provide an extensive information on various features of human bodies. In order to compose a smart system that can actively participate in medical examinations and provide clinical decision support or simply support users in a daily routine, the classification models need to be able to use data collected in the process of automated medical examinations.

The mobile health applications are on the rise, with many people utilizing apps successfully in diverse ranges of medical practices. The apps range from medical information systems [2], physiological health care systems [3], health alert/identification systems and connection with healthcare providers [4] to calculating required doses of medicine, such as insulin [5] and often in combination with some e-health wearable devices [6], sometimes with a very specific application in detecting premature ventricular contraction [7]. The diagnostic apps have a huge potential to provide access to diagnostic definitions and might appeal to both healthcare professionals, as it can invite a patient for early diagnosis, and patients by offering a diagnostic adjunct [8]. The use of such systems can improve the treatment of the Huntington Disease (HD), support medical doctors, caregivers in remote monitoring of the HD patients [9].

The central nervous system motor disorders (CNSD) are mainly related to problems that people may have with the control of their bodies and mind actions. The number of people suffering from pathological tremor has been increasing steadily with the ageing of the population. As for numbers, 4% of the elderly (>65 years) suffer from the essential tremor (ET); whereas, 1% of people aged more than 50 years suffer from the Parkinson's disease (PD) [10]. The Alzheimer's Disease (AD) is a

neurodegenerative disease, which symptoms are memory loss and behaviour problems that gradually worsen in time to language difficulties, disorientation and behavioural problems with managing self-care, what finally lead to dementia. The Huntington Disease (HD) is mainly inherited from the family and reveals itself in the death of brain cells. In the early stage, people have subtle problems with mental abilities, lack of coordination or unsteady gait.

At the later stages, patients have uncoordinated moves that gradually worsen to serious difficulties in talk and dementia. Mild Cognitive Impairment (MCI) is a condition in which an individual's thinking ability shows some mild changes that can be easily noticed by the people who are close to the affected person. Some cases of MCI are actually considered as an early stage of AD. Therefore, it is an interest of researchers to detect these diseases in the early stage before they progress further and make medical treatments ineffective.

9.1.2. Aim of the research

To develop a computerized model, which predicts and identifies the deficits for patients at the early stage of central nervous system disorders (CNSD).

9.1.3. Tasks of the research

Five tasks were set in order to achieve the aim:

1. Analyse and compare the state of the art data mining and machine learning approaches that involve patients with central nervous system disorders (CNSD) in the classification problematics studies;
2. Create a computerized extended version of data collection methodology via smart interface (mobile phone and tablet), adapted for the early stage patients with CNSD;
3. Design a system for feature extraction mechanism based on cognitive, speech, energy expenditure impairments occurring in CNSD patients and collect features to a dataset.
4. Develop binary classification models based on the proposed methodological tasks to evaluate sick vs. healthy condition of CNSD patients;

5. Fuse different types (functions, Bayesian, trees, rules, lazy) of single supervised learning classifiers (hybrid model) that improve sick vs. healthy classification accuracy results and decreases performance errors.

9.1.4. Scientific novelty

These novelty factors are listed:

1. Computerization of predefined Self-Administered Gerocognitive exam (SAGE) methodology for the detection of early signs in memory or thinking cognitive impairments to evaluate patient's health state automatically with the proposed smart interface (mobile application).
2. Extension of SAGE methodology with extra tests for tremor, speech and energy expenditure impairments.
3. Innovative patient health state monitoring method for self-assessment and comparison to the measurements of healthy subjects.
4. Implementation of hybrid classification model (all tasks are considered from the proposed extended methodology) with improved accuracy (compared to standalone models) to predict possible deterioration of CNSD patient health status.

9.1.5. Research object

Data mining techniques for pattern recognition in showing healthy vs. sick (person suffering from neurological disorders such as Huntington Disease) classification intuition and machine learning models for the prediction of future outcomes.

9.1.6. Practical significance

The developed mobile application “Neural Impairment Test Suite” (NITS) [11] for CNSD patient health state monitoring is published in Google Play free of charge. The app is disseminated through various media channels in Lithuania and worldwide [12, 13, 14], presented to the medical community. At the time of writing this dissertation, more than 500 users are currently using NITS app.

9.1.7. Dissertation statements

1. Proposed extended SAGE methodology enables innovative health state monitoring for people, suffering from central nervous system disorders (e.g., Huntington, Parkinson Diseases, Mild-Cognitive Impairment, cerebral palsy), based on tremor, cognitive, speech and energy expenditure impairments.
2. Created NITS system can be adapted for CNSD patients as an assistive device for self-assessment (e.g., at home or medical institution) and give feedback on the early diagnosis of the disease status.
3. Implemented classifier fusion mechanism results in more accurate and improved performance (less prone to errors) hybrid model for solving binary (sick vs. healthy) classification problem.

9.1.8. Scientific approval

Articles indexed in the Web of Science and Scopus with Impact Factor/SNIP International Publishers

1. Lauraitis, A., Maskeliūnas, R., Damaševičius, R., Polap, D., Wozniak, M. (2019). A smartphone application for automated decision support in cognitive task based evaluation of central nervous system motor disorders. *IEEE Journal of Biomedical and Health Informatics*. Vol. 23, issue 5. Electronic ISSN: 2168-2208. DOI 10.1109/JBHI.2019.2891729. Author contribution: 0.400.
2. Lauraitis, Andrius; Maskeliūnas, Rytis; Damaševičius, Robertas. ANN and fuzzy logic based model to evaluate Huntington disease symptoms // *Journal of Healthcare Engineering*. New York: Hindawi. ISSN 2040-2295. eISSN 2040-2309. 2018, vol. 2018, art. no. 4581272, pp. 1–10. DOI: 10.1155/2018/4581272. [DOAJ; Scopus; MEDLINE; Science Citation Index Expanded (Web of Science)]. Author contribution: 0.334.
3. Ivanavičius, Arnas; Simonavičius, Henrikas; Gelšvartas, Julius; Lauraitis, Andrius; Maskeliūnas, Rytis; Cimmperman, Piotras; Serafinavičius, Paulius. Real-time CUDA-based stereo matching using Cyclops2 algorithm: Research // *EURASIP Journal on*

Image and Video Processing. Cham: Springer Open. ISSN 1687-5176. eISSN 1687-5281. 2018, vol. 2018, art. no. 12, pp. 1–15. DOI: 10.1186/s13640-018-0253-2. [DOAJ; Scopus; Science Citation Index Expanded (Web of Science)]. Author contribution: 0.142.

Articles indexed in the Web of Science and Scopus without Impact Factor/SNIP International Publishers

1. Lauraitis, Andrius; Maskeliūnas, Rytis. Investigation of predicting functional capacity level for Huntington disease patients // Information and software technologies: 23rd international conference, ICIST 2017, Druskininkai, Lithuania, October 12–14, 2017: proceedings / Robertas Damaševičius, Vilma Mikašytė (Eds.). Cham: Springer, 2017. ISBN 9783319676418. eISBN 9783319676425. pp. 142–149. (Communications in computer and information science, ISSN 1865-0929, eISSN 1865-0937; vol. 756). DOI: 10.1007/978-3-319-67642-5_12. Author contribution: 0.500.

Articles in other peer reviewed research publications (Articles in periodicals, serials, collections of articles etc.) International Publishers

1. Gelšvartas, Julius; Simonavičius, Henrikas; Lauraitis, Andrius; Maskeliūnas, Rytis; Cimmperman, Piotras; Serafinavičius, Paulius. Realtime stereo matching using maximally stable extremal regions // Transylvanian review. Author contribution: 0.166.

Articles in conference proceedings. National Publishers

1. Gelšvartas, Julius; Lauraitis, Andrius; Simutis, Rimvydas; Maskeliūnas, Rytis. Review of assistive technologies for disabled people // Biomedical engineering 2016 : Proceedings of international conference / Kaunas University of Technology, Lithuanian Society for Biomedical Engineering. Kaunas: Kauno technologijos universitetas. ISSN 2029-3380. 2016, pp. 142–146. Author contribution: 0.250.

PRESENTATION OF RESEARCH RESULTS AT CONFERENCES (Conference abstracts and non-peer-reviewed conference papers).

Other

abstracts and papers

1. Maskeliūnas, Rytis; Lauraitis, Andrius. Neural impairment test suite mobile app // Technorama 2019: From vision to innovation! : [9 May 2019]: innovation catalogue. [Kaunas]: KTU. 2019, project no. 49, p. 59. Author contribution: 0.250.

In 2018, the author participated in the summer school of Tallinn University and completed the course “EXPERIMENTAL INTERACTION DESIGN”, 4 ECTS. The author won R1 price in Technorama 2019 conference.

9.1.9. Dissertation structure

The dissertation is outlined to 5 chapters. Chapter 2 provides an overview of central nervous system disorders symptoms, including those occurring in Huntington Disease as well as related work of state of art data mining and machine learning. Chapter 3 covers materials and methods that are used in the dissertation, i.e., the design of computerized extended SAGE (Self-administered cognitive testing) methodology, methods for feature extraction and analysis of mathematical models for the investigated classifiers. Chapter 4 describes four conducted experiments to solve sick vs. healthy classification problem. Experiment one (E1) applies the feature distribution approach for individual tasks. Experiment two (E2) uses integrated feature set for building single unified classifiers and implies a fusion mechanism for developing a hybrid model. Experiment three (E3) and experiment four (E4) deal with voice signals to track speech impairments. Chapter 5 contains a discussion and conclusion. Chapter 6 provides a list of literature references.

9.2. Literature Review

A lot of research in the computer science field is being conducted that involves studies with central nervous system disorder (CNSD) patients. Mostly, such investigations include Huntington Disease (HD), Parkinson Disease (PD), Alzheimer Disease (AD), mild cognitive impairment (MCI) and dementia test subjects. These diseases cover a broad range of symptoms that are occurring in CNSD patients. In particular, tremor (involuntary movements, body balance disorders, muscle stagnancy etc.), cognitive (decision-making difficulties, behavioural distractions, problems to focus attention, memory loss, poor self-care etc.), speech (lack of words pronounced, the use of shorter phrases, pauses) and energy expenditure (weight loss, mainly due to the development of early negative energy balance) impairments.

The computer scientists aim at the development of diagnostic aided tools for neural impairment screening and models for decision support making [19], [11]. In many cases, such aims to be digitalized are based on a set of methodological tasks, approved and validated by medical staff from clinical institutions. Most commonly known methodologies are Montreal Cognitive Assessment (MoCA) [17], Mini-Mental State Examination (MMSE) [18], Self-Administered Gerocognitive Examination (SAGE) [15] and its electronic version (e-SAGE) [16]. MoCA was developed to target cognitive impairments, which are best adapted for a screening test, in MCI and AD patients. MoCA covers short-term memory recall tasks, visuospatial abilities, phonemic fluency, repetition, orientation, trial making (executive), target detection, language assessment. MMSE test includes the orientation in time and place, repeating named prompts, spelling a given word backwards, speaking a phrase or drawing a shown figure. Similarly, SAGE methodology covers general insights task, orientation (current date), picture naming, similarities, basic calculations (mathematical division and subtraction), construction (3-D figure, analogue clock), language (verbal fluency), executive (modified trials, problem solving) and memory. All three methodologies have a scoring (task evaluation) system, i.e., if the total score of a patient is less than a predefined threshold value, an impaired neuropsychological health state is assessed. The decision support making in the presented context is a process of integrating classification models to the proposed screening tools.

The literature analysis chapter is structured as follows. Firstly, an overview of Huntington Disease (HD) symptoms is provided. The reason for considering HD is that it covers all four types of targeted impairments. In addition, HD symptoms are most versatile, as compared to other neurodegenerative disorders. Moreover, the majority of CNSD test subjects that were involved in the work of this dissertation are HD patients. Finally, the analysis summary is written to emphasise the drawbacks of existing solutions and give motivation for the proposed approach.

9.3. Materials and methods

A dataset collection tool is proposed based on the self-administered cognitive testing methodology (SAGE), which is used to identify the signs of mild cognitive impairment (MCI) and early dementia in, e.g., Huntington disease (HD), Parkinson disease (PD), Alzheimer disease (AD) patients. SAGE [15], [16] is applied in practice (mostly in the USA) by medical practitioners-neurologists, i.e., submitting questionnaires to patients in paper form and self-assessing their condition manually.

The author presents an innovative approach for computerizing all 12 tasks that are defined in SAGE questionnaire with the integration of 4 different group variants as well as extending the methodology by considering related research about tremor (motor) and energy expenditure impairments that are occurring in HD, PD and AD patients. All these indicators lead to the improved patient health state monitoring by using automatic manner, i.e., without direct supervision of the doctor. The proposed system is easily adaptable to various neurodegenerative diseases (HD, PD, AD) or dementia conditions (MCI).

In addition, the adapted methods for feature extraction in tremor, cognitive and energy expenditure impairments are Euclidean Distance, Frechet distance for curve comparison, Jaro algorithm to compare string input symbol by symbol, Neighbour matching algorithm for graph similarity measure, parallel line angle calculation of 3-D figure edges, finger motion tracking methods, daily calories gained and burned calculation by using Basal Metabolic Rate (BMR). The feature extraction methods for speech analysis are the audio processing and extraction methods: pitch, Mel-frequency cepstral coefficients (MFCC), gammatone cepstral coefficients (GTCC), Wavelet Scattering Transform (WST),

spectral methods in frequency domain (slope, skewness, spread, flux, centroid, rolloff, decrease, flatness, kurtosis, entropy).

9.4. Experimental Research

In this dissertation, a fusion strategy is provided, i.e., all extracted materials (features) are considered in healthy vs. sick (two target classes, 0 vs. 1) test subject evaluation. Thus, 4 experiments (all of them are adapted on the built-in dataset) were carried out:

1. Experiment1 (E1): feature set is distributed to individual tasks only (14 different classifiers);
2. Experiment2 (E2): all 238 features are combined (integrated) and fed to a classifier, combinations of classifiers are used to propose a hybrid model;
3. Experiment3 (E3): audio files (from T14 task) are used for further processing to extract features with a combination of methods (Pitch, MFCC, GTCC and spectral, e.g., skewness) and classify samples with deep learning networks (in particular recurrent neural network (RNN) with Long short-term memory (LSTM));
4. Experiment4 (E4): audio files (from T14 task) are used for further processing to extract features with WST method and classify samples with support vector machine (SVM).

For decision support making of health state evaluation, the author targets a wide range of existing machine learning (ML) methods. Four main types of ML algorithms can be defined for dealing with classification (categorization of certain instances to types) problems [20], [21]: supervised (input data has a known label), semi-supervised (not all data instances have labels), unsupervised (input data has no label) and reinforcement learning (labels are gathered from the environment in an iterative fashion). Three latter approaches are excluded, i.e., leaving with supervised learning, as labelled data (0 or 1) was collected from patients with direct supervision for test execution.

In addition, the addressed classification problematics in this dissertation is considered to be of innovative and non-linear origin because of uniquely formed tremor, cognitive and energy expenditure

feature combinations for model inputs. Thus, the motivation for selecting supervised learning algorithms that fit the collected data properly can be indeterminate. In such case, the author of this dissertation chooses a strategy to investigate a number of the most commonly used supervised learning algorithms and search for appropriate solution. Firstly, five groups of supervised learning methods are examined: functions, Bayesian, trees, rules and lazy classifiers [22] [23]. Function group classifiers can be defined by using mathematical equations. Bayesian group classifiers are based on Bayes Theorem of Probability [24] and allocate the element value to a population from one of the categories that are available. Tree group classifiers apply graphical representation that makes use of branching methodology to exemplify all possible outcomes of a decision, based on certain condition. The supervised learning based on rules is a method to represent knowledge of the system by a set of expressions (e.g., IF: THEN). Lazy classifiers store the training instances and do the real work only when the actual classification starts.

9.5. Conclusions and future work

9.5.1. Conclusions

1. The related work analysis provides a comprehensive study of existing data mining and machine learning methods for health state assessment for patients with central nervous system disorders (CNSD). The minority of these solutions target Huntington Disease (HD) patients in the described investigations. In addition, CNSD patient monitoring approaches are very different and have limitations, e.g., the first does not use a mobile device; the second adapts extra; the third uses party wearable devices that usually increase the stress factor for patients and are uncomfortable. None of the related solutions proposes an extended methodology that applies a combination of feature extraction methods to target tremor, cognitive, speech and energy expenditure impairments occurring in CNSD patients as it is presented in this dissertation.
2. The restrictions of existing solutions have been addressed by gathering a priori knowledge from the medical staff and developing a computerized version of SAGE (Self-administered

cognitive testing) methodology. The proposed system is an actual framework that requires only one smart non-invasive interface, i.e., a mobile device or a tablet for CNSD patient neural impairment screening. The framework supports 16 tasks: 12 of them are based on SAGE methodology; others adapt finger tapping, spiral contour following (drawing), voice recording and energy expenditure tasks as an extension. In total, more than 230 features are automatically extracted in the framework, i.e., finger motion tracking, task duration, distance estimation of geometrical shapes, graph similarity evaluation, image collection from clock drawing (CDT) task, SAGE scores, voice recordings, daily calorie balances etc. Another innovative aspect of this dissertation is the creation of newly designed dataset, which opens a gateway for using it in a machine learning repositories (e.g., UCI) by other computer scientists.

3. The practical significance of this dissertation has been evaluated by the implementation of Neural Impairment Test Suite (NITS) mobile application, which provides feedback on the patient's health status in graphical form and compares results with healthy subject's estimates. NITS is available on Google Play web portal. More than 500 users have already installed NITS app at the time of writing this dissertation. However, the author notes that the proposed mobile app is designed as a curiosity tool to self-evaluate behaviour changes for people in a risk group of suspected symptoms of neural impairments (e.g., Huntington, Parkinson diseases, MCI, dementia) and does not replace professional medical assessment. Moreover, only reliable data, which is approved by the data collection permit of Kaunas University of Technology (KTU) and signed with a test subject on agreement document, is used in this work. Other, unknown sources are ignored.
4. Four experiments were carried out to solve sick vs. healthy binary classification problem. The collected dataset was used to validate the proposed experiments. Tremor, cognitive, energy expenditure features are considered in experiments E1 and E2. E1 applies innovative feature distribution method, i.e., for each proposed individual methodological task, the features are

grouped separately. Furthermore, three extra unique feature combinations were composed for the designed classification models: proper spelling, SAGE points and task durations in seconds. E1 showed the best results for task T9 with LWL classifier = 91.50% (lowest achieved accuracy was 74.52). In experiment E2, combining all 238 features into a single classifier, it resulted in 3% increased accuracy model, compared to the best performance in E1 with AdaBoostM1 (Random Forest) = 94.57% classifier. In addition, all 12 proposed classifiers for E2 experiment achieved at least 91.47% accuracy. E3 and E4 experiments deal with speech impairments. E3 applies bidirectional recurrent neural network with long short-term memory (BiLSTM) and achieved an accuracy of 86.63% for a test set (100% for training model and network converge). E4 uses wavelet scattering transform (WST) method to achieve an accuracy of 100% for a test set. WST overcomes BiLSTM, considering the fact that collected voice recording from CNSD patients were significantly long, i.e., up to 47 seconds.

5. In a classifier ensemble (hybrid model) design, multiple combinations with ~ 2% increased accuracy, i.e., 96.12%, were found by using heuristic approach. Fusing 3 classifiers: AdaBoostM1 (SMO), AdaBoostM1 (BayesNet) and Voted Perceptron (PCA) with Vote method (Average of Probabilities combination rule), results in the best-fit combination based on the highest accuracy value and minimum time expenditure (0.59 seconds) for model training. The hybridization approach targets tremor, cognitive and energy expenditure feature set, leaving out the audio features from standalone experiments E3 and E4, as they require special software environment for voice signal processing and extra computational resources.

9.5.2. Future Work

The built models cover a very wide range of possible disorders, occurring to those in Huntington, Parkinson or MCI disease patients: one can have tremor in hands or body; the others may have memory loss, voice problems or weight loss. This implies a high potential of determination of patient's health state deterioration status. The designed theoretical models (based on experiments E1, E2, E3 and E4) can be integrated into the proposed NITS mobile app for the prediction of alerts (alarms) to a patient on his disease progression before the symptoms occur visually.

Acknowledgements

The author would like to thank the president of Lithuania Huntington disease association Dr. Zivile Navikiene for contacting family members of HD patients and disseminating the proposed research in the official website and the Facebook channel of the association to help and carry out with the experiments described in this dissertation as well as for practical support and advices.

Human Studies

The research was approved by the Institutional Review Board of the Faculty of Informatics, Kaunas University of Technology.

UDK 616.8:004.5 + 004.5:616.8](043.3)

SL344. 2020-02-05, 5 leidyb. apsk. l. Tiražas 50 egz.

Išleido Kauno technologijos universitetas, K. Donelaičio g. 73, 44249 Kaunas
Spausdino leidyklos „Technologija“ spaustuvė, Studentų g. 54, 51424 Kaunas