



**IBSU**

**შავი ზღვის საერთაშორისო უნივერსიტეტი  
კომპიუტერული ტექნოლოგიებისა და საინჟინრო საქმის ფაკულტეტი  
კომპიუტერული მეცნიერებების პროგრამა**

**ადამიანის ყოველდღიური აქტივობის ამოცნობა სხეულზე  
დამაგრებული სენსორების გამოყენებით**

**მარიამ დედაბრიშვილი  
კომპიუტერული მეცნიერებების სადოქტორო დისერტაციის ავტორეფერატი**

**თბილისი, 2018**

**სამეცნიერო ხელმძღვანელი:**

ირაკლი როდონია

---

შავი ზღვის საერთაშორისო უნივერსიტეტის პროფესორი, დოქტორი

---

(ხელმძღვანელის ხელმოწერა)

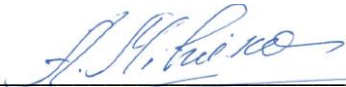
**სამეცნიერო თანახელმძღვანელი:**

ალექსანდრე მილნიკოვი

---

დოქტორი

---



(თანახელმძღვანელის ხელმოწერა)

**ექსპერტები (სახელი, გვარი & აკადემიური წოდება):**

1. პროფ. დრ. ნოდარ მომცელიძე

---

2. პროფ. დრ. ბესიკ დუნდუა

---

3. ასოც. პროფ. დრ. გიორგი მანდარია

---

**ოპონენტები (ექსპერტები (სახელი, გვარი & აკადემიური წოდება):**

1. პროფ. დრ. დაიანა კუკი

---

2. ასისტ. პროფ. დრ. დანიელა მიკუჭი

---

3. ასოც. პროფ. დრ. ხათუნა ბარდაველიძე

---

## შესავალი

წინამდებარე ნამუშევარი განიხილავს ადამიანის ყოველდღიური აქტივობის კლასიფიკაციას სხეულზე დამაგრებული ინერციული სენსორების მეშვეობით. სიარული, წოლა, ფეხზე წამოდგომა და ა.შ. ამ აქტივობების მაგალითებია. კვლევაში გამოყენებული მონაცემები მოიცავს სამი ინერციული სენსორის გამოყენებით განხორციელებულ 12 აქტივობას: კიბეზე ჩასვლა, დგომა, დაჯდომა, ჯდომა, ჯდომითი მდგომარეობიდან მიწაზე გადაჯდომა, მიწაზე ჯდომა, დაწოლა, წოლა, წოლითი მდგომარეობიდან მიწაზე მჯდომიარე მდგომარეობაში გადასვლა, ადგომა, სიარული და კიბეზე ასვლა. კვლევის ფარგლებში მოხდება შედარება ოთხი კონტროლირებადი კლასიფიკაციის ტექნიკის, კერძოდ, k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), Supervised Learning Gaussian Mixture Models (SLGMM) და Random Forest (RF) აგრეთვე, სამი არაკონტროლირებადი კლასიფიკაციის ტექნიკის, მათ შორის, k-Means, Gaussian Mixture Models (GMM) და Hidden Markov Model (HMM) მათი კლასიფიკაციის სიზუსტის, F-measure, recall, precision და specificity-ის თვალსაზრისით. თითოეულ კლასიფიკატორში გამოყენებული იქნება, როგორც დაუმუშავებელი, ასევე დამუშავებული მონაცემები. ინერციული სენსორები დამაგრებულია რამოდენიმე ჯანმრთელი სუბიექტის სხეულის ზედა და ქვედა კიდურების საკვანძო ადგილებში (მკერდი, მარჯვენა ბარძაყი და მარცხენა ტერფი). აქტივობის ამოცნობის პროცესში სამი ძირითადი ეტაპია: სენსორების განლაგება, მონაცემთა წინასწარი დამუშავება და მონაცემთა კლასიფიკაცია. კვლევაში გამოყენებულია მხოლოდ აჩქარების მონაცემები, როგორც აქტივობის შეფასების მოდალობა (Altun, Barshan, & Tunçel, 2010; Chamroukhi, Mohammed, Trabelsi, Oukhellou, & Amirat, 2013). კვლევის ფარგლებში წარმოდგენილი და გაანალიზებული იქნება კონტროლირებადი და არაკონტროლირებადი კლასიფიკაციის ალგორითმების მეშვეობით მიღებული შედეგები. უფრო მეტიც, შემუშავებული იქნება ჰიბრიდული მეთოდები კლასიფიკაციის ალგორითმების გაუმჯობესებული შესრულების მიზნით.

## პრობლემის დასმა

საიმედო და ზუსტი მეთოდების შემუშავება ადამიანის აქტივობის ამოცნობაში უაღრესად მნიშვნელოვანია, ვინაიდან აქტივობის არასწორმა ან არაზუსტმა ამოცნობამ შეიძლება გამოიწვიოს ადამიანის ჯანმრთელობისთვის საზიანო შედეგები. HAR-ის სფეროში მომუშავე მეცნიერები ცდილობენ აღმოაჩინონ გზები ამოცნობის სიზუსტის მიღწევების გასაუმჯობესებლად. აღნიშნულის გათვალისწინებით, არსებითი მნიშვნელობის მქონეა ისეთი ალგორითმების შერჩევა, რომლებიც ახორციელებენ კლასიფიკაციას საიმედო მაჩვენებლებით. თუმცა, ახლახან არსებული ალგორითმების ლიმიტირებულობა და მათი სიზუსტის დონის დაბალი მაჩვენებელი, კითხვის ნიშნის ქვეშ აყენებს მათ გამოყენებადობას აღნიშნულ სფეროში. კერძოდ, დამალულ მარკოვის მოდელებს გააჩნიათ გარკვეული შეზღუდვები, რაც გამოწვეულია შემთხვევითობის პრინციპით შერჩეული პარამეტრებით, რომელიც თავის მხრივ ართულებს კლასებს შორის დისკრიმინაციას მაღალი სიზუსტით.

სხვა კარგად ცნობილ მეთოდებსაც აგრეთვე გააჩნია ზოგიერთი ნაკლოვანება მათი ბუნებიდან გამომდინარე. აღნიშნული პრობლემების გადაჭრის მიზნით და საჭირო შედეგების მისაღწევად გადაწყვეტილი იქნა კლასიფიკატორების ჰიბრიდული კომპლესების შექმნა გაუმჯობესებული სიზუსტის და სწავლების სფეროს ადექვატური მოდელის უზრუნველსაყოფად. დისერტაციის მეთოდოლოგიის შემუშავებისას, გათვალისწინებული იქნა საფრანგეთში, პარიზში, ლისის ლაბორატორიაში მომუშავე ცნობილი მკვლევარების რეკომენდაციები.

შედეგად, წინამდებარე თეზისში შემუშავებული იქნა მეთოდოლოგია და ალგორითმები, რომლებიც წარმატებით წყვეტენ დასმულ პრობლემას. კერძოდ, შემუშავებული ალგორითმების ჰიბრიდი უფრო მეტად ბუნებრივი და გამოყენებადია HAR-ის კონტექსტში, სხვა არსებულ მეთოდებთან შედარებით.

შედეგების და მიღწევების ნაწილი გამოქვეყნებული იქნა სამეცნიერო ჟურნალ “Sensors”-ში, რომელსაც გააჩნია მაღალი impact factor. სტატიაში დაიმსახურა სამეცნიერო საზოგადოების დიდი დაინტერესება, რაც შეიძლება დადასტურდეს ციტირებების მაღალი მაჩვენებლით. სტატიისადმი გამომჟღავნებული ინტერესის მხედველობაში მიღებით, გადაწყვეტილი იქნა მეთოდოლოგიის შემუშავების

გაგრძელება ცნობილი კლასიფიკაციის მეთოდების კომბინაციების თვალსაზრისით, რაც მიზნად ისახავს ადამიანის ყოველდღიური, შერჩეული აქტივობების ამოცნობის სიზუსტის მნიშვნელოვან გაუმჯობესებას.

### **სწავლის მიზნობრიობა**

კვლევის ძირითადი მიზანია იმ კლასიფიკატორების აღმოჩენა, რომლებიც ყველაზე მეტად მიესადაგება კონტროლირებად და არაკონტროლირებად გარემოში ადამიანის ყოველდღიური ფიზიკური ქმედების ამოცნობას.

კვლევის მიზნების განსახორციელებლად შერჩეული იქნება ადამიანის მიერ ყოველდღიურ ცხოვრებაში განხორციელებული თორმეტი ძირითადი აქტივობა. მონაცემები შეგროვდება მობილური ტექნოლოგიების გამოყენებით, კერძოდ, ადამიანის სხეულზე დამაგრებადი ინერციული სენსორებით.

კვლევის ფარგლებში გამოკვლეული იქნება აქტივობის ამოცნობის პროცესის სამი ძირითადი ეტაპი: სენსორების დამაგრება, მონაცემების დამუშავება და მონაცემების კლასიფიკაცია. სწავლებაში გამოიყენება ოთხი კონტროლირებადი კლასიფიკაციის ალგორითმი, კერძოდ, k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), Supervised Learning Gaussian Mixture Models (SLGMM) და Random Forest (RF) და ამასთან, სამი არაკონტროლირებადი კლასიფიკაციის მეთოდი, მათ შორის, k-Means, Gaussian Mixture Models (GMM) და Hidden Markov Model (HMM), რომელთა შედარება მოხდება მათი კლასიფიკაციის სიზუსტის, F-measure, recall, precision და specificity-ის მაჩვენებლებით. თითოეულ კლასიფიკატორში ცალ-ცალკე იქნება გამოყენებული დაუმუშავებელი და დამუშავებული მონაცემები. მონაცემები აგრეთვე იქნება შესწავლილი კონტროლირებადი და არაკონტროლირებადი კლასიფიკაციის ტექნიკების კომბინაციების მეშვეობით.

### **სწავლის ორიგინალურობის მაჩვენებლები**

- ✓ სენსორების გადანაწილება – სხვა კვლევებში წარმოდგენილია განსხვავებული კომბინაციები
- ✓ შერჩეული აქტივობები და მათი თანმიმდევრობა
- ✓ მონაცემების გაფილტვრის და ზომაში შემცირების ხერხები

- ✓ ნაპოვნი მახასიათებლების რიცხვი
- ✓ სხვა კვლევებში მიღებულ შედეგებზე დაყრდნობით HAR-ში გამოყენებული ყველაზე ცნობილი ალგორითმების შერჩვა
- ✓ კონტროლირებადი ალგორითმების გარემოსთვის მეტად შესაფერისი კლასიფიკატორის გამორკვევა შერჩეულ ალგორითმებს შორის მათი სიზუსტის მაჩვენებლის მიხედვით
- ✓ არაკონტროლირებადი ალგორითმების გარემოსთვის მეტად შესაფერისი კლასიფიკატორის გამორკვევა შერჩეულ ალგორითმებს შორის მათი სიზუსტის მაჩვენებლის მიხედვით
- ✓ სიზუსტის მაჩვენებლის გაუმჯობესება 9%-ით სენსორების განლაგებისა და შერჩეული აქტივობების კერძო შემთხვევაში NN-HMM-ის ჰიბრიდული ალგორითმის გამოყენებისას
- ✓ განსხვავებული ბუნების მქონე ალგორითმების კომბინაციების გამოყენება HAR-ში, რომელიც ჩვენი ცოდნით აქამდე არ არის გამოყენებული აღნიშნულ სფეროში.

### **კვლევის მიზნები**

თეზისის კვლევის მიზნებია HAR-ისთვის ყველაზე მეტად შესაფერისი კლასიფიკატორის დადგენა, რომელიც გააუმჯობესებს მიღწევებს ამავე სფეროში და საშუალებას მისცემს მომხმარებლებს ამოიცნონ ადამიანის ყოველდღიური საქმიანობა მაღალი სიზუსტით.

### **სიახლე და აქტუალობა**

კვლევის მიღწევებზე დაყრდნობით თეზისის სიახლე და აქტუალობა შეიძლება აიხსნას მოცემული ფაქტორებით:

- ✓ სწავლის ფარგლებში გამოყენებული ცნობილი კონტროლირებადი ალგორითმებიდან HAR-ისთვის ერთ-ერთი საუკეთესო კლასიფიკატორის დადგენა.

- ✓ სწავლის ფარგლებში გამოყენებული ცნობილი არაკონტროლირებადი ალგორითმებიდან HAR-ისთვის ერთ-ერთი საუკეთესო კლასიფიკატორის დადგენა.
- ✓ შემუშავდა ალგორითმების ახალი ჰიბრიდები მონაცემთა კლასიფიკაციისათვის, რომლებიც აუმჯობესებენ ცალკეულად გამოყენებულ კლასიფიკატორთა შედეგებს HAR-ის სფეროში.
  - ა. შემუშავდა Instance-based ალგორითმისა და Naïve Bayes-ის კომბინირებული კლასიფიკატორი.
  - ბ. შემუშავდა Hidden Markov Model-ისა და Artificial Neural Networks – ის კომბინირებული კლასიფიკატორი
- ✓ კვლევებმა ცხადყო, რომ ერთი ალგორითმის სიძლიერე შეიძლება გამოყენებულ იქნას სხვა ალგორითმის სუსტი მხარის ჩასანაცვლებლად.
- ✓ კვლევის შედეგებმა აჩვენა, რომ სხვადასხვა კლასიფიკაციის ალგორითმებში დატრენინგებული მონაცემების გაერთიანებით იზრდება სიზუსტის მაჩვენებელი.
- ✓ სამომავლოდ შემუშავდა ჰიბრიდული კლასიფიკატორის ფუნქციის შექმნის გზა სხვა მომხმარებლებისთვის მომიჯნავე სფეროებში გამოსაყენებლად.

### **პრობლემის მნიშვნელობა**

ადამიანის აქტივობის ამოცნობის პრობლემა მაღალი სიზუსტის მაჩვენებლით ძალზედ მნიშვნელოვანია მისი შედეგების მგრძობელობის გამო, რაც შეიძლება აისახოს ადამიანის ჯანმრთელობაზე მავნე ზემოქმედებით.

### **პრაქტიკული და თეორიული მნიშვნელობა**

მონაცემთა კლასიფიკაციისათვის ალგორითმების შერჩევა და ჰიბრიდული კლასიფიკატორის შექმნა, რომელიც აუმჯობესებს ადამიანის აქტივობის ამოცნობაში სიზუსტის მაჩვენებელს, ხელს უწყობს სხვადასხვა დაავადების მქონე ინდივიდებისა და ხანდაზმული მარტოხელა მოსახლეობისთვის, რომელთა რიცხვიც დღეისათვის

მნიშვნელოვნად გაიზარდა, დამოუკიდებლობისა და კეთილდღეობის გახანგრძლივებას.

## **კვლევის მეთოდები**

ჩატარებულ კვლევებში ადამიანის აქტივობა შეფასებულია Xbus Kit-ის გამოყენებით (Enschede, Netherlands), რომელიც უზრუნველყოფს ადამიანის მოძრაობის ამბულატორულ გაზომვას. იგი მოიცავს პორტაბელურ ჩარჩოს, რომელიც აერთიანებს Xbus Master-ს და სამ MTX ინერციული სენსორის ერთეულს. სწავლის ფარგლებში სენსორები დამაგრებულია სუბიექტების გულმკერდის არეში, მარჯვენა ბარძაყზე და მარცხენა კოჭზე.

## **დისერტაციის სტრუქტურა**

დისერტაცია მოიცავს შესავალს, ლიტერატურის მიმოხილვას, თეორიულ ნაწილს, პრაქტიკულ იმპლემენტაციას, დასკვნას, 39 ფიგურას, 21 ცხრილს და 181 გამოყენებულ ლიტერატურულ წყაროს.

### **კვლევის ძირითადი შინაარსი**

#### **თავი 1. კვლევის მიზნები და ლიტერატურის მიმოხილვა**

თეზისის პირველ თავში გადმოცემულია ადამიანის აქტივობის ამოცნობის მეთოდების კრიტიკული ანალიზი. კვლევამ ცხადყო, რომ სხვადასხვა სწავლაში გამოყენებული ალგორითმების ეფექტურობა რთულად შესადარებელია რამოდენიმე მიზეზის გამო. სიძნელე უმეტესად დაკავშირებულია (i) ექსპერიმენტალური პროტოკოლის მრავალგვარობასთან (მონაწილე სუბიექტების რაოდენობა, ამოსაცნობი აქტივობების სახე და რიცხვი – ამბულატორიული, გადაადგილებადი, ყოველდღიური აქტივობები, ვარჯიში/ფიტნესი, სხვადასხვა აქტივობების ხანგრძლივობა და თანმიმდევრობა და ა. შ.); (ii) ადამიანის აქტივობის ამოცნობის მოტივებთან (მონიტორინგი, წაქცევის ამოცნობა, სახლის პირობებში მკურნალობა და ა. შ.); (iii) გამოყენებული სენსორების სახეობასთან (აქსელერომეტრი, პლანტარული წნევის სენსორი, გიროსკოპი) და მათი განლაგებასთან სხეულზე (მაჯა, გულმკერდი, თეძო, ბარძაყი, ყელი) (iv) მონაცემების მოპოვების (mining) სახესხვაობასთან სხვადასხვა

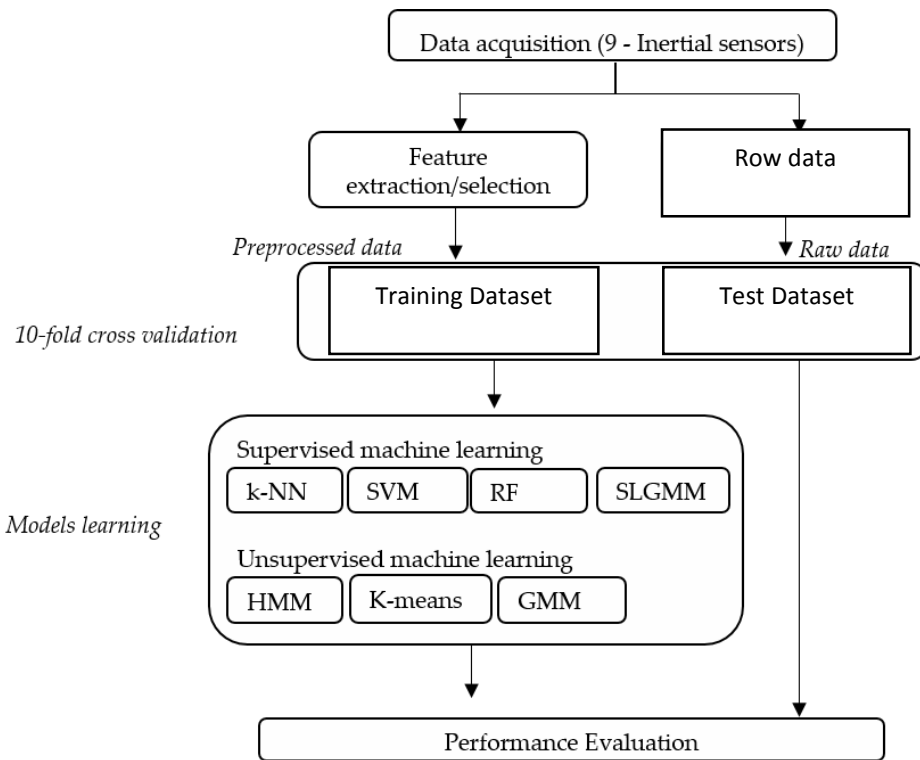


ლიტერატურაში (მახასიათებლების გამოყოფა, ამორჩევა, მონაცემების შემცირება, კლასიფიკაცია); (v) შეფასების კრიტერიუმთან (accuracy, F-measure, recall, precision, specificity და ა.შ.) დადასტურების (ვალიდაციის) პროცედურასთან (P-fold, leave one out, repeated random sub-sampling, bootstrap, და ა.შ.). ადამიანის აქტივობის ამოცნობის თანამედროვე მიდგომების გარემოებებისა და ლიმიტაციების გაანალიზების საფუძველზე და მათი მხედველობაში მიღებით, განისაზღვრა კვლევის შემდეგი მიზნები, ოპტიმალური გზის მოსაძებნად და მაღალი სიზუსტის მაჩვენებელს მისაღწევად:

### სწავლის მიზნები

- სწავლის ფარგლებში დადგინდეს ცნობილ კონტროლირებად ალგორითმებს შორის HAR-ზე მორგებული, უმჯობესი კლასიფიკატორი.
- სწავლის ფარგლებში დადგინდეს ცნობილ არაკონტროლირებად ალგორითმებს შორის HAR-ზე მორგებული, უმჯობესი კლასიფიკატორი.
- შემუშავდეს ალგორითმების ახალი ჰიბრიდები მონაცემთა კლასიფიკაციისათვის, რომლებიც გააუმჯობესებენ ცალკეულ კლასიფიკატორთა შედეგებს HAR-ის სფეროში.
  - შემუშავდეს Instance-based ალგორითმისა და Naïve Bayes-ის კომბინირებული კლასიფიკატორი.
  - შემუშავდეს Hidden Markov Model-ისა და Artificial Neural Networks – ის კომბინირებული კლასიფიკატორი
- დადასტურდეს, რომ ერთი ალგორითმის სიძლიერე შეიძლება გამოყენებულ იქნას სხვა ალგორითმის სუსტი მხარის ჩასანაცვლებლად.
- მოხდეს თვალსაჩინოება, რომ სხვადასხვა კლასიფიკაციის ალგორითმებში დატრენინგებული მონაცემების გაერთიანებით შესაძლებელია სიზუსტის მაჩვენებლის გაზრდა.
- სამომავლოდ შემუშავდეს ჰიბრიდული კლასიფიკატორის ფუნქციის შექმნის გზა სხვა მომხმარებლებისთვის მომიჯნავე სფეროებში გამოსაყენებლად.

*ფიგურა 1* თავს უყრის შემუშავებული მიდგომის პირველი ნაწილის შემადგენელ ეტაპებს, რომლებიც ნაბიჯ-ნაბიჯ განხორციელდება სწავლის პროცესში.



ფიგურა 1. ადამიანის აქტივობის ამოცნობა სხვადასხვა ალგორითმის გამოყენებით.

## თავი 2. თეორიული მოდელები და პრობლემის გადაჭრის გზები

მეორე თავში მოკლედ აღწერილი და გაანალიზებულია ადამიანის აქტივობის ამოცნობისთვის სწავლაში გამოყენებული კლასიფიკაციის ტექნიკები მათი მახასიათებლების მიხედვით (GMMs, k-Nearest Neighbors (k-NN), Naïve Bayes, SVMs, Random Forests (RFs), K-means, ANN და HMMs), რასაც მივყავართ გადაწყვეტილებამდე, რომ მოხდეს სხვადასხვა მეთოდების გაერთიანება ალგორითმის შესრულების გაუმჯობესების მიზნით.

## კონტროლირებადი და არაკონტროლირებადი დასწავლის ალგორითმების ზოგადი საკითხები

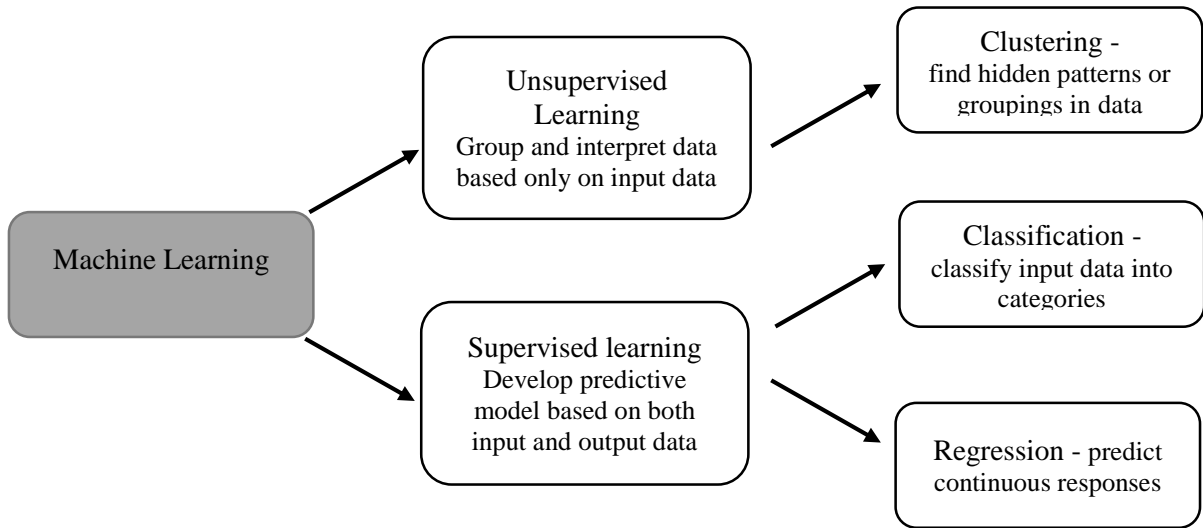
პროგნოზირებადი მონაცემების მოპოვება (mining) წარმოადგენს ყველაზე მნიშვნელოვან მიმართულებას Machine Learning (ML)– ის ისფეროში. ნიმუშები, რომლებიც გამოიყენება ML ალგორითმებში წარმოადგენილია მათი მახასიათებლების მეშვეობით. გამოყოფილი/შერჩეული მახასიათებლები სენსორის პირდაპირი მონაცემებიდან გამოიყენება, როგორც შესაყვანი (input) მონაცემები კლასიფიკაციის

ალგორითმებში. ადამიანის აქტივობის ამოცნობის შემთხვევაში შესაყვანი მონაცემების ნიმუშები ასოცირდება ცალკეულ აქტივობებთან (კლასებთან). ზოგადად, კლასიფიკაციის ამოცანა მოითხოვს გადაწყვეტილების მიღების ან ფუნქციის დასწავლას, რომელიც დაუკავშირებს შესაყვან მონაცემებს კლასებს (Duda, Hart, & Stork, 1999; Webb, 2003; Theodoridis, Piskris, Koutroumbas, & Cavouras, 2010). ნიმუშების კატეგორიზაცია შეიძლება აიხსნას ორი გზით: პირველი, რომელსაც გააჩნია სახელი (label) ან შესაბამისი სწორი შედეგი (output) და ხორციელდება სუბიექტის ზედამხედველობის ქვეშ, იწოდება როგორც, კონტროლირებადი (Supervised) და მეორე, არაკონტროლირებადი (Unsupervised) დასწავლა, სადაც ნიმუშები სახელების გარეშეა (unlabeled) და კლასიფიკაცია ხორციელდება სხვადასხვა პროგნოზირებადი მეთოდების საშუალებით (Jain AK, 1999).

კონტროლირებად დასწავლაში, სახელის მქონე მონაცემებზე დაყრდნობით ხდება ალგორითმის დატრენინგება წინასწარ განსაზღვრული კონცეფციებისა და ფუნქციების გამოყენებით (Zoila Ruiz, 2017). კონტროლირებადი დასწავლის მეთოდები ცდილობენ აღმოაჩინონ ურთიერთდამოკიდებულება შესაყვან ატრიბუტებსა (ე.ი. დამოუკიდებელი ცვლადი) და მიზნობრივ ატრიბუტებს (ე.ი. დამოკიდებული ცვლადი) შორის. აღმოჩენილი მოდელები ზოგადად განსაზღვრავენ და ხსნიან ფენომენას მონაცემებში არსებული სურათის შესახებ და იძლევიან საშუალებას დამოუკიდებელი ცვლადის ცნობილი მნიშვნელობის გათვალისწინებით მოხდეს დამოკიდებული ცვლადის მნიშვნელობის პროგნოზირება (Maimon & Rokach, 2005). ნიმუშებიდან (მაგალითები დასატრენინგებელი მონაცემებიდან) წესების დასწავლის პროცესი იწოდება ინდუქციურ მანქანურ სწავლებად, ან სხვა სიტყვებით, კლასიფიკატორის შექმნა, რომელიც გამოიყენება ახალი ნიმუშების განსაზღვრებად (Kotsiantis, Zaharakis, & Pintelas, 2007).

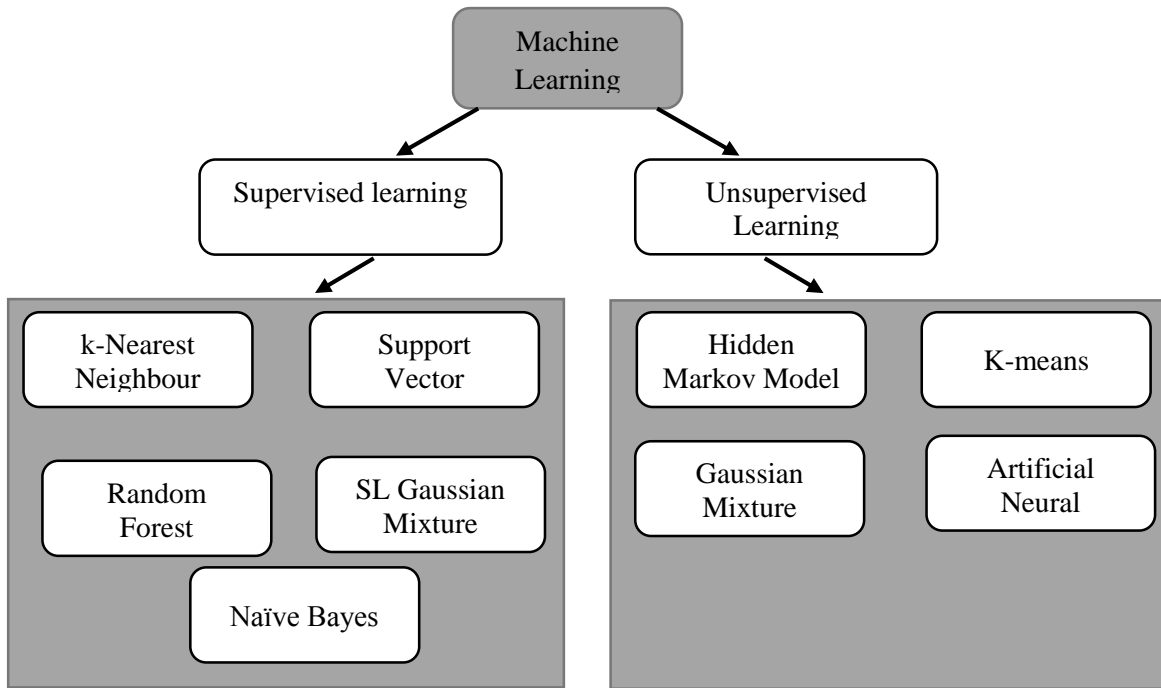
არაკონტროლირებად დასწავლაში, ალგორითმებს ევალებათ აღმოაჩინონ მოცემულ ნიმუშებში საინტერესო ატრიბუტები (Attal, Mohammed, Dedabrishvili, & Chamroukhi, 2015). არაკონტროლირებადი დასწავლის ამოცანაა დაადგინოს, თუ როგორ შეუძლიათ სისტემებს კონკრეტული შესაყვანი ნიმუშის წარმოჩენა იმგვარად, რომ არ დაირღვეს მთელი შესაყვანი ნიმუშების სტატისტიკური სტრუქტურა. კონტროლირებად

დასწავლასთან შედარებით, არ არსებობს კონკრეტული მიზნობრივი შედეგი (output), არც გარემოებაზე დაფუძნებული გამოთვლები, რომლებიც უკავშირდება თითოეულ შენატანს (input); ნაცვლად ამისა, არაკონტროლირებადი დასწავლის ალგორითმი ინარჩუნებს პირველად კავშირს თუ შენატანის სტრუქტურის რა მახასიათებლები აისახება შედეგზე (output). არაკონტროლირებადი დასწავლა უფრო ახალა ადამიანის ტვინის სტრუქტურასთან და ამიტომაც არის მნიშვნელოვანი.



ფიგურა 2. მანქანური დასწავლის ტექნიკები.

მაგალითად, Artificial Neural Networks (Webb, 2003) და Support Vector Machines (SVM) (Vapnik, 2000) წარმოადგენენ კონტროლირებადი დასწავლის მეთოდებს და საჭიროებენ მთლიანად დასათაურებელ მონაცემებს კლასიფიკაციისთვის; მაშინ როდესაც არაკონტროლირებადი დასწავლის მეთოდები, როგორებიცაა Gaussian Mixture Models (GMMs) და Hidden Markov Models (HMMs) (Rabiner, 1989) საშუალებას იძლევიან ავტომატურად მოახდინონ მონაცემების დასათაურება.



ფიგურა 3. კვლევისთვის შერჩეული ალგორითმების სია.

თეზისის ეს ქვეთავი მთლიანობაში წარმოადგენს კლასიფიკაციის ტექნიკების თეორიულ მიმოხილვას ადამიანის აქტივობის ამოცნობისთვის, რაც შემდგომ გამოყენებულია სწავლის ფარგლებში. განხილული და შეჯამებულია კონტროლირებადი და არაკონტროლირებადი დასწავლის ტექნიკების ზოგადი საკითხები და ამასთან ცალკეული ალგორითმების აღწერილობა. რამდენადაც ალგორითმებს გააჩნიათ სუსტი და ძლიერი მხარეები ადამიანის აქტივობის ამოცნობის საქმეში, თეზისში ერთის მხირვ წარმოდგენილია distance-based და Naïve Bayes ალგორითმების კომბინაცია, და მეორეს მხირვ, ANN-ის და HMM-ის ჰიბრიდი. სხვადასხვა ალგორითმების გამოყენება მონაცემებზე ცხადყოფს, რომ ალგორითმების ცალკეული გამოყენება შედარებით დაბალ შედეგს გვაძლევს, ვიდრე კომბინირებული ტექნიკები. ამგვარად, ალგორითმი, რომელიც იყენებს სხვა მეთოდის ძლიერ მხარეს თავისი სუსტი მხარის ჩასანაცვლებლად, გაუმჯობესებული შედეგის მომცემია.

**პრობლემა 1: კონტროლირებადი ალგორითმების კომბინაციის მეთოდოლოგია**

ერთიანი კლასიფიკატორების შექმნისთვის არსებობს სხვადასხვა მეთოდები (Sergey Tulyakov, 2008). მიუხედავად იმისა, რომ წარმოდგენილია ალგორითმების გაერთიანების რამოდენიმე ტექნიკა, რთული სათქმელია თუ რომელი მეთოდია

უმჯობესი (Villada & Drissi, 2002). აქედან გამომდინარე, კლასიფიკატორების კარგი კომბინაციის შექმნა წარმოადგენს კვლევის აქტიურ სფეროს კონტროლირებად დაწესავლაში. არსებობს სამი ძირითადი მეთოდოლოგია კლასიფიკატორების გაერთიანების თვალსაზრისით: (i) სხვადასხვა დასატრენინგებელი მონაცემების საფუძველზე კონკრეტული დასწავლის ტექნიკით, (ii) დასატრენინგებელ მონაცემებზე სხვადასხვა პარამეტრების მინიჭებით კონკრეტული training ტექნიკით (მაგ. სხვადასხვა initial weights) და (iii) სხვადასხვა დასწავლის ტექნიკების გამოყენებით. კლასიფიკატორების შერწყმისას შეიძლება დამატებითი ინფორმაციის მიღება სხვადასხვა წყაროების შერევით. ყველა ზემოთ აღწერილი კომბინაციის მეთოდი შეიძლება გამოყენებულ იქნას გაუმჯობესებული შედეგების მისაღებად (Lazkano & Sierra, 2003; Sierra, Lazkano, Martinez-Otzeta, & Astigarraga, 2003).

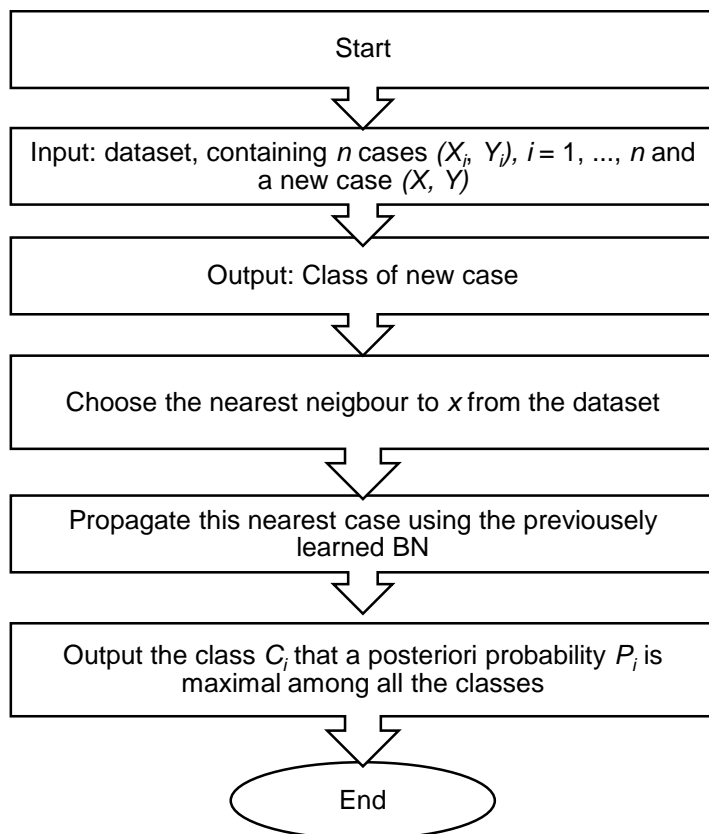
სტატისტიკაზე დაფუძნებული Naïve Bayes კლასიფიკატორი და მანძილზე დაფუძნებული K-Nearest Neighbor ხშირად გამოიყენება პროგნოზირების პრობლემებთან (Ferdousy, Islam, & Matin, 2013). Naïve Bayes ალგორითმთან დაკავშირებით აღსანიშნავი ფაქტორია ციფრულ ატრიბუტებთან მუშაობა, რადგან ალგორითმში უნდა განისაზღვროს პირობითი ალბათობა ყველა ატრიბუტის თითოეული შესაძლო მნიშვნელობისთვის. ამ პრობლემის გადასაჭრელად, რიცხვითი ატრიბუტები უნდა დაიყოს მრავალრიცხოვან კლასებად სხვადასხვა არსებული რომელიმე დისკრეტიზაციის მეთოდით. შედეგად, დისკრეტიზაციისთვის შერჩეულ ტექნიკას მნიშვნელოვანი როლი აკისრია მეთოდის სიზუსტესთან მიმართებაში. ახალი დისკრეტული სტრუქტურის შექმნის გზით განხორციელდა რამოდენიმე მცდელობა Naïve Bayes ალგორითმის სიზუსტის გასაზრდელად (Yang & Webb, 2002).

სიტუაცია მკვეთრად განსხვავდება K Nearest Neighbor ალგორითმის შემთხვევაში. აქ საქმე ეხება კატეგორიულ ატრიბუტებს. ვინაიდან ალგორითმი ირჩევს დასატრენინგებელი მონაცემების კონკრეტულ სემპლს დისტანციის მიხედვით, საჭიროა შეიქმნას დაშორების გაზომვის სქემა კატეგორიული მონაცემებისთვისაც. როგორც წესი, ეს ხორციელდება სხვადასხვა მსგავსების დასადგენი ტექნიკით. ალგორითმი, რომელიც გამოიყენება თეზისში აერთიანებს ამ ორ კლასიფიკატორს

იმგვარად, რომ ორივე ზემოთ აღნიშნული პრობლემატური საკითხი იქნას გადაჭრილი.

კერძოდ, ალგორითმში აღარ არსებობს დამატებითი მოთხოვნა უწყვეტი ცვლადების დისკრედიტაციისთვის და, ამავდროულად, არ საჭიროებს დაშორების გაზომვას კატეგორიულ ატრიბუტებს შორის. კომბინაცია ზრდის ალგორითმის ეფექტურობას და საიმედო ხასიათს ატარებს.

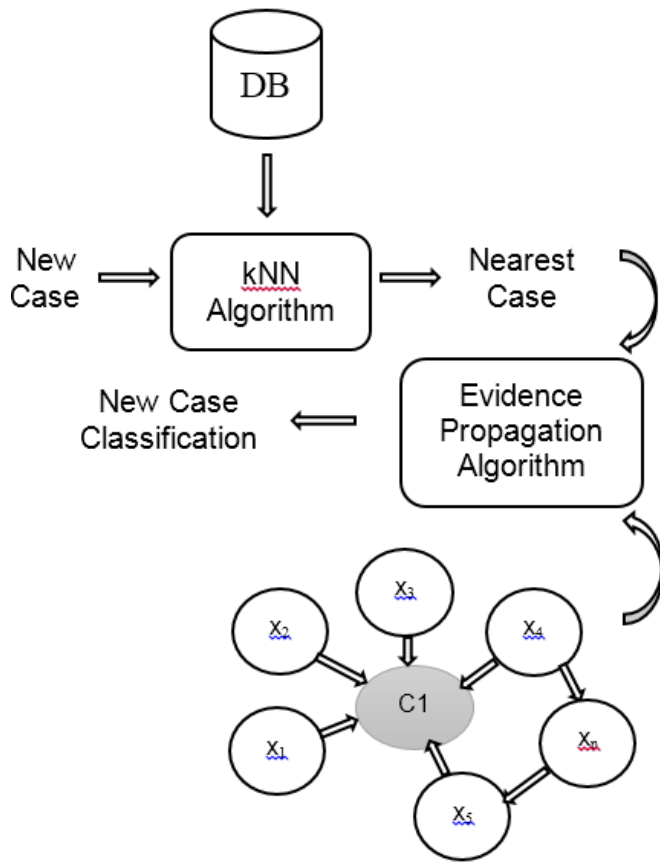
ეს ტექნიკა მიეკუთვნება მესამე (iii) ზემოთ აღწერილ მეთოდოლოგიას. ამგვარად, მონაცემებზე გამოიყენება Bayesian Network (BN-ის კერძო შემთხვევა - Naïve Bayes) და K Nearest Neighbor -ის ჰიბრიდული კლასიფიკატორი. Bayesian Network-ის სტრუქტურა იქმნება მონაცემებზე დაყრდნობით და Nearest Neighbor ალგორითმი გამოიყენება Bayesian Network-თან კომბინაციაში (Dedabrishvili, 2017).



ფიგურა 4. kNN-NB-ის ჰიბრიდული ალგორითმის ფსევდო-კოდი.

კომბინირებული ალგორითმის ეტაპები მოცემულია მე-4 ფიგურაში. ახალი ნიმუშები დასატრენინგებულ მონაცემებში კლასიფიცირდება უახლოესი შემთხვევის მიხედვით და საბოლოო გადაწყვეტილების მიღება ხდება წინასწარ განსწავლული Bayesian

Network-ის მიერ (by propagating the evidence). ახალი ნიმუშის კლასიფიკაციის სქემა მოცემულია მე-5 ფიგურაზე.



ფიგურა 5. ახალი შემთხვევის კლასიფიკაციის სქემა.

შესაჯამებლად, ახალი ობიექტის კლასიფიკაცია ხდება ამგვარად: პირველად, დასატრენინგებელ მონაცემებში KNN ალგორითმი ეძებს K უახლოეს მეზობელს. KNN-ის იმპლემენტაციის დროს მხედველობაში არ მიიღება კატეგორიული ატრიბუტები. მანძილის გაზომვა ხორციელდება მხოლოდ რიცხვითი ატრიბუტების გამოყენებით. K უახლოესი ობიექტის შერჩევის შემდეგ Naïve Bayes ალგორითმის მეშვეობით იქმნება მოდელი, სადაც ამ ეტაპზე მხოლოდ კატეგორიული ატრიბუტები განიხილება. ახალი ობიექტის კლასიფიკაციას კი ახდენს კომბინირებული მეთოდი. შესაბამისად, პროცესი ორ ეტაპიანია, სადაც პირველი ეტაპი მოიცავს მხოლოდ რიცხვითი ატრიბუტებით შერჩეულ უახლოეს მონაცემებს ახალი ობიექტისთვის. ეს ლოგიკურია, რადგან რიცხობრივად ახლო ობიექტებს ერთნაირი მახასიათებლები გააჩნიათ. მას შემდეგ, რაც მოპოვებულია K უახლოესი ობიექტი ახალი ობიექტისთვის,



მეორე ეტაპად, ნაცვლად ძირითადი კენჭისყის სქემისა, ისე როგორც ხდება  $k$ -NN-თან, კატეგორიული მონაცემების მახასიათებლები და მათი კლასებთან დამოკიდებულება მოხდება Naïve Bayes კლასიფიკაციის ტექნიკით. ამგვარად, ახალი ობიექტის კლასიფიკაციისთვის გამოიყენება ორივე, რიცხვითი და კატეგორიული ატრიბუტები და რაც მთავარია, არანაირი დისკრეტიზაციის ან კომპლექსური გაზომვები მსგავსების დასადგენად აღარ არის საჭირო.

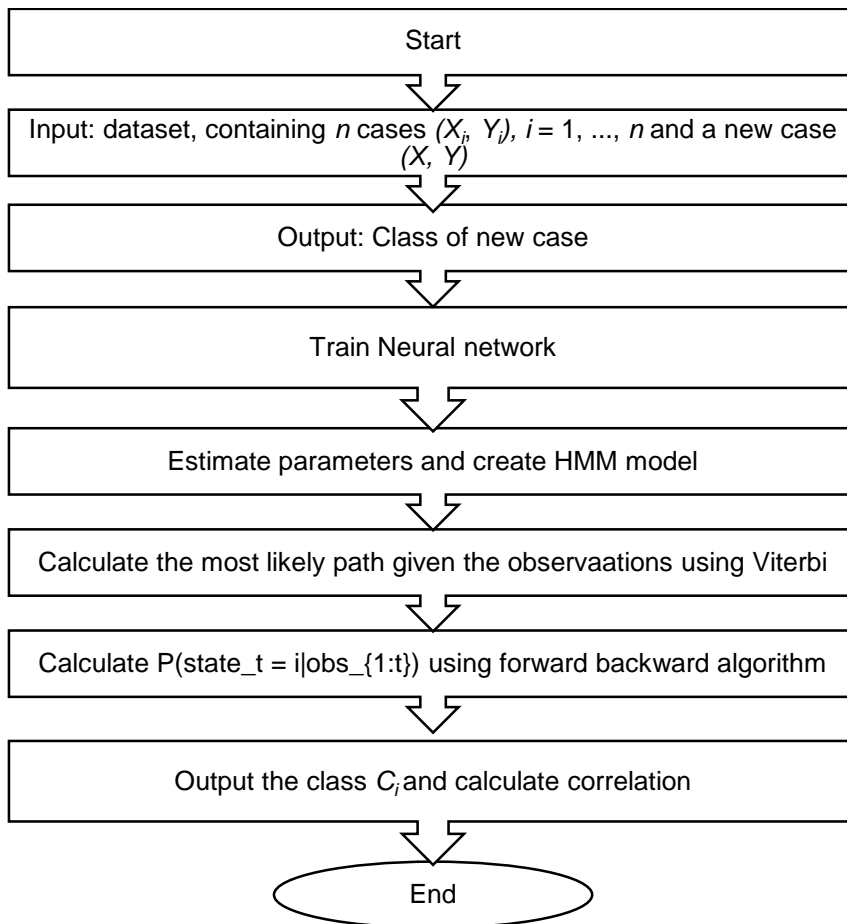
## პრობლემა 2: კონტროლირებადი და არაკონტროლირებადი ალგორითმების კომბინაციის მეთოდოლოგია

ცნობილი ფაქტია, რომ HMM გააჩნია შიდა შეზღუდვები, რაც ძირითადად გამოწვეულია მათი შემთხვევითობის პრინციპით განსაზღვრული პარამეტრების გამო (Trentin & Gori, 2003). ამასთან მიმართებში Artificial Neural Networks პერსპექტიული ალტერნატივაა. კვლევაში გამოყენებული ალგორითმების კომბინაცია ემყარება gradient-ascent მეთოდს ANN-HMM ჰიბრიდული სისტემის დატრენინგებისთვის, სადაც ANN ატრენინგებს HMM-ის ემისიის ალბათობას (emission probabilities) მდგომარეობის (state) შეფასებისათვის. მეთოდი ეფუძნება ძირითად ჰიბრიდულ სისტემას, რომელიც განავითარა ბურლარდმა, მორგანმა და ბენგიომ ალგორითმის გაერთიანების გზით მათი შეზღუდვების დასაძლევად.

გამოყენებული მეთოდი მოიცავს რამოდენიმე ფუნქციას (Palm, 2012):

```
hmmgenerate (Matlab function)
hmmest      (Estimates A (Transition Matrix) and PI)
hmmfbNN     (Forward backward algorithm hmm-nn hybrid)
hmmfbEMIS   (Forward backward algorithm hmm)
viterbiNN   (Find most probable path hmm-nn hybrid)
viterbiEMIS (Find most probable path hmm)
```

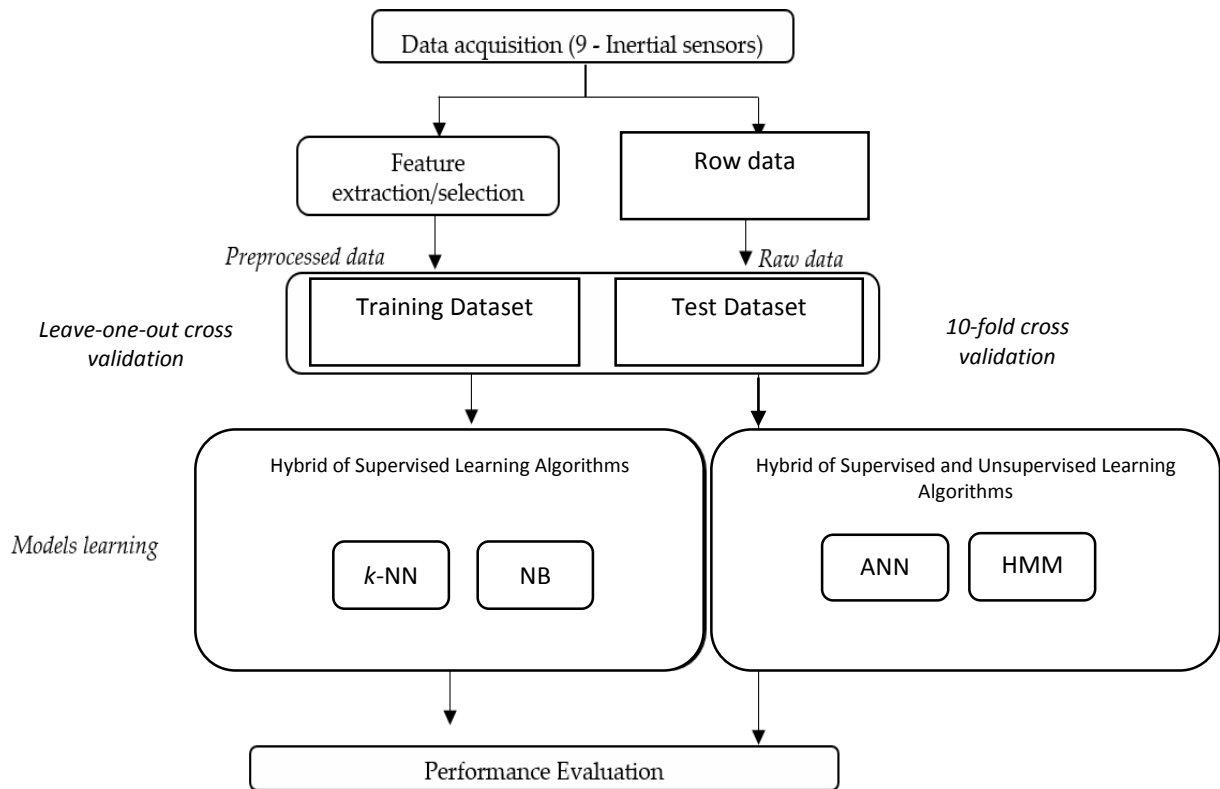
ალგორითმის სტრუქტურა გადმოცემულია მე-6 ფიგურაზე.



ფიგურა 6. NN-HMM ჰიბრიდული ალგორითმის ფსევდო-კოდი.

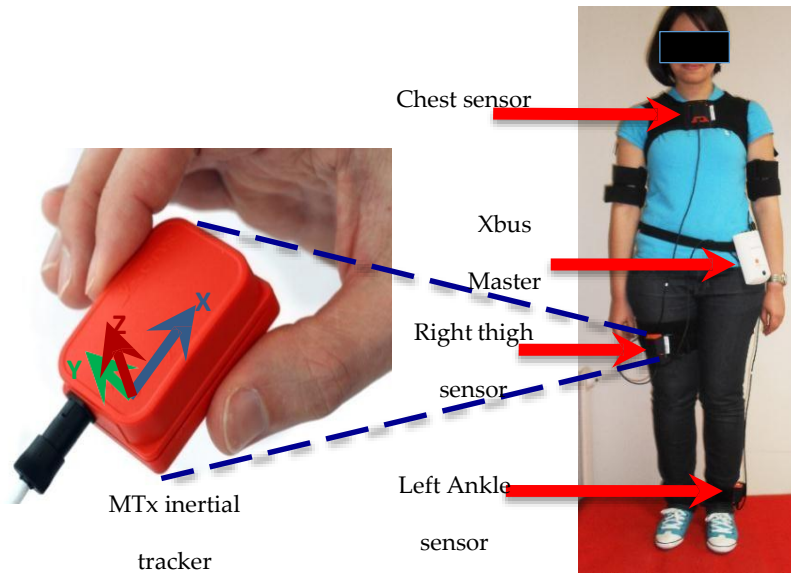
### თავი 3. შემუშავებული HAR-ის ტექნიკა

ამ ქვეთავში, წარმოდგენილია შემუშავებული მეთოდოლოგია, რაც მოიცავს მონაცემების შეგროვებას, კლასიფიკატორების გამოყენებას და ეფექტურობის შეფასებას 10-fold cross validation-ის ტექნიკით. ფიგურა 1 აჩვენებს კვლევის პირველი ნაწილის ეტაპებს ცალკეულად გამოყენებული ალგორითმებისთვის შემუშავებული მეთოდით, მაშინ როდესაც ფიგურა 7 წარმოადგენს სხვადასხვა ალგორითმების კომბინაციებს ეტაპობრივად, რომელიც მიეკუთვნება სწავლების მეორე ნაწილს.



ფიგურა 7. ადამიანის აქტივობის ამოცნობის ეტაპები ჰიბრიდული ალგორითმების მეშვეობით.

კვლევის ფარგლებში, ადამიანის აქტივობა ფასდება Xbus Kit from Xsens (Enschede, Netherlands) მოწყობილობის საშუალებით, რომელიც საშუალებას იძლევა ამბულატორიურ პირობებში მოვახდინოთ ადამიანის მოძრაობის ამოცნობა. ის მოიცავს მოძრავ სისტემას – Xbus Master –ის და სამი MTx ინერციული სენსორის შემადგენლობით, რომელიც დამაგრებულია სუბიექტის გულმკერდის არეში, მარჯვენა ბარძაყზე და მარცხენა კოჭზე, იხ. სურათი 8.



ფიგურა 8. MTx-Xbus ინერციული ტრეკერი და სენსორების განთავსება.

მონაცემები შეგროვდა ლისის ლაბორატორიაში, პარიზის უნივერსიტეტში (UPEC). ექსპერიმენტში მონაწილეობა მიიღო ექვსმა ჯანმრთელმა ინდივიდმა განსხვავებული პროფილით (საშუალო ასაკი –26 წელი, საშუალო წონა – 65 კგ.). სუბიექტებს მიეცათ ინსტრუქციები, რათა განახორციელონ აქტივობა კონკრეტული შეზღუდვების გარეშე. თითოეულმა სუბიექტმა სულ თორმეტი აქტივობა შეასრულა. მონაცემთა მოპოვება განხორციელდა ოფისის პირობებში, დაახლოებით 30 წუთის მანძილზე. აქტივობების ჩამონათვალი და განმარტება მოცემულია ცხრილში 1. მოპოვებული მონაცემები ხელით შეფასდა დამოუკიდებელი ოპერატორის მიერ (Attal, Mohammed, Dedabrishvili, & Chamroukhi, 2015).

ცხრილი 1. შერჩეული აქტივობების ჩამონათვალი (A1. . .A12).

Activity Reference	Description of Activity
A1	Stair descent
A2	Standing
A3	Sitting down
A4	Sitting
A5	From sitting to sitting on the ground
A6	Sitting on the ground
A7	Lying down
A8	Lying
A9	From lying to sitting on the ground
A10	Standing up
A11	Walking
A12	Stair ascent

### ექსპერიმენტული შედეგები

ამ ნაწილში განხილული და შედარებულია ყოველდღიური აქტივობების ამოსაცნობად გამოყენებული სტანდარტული კონტროლირებადი და არაკონტროლირებადი ML ალგორითმების ეფექტურობა. ეს შედარება წარმოაჩენს სხვადასხვა ალგორითმებს მათი average accuracy rate (R) + standard deviation (std), F-measure, recall, precision and specificity-ის თვალსაზრისით. ამ შედარებითი კვლევისას განიხილება რამოდენიმე შემთხვევა:

### შემთვევა 1: პირველადი მონაცემები

პირველად მონაცემებზე (row data) დაყრდნობით მოპოვებული შედეგები მოცემულია ცხრილში 2 და 3. კონტროლირებადი ალგორითმების ეფექტურობა გადმოცემულია მეორე ცხრილში. შეიძლება აღინიშნოს, რომ სხვადასხვა ტექნიკით მიღებული კლასიფიკაციის მაჩვენებლები აღემატება 84%-ს. *k*-NN ალგორითმი იძლევა საუკეთესო შედეგს global correct classification rate, F-measure, recall, და precision -ის გათვალისწინებით, შემდეგ მოდიან SVM და SLGMM ალგორითმები შედარებით დაბალი მაჩვენებლებით.

ცხრილი 3 გადმოსცემს არაკონტროლირებადი დასწალის ალგორითმებით მიღებულ შედეგებს.

ცხრილი 2. კონტროლირებადი ალგორითმების შედეგები პირველად მონაცემებზე.

	Accuracy $\pm$ std	F-measure	Recall	Precision	Specificity
k-NN (%)	96.53 $\pm$ 0.20	94.6	94.57	94.62	99.67
RF (%)	94.89 $\pm$ 0.57	82.87	82.28	83.46	99.43
SVM (%)	94.22 $\pm$ 0.28	90.66	90.98	90.33	99.56
SLGMM (%)	84.54 $\pm$ 0.30	69.94	69.99	69.88	98.39

ცხრილი 3. არაკონტროლირებადი ალგორითმების შედეგები პირველად მონაცემებზე.

	Accuracy $\pm$ std	F-measure	Recall	Precision	Specificity
HMM (%)	80.00 $\pm$ 2.10	67.67	65.02	66.15	97.68
K-means (%)	68.42 $\pm$ 5.05	49.89	48.67	48.55	93.21
GMM (%)	73.60 $\pm$ 2.32	57.68	57.54	58.82	96.45

k-NN-ის და HMM-ის შემთხვევებში, როგორც ამოსაცნობი ნიმუშების იდენტიფიცირებისთვის მე-4 და მე-5 ცხრილებში მოცემულია global confusion matrix. შეიძლება ითქვას, რომ უმეტეს წილად შეცდომები ხდება გარდამავალ აქტივობებს შორის, როგორებიცაა (A9, A7) და დინამიურ აქტივობებს შორის, როგორებიცაა (A1, A11), (A1, A12) და (A11, A12). ეს შეცდომები მეტად მნიშვნელოვანია HMM-ის შემთხვევაში. აგრეთვე აღსანიშნავია, რომ ძირითადი აქტივობები, როგორებიცაა A2, A4, A8 უფრო ადვილად ამოსაცნობია, ვიდრე A3, A5 და A7 გარდამავალი აქტივობები.

ცხრილი 4. Global confusion matrix მიღებული k-NN-ის მეშვეობით პირველად მონაცემებზე (on row data).

		Obtained Classes											
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
True Classes	A1	88.98	0.41	0.04	0	0.04	0	0	0	0	0.78	4.34	5.41
	A2	0.40	98.52	0.08	0	0	0	0	0	0	0.21	0.56	0.23
	A3	0.21	0.64	95.73	0.53	0.64	0	0	0	0	0.96	0.85	0.43
	A4	0	0	0.77	98.92	0.31	0	0	0	0	0	0	0
	A5	0.08	0	0.55	0.16	97.98	0.47	0.08	0	0.16	0.55	0	0
	A6	0	0	0	0	0.22	99.41	0.03	0	0.25	0.08	0	0
	A7	0	0	0	0	0.22	0.15	95.71	1.53	2.33	0.07	0	0
	A8	0	0	0	0	0	0	1.58	97.62	0.80	0	0	0
	A9	0	0	0	0	0.25	0.34	3.96	0.67	94.44	0.34	0	0
	A10	1.58	0.46	0.19	0	0.65	0.28	0	0	0.19	94.07	0.93	1.67
	A11	4.07	0.41	0.03	0	0	0	0	0	0	0.55	92.57	2.37
	A12	5.05	0.43	0	0	0	0	0	0	0	1.03	3.08	90.42

ცხრილი 5. Global confusion matrix მიღებული HMM-ის მეშვეობით პირველად მონაცემებზე (on row data).

		Obtained Classes											
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
True Classes	A1	55.33	1.70	1.08	0	0.62	0	0	0	0	3.19	23.52	14.57
	A2	2.83	86.22	0.47	0	0	0	0	0	0	1.50	6.97	2.01
	A3	0.12	0	39.86	32.82	12.53	0	0	0	0	10.62	0.24	3.82
	A4	0.10	0	9.58	87.21	3.11	0	0	0	0	0	0	0
	A5	0.67	0	7.20	0.29	73.61	0.10	1.06	0	1.44	15.55	0	0.10
	A6	0	0	0	0	3.15	91.63	0.88	0	2.18	2.16	0	0
	A7	0	0	0	0	2.24	0.50	29.74	35.33	27.95	4.25	0	0
	A8	0	0	0	0	0	0	13.14	81.38	5.48		0	0
	A9	0	0	0	0	2.13	0	37.03	16.70	33.75	10.39	0	0
	A10	0	0	0	0	9.20	0	0	0	1.15	89.66	0	0
	A11	19.59	1.38	2.53	0	0	0	0	0	0	2.38	56.95	17.17
	A12	16.65	0	3.72	0	2.44	0	0	0	0	5.75	11.10	60.34

**შემთვევა 2: პირველად მონაცემებზე მოპოვებული და შერჩეული მახასიათებლები**

ზემოთ წარდმოგენილი მიღებული შედეგების გაუმჯობესების მიზნით განხორციელდა წინადადებების ეტაპი, რომელიც მოიცავს მახასიათებლების მოპოვებას და შერჩევას. სამი MTx IMUs სენსორიდან მოპოვებულია აქსელერომეტრის ცხრა სიგნალზე დაყრდნობით გამოთვლილია შემდეგი დროისა და სიხშირის მახასიათებლები:

- თერთმეტი დროის–დომენი მახასიათებელი, სახელდობრ: mean, variance, median, interquartile rang, skewedness, kurtosis, root mean square, zero crossing, peak to peak, crest factor and rang.
- ექვსი სიხშირის–დომენი მახასიათებელი, მათ შორის: DC component in FFT spectrum, energy spectrum, entropy spectrum, sum of the wavelet coefficients, squared sum of the wavelet coefficients და energy of the wavelet coefficients.

ცხრილი 6. კონტროლირებადი ალგორითმების შედეგები მოპოვებული

მახასიათებლების მქონე მონაცემებზე.

	Accuracy $\pm$ std	F-measure	Recall	Precision	Specificity
k-NN (%)	99.25 $\pm$ 0.17	98.85	98.85	98.85	99.96
RF (%)	98.95 $\pm$ 0.09	98.27	98.24	98.25	99.90
SVM (%)	95.55 $\pm$ 0.30	93.02	93.15	92.90	99.92
SLGMM (%)	85.05 $\pm$ 0.57	73.44	74.44	73.61	99.88

ცხრილი 7. არაკონტროლირებადი ალგორითმების შედეგები

მოპოვებული მახასიათებლების მქონე მონაცემებზე.

	Accuracy $\pm$ std	F-measure	Recall	Precision	Specificity
HMM (%)	83.89 $\pm$ 1.30	69.19	68.27	67.74	98.38
K-means (%)	72.95 $\pm$ 2.80	50.29	52.20	51.22	97.04
GMM (%)	75.60 $\pm$ 1.25	65.00	66.29	64.30	97.12

მე-8 და მე-9 ცხრილები წარმოადგენენ confusion matrix-ებს k-NN-ის და HMM-ის შემთხვევაში მოპოვებული მახასიათებლების გამოყენებით.

ცხრილი 8. Global confusion matrix მიღებული k-NN-ის მეშვეობით შერჩეული მახასიათებლების მქონე მონაცემებზე (using selected features).

		Obtained Classes											
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
True Classes	A1	99.00	0.32	0	0	0	0	0	0	0	0.08	0.48	0.12
	A2	0.06	99.75	0.04	0	0	0	0	0	0	0.03	0.07	0.04
	A3	0	0.43	99.15	0.43	0	0	0	0	0	0	0	0
	A4	0	0	0.11	99.79	0.11	0	0	0	0	0	0	0
	A5	0	0	0	0.23	99.38	0.23	0	0	0.08	0.08	0	0
	A6	0	0	0	0	0.07	99.78	0.07	0.03	0.05	0	0	0
	A7	0	0	0	0	0	0.21	99.65	0.14	0	0	0	0
	A8	0	0	0	0	0	0.15	99.79	0.06	0	0	0	0
	A9	0	0	0	0	0.08	0.17	0.33	99.42	0	0	0	0
	A10	0.35	0.18	0	0	0.09	0.09	0	0	0	99.20	0.09	0
	A11	0.22	0.17	0	0	0	0	0	0	0	0	99.34	0.28
	A12	0.08	0.17	0	0	0	0	0	0	0	0.04	0.25	99.45

Table 9. Global confusion matrix მიღებული HMM-ის მეშვეობით შერჩეული მახასიათებლების მქონე მონაცემებზე (using selected features).

		Obtained Classes											
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
True Classes	A1	57.74	0.06	0.43	0	0.31	0	0	0	0	4.07	20.17	17.21
	A2	1.36	94.66	0.31	0	0	0	0	0	0	0.89	1.98	0.80
	A3	3.82	0	55.30	5.69	15.42	0	0	0	0	1.64	4.91	13.24
	A4	0	0	2.85	96.31	0.83	0	0	0	0	0	0	0
	A5	2.05	0	1.80	0.66	71.62	4.35	2.21	0	5.50	11.48	0	0.33
	A6	0	0	0	0	1.39	97.09	0.30	0	0.94	0.28	0	0
	A7	0	0	0	0	1.54	0	59.91	4.25	32.30	1.99	0	0
	A8	0	0	0	0	0	0	3.30	94.69	2.01	0	0	0
	A9	0	0	0	0	4.02	1.75	32.68	0.10	50.41	11.03	0	0
	A10	13.56	0	1.51	0	6.44	0	1.92	0	2.19	60.68	7.12	6.58
	A11	19.87	4.45	1.50	0	0	0	0	0	0	3.45	57.02	13.73
	A12	16.37	0.17	0	0	0.34	0	0	0	0	1.90	17.26	63.97



**შემთხვევა 3: ექსპერიმენტული შედეგები კონტროლირებადი დასწავლის ალგორითმების კომბინაციის გამოყენებით**

ამ ქვეთავში წარმოდგენილია ექსპერიმენტული შედეგების მიმოხილვა და შედარება, რომელიც განხორციელდა მონაცემებზე მულტი კლასიფიკატორის ან იგივე ჰიბრიდული კლასიფიკატორის NB-სა და  $k$ -NN მეშვეობით.

მონაცემების შესწავლისას, ახალი შემთხვევების კლასიფიკაცია ხდება შემდეგი თანმიმდევრობით:

(i) პირველ რიგში, იძებნება უახლოესი მეზობელი დასატრენინგებელ მონაცემთა ბაზაში  $k$ -NN ალგორითმის საშუალებით, სადაც  $K_i$  წარმოადგენს უახლოეს შემთხვევას,

(ii) შემდგომ,  $K_i$  შემთხვევის გადაცემა წინასწარ ნასწავლები BN-ისთვის, როგორც ახალი შემთხვევისა,

(iii) და ბოლოს, ახალი შემთხვევის დასათაურება შესაბამისი კლასის მნიშვნელობით, რომელიც მიიღწევა Bayesian network-ის ორი ქვე-მიზნით: ქსელის სტრუქტურის დადგენით და თითოეული კვანძისთვის ალბათობის ცხრილის შექმნით (Dedabrishvili, 2017).

ექსპერიმენტის შედეგები მოცემულია ცხრილში 10. მონაცემებმა, როგორც ზემოთ არის აღწერილი, გაიარა წინა-დამუშავების ფაზა (preprocessing) და მისი განზომილება შემცირებულია Principal Component Analysis-ის მეშვეობით (Attal, Mohammed, Dedabrishvili, & Chamroukhi, 2015).

*ცხრილი 10. ცალკეულად და კომბინირებულად ( $k$ -NN-NB) გამოყენებული ალგორითმების შედეგები მოპოვებული მახასიათებლების მქონე მონაცემებზე.*

(%)	Accuracy	Error Rate	Precision	Recall
kNN	0.99253	0.00747	0.98851	0.98851
NB	0.94286	0.05714	0.94286	0.95887
kNN-NB	<b>0.99526</b>	0.00474	0.99526	0.99527

**შემთხვევა 4: ექსპერიმენტული შედეგები კონტროლირებადი და**

**არაკონტროლირებადი დასწავლის ალგორითმების კომბინაციის გამოყენებით**

აქ წარმოდგენილია კონტროლირებადი და არაკონტროლირებადი ალგორითმების კომბინაციით მიღებული შედეგები. ექსპერიმენტი განხორციელდა ANN და HMM

ალგორითმების გამოყენებით HAR მონაცემთა ნაკრებზე (dataset).

მონაცემების შესწავლის პროცესში, ახალი შემთხვევების კლასიფიკაცია ხდება შემდეგნაირად:

- (i) უპირველეს ყოვლისა Neural Networks –ის დატრენინგებით,
- (ii) შემდგომ, პარამეტრების განსაზღვრით (Transition Matrix, Emission Matrix and PI) და HMM მოდელის შექმნით,
- (iii) ამის შემდეგ, მეტად სავარაუდო გზის გამომანგარიშება Viterbi ალგორითმის დაკვირვების საფუძველზე,
- (iv) და ბოლოს, ალბათობის გამოთვლით  $P(\text{state}_t = i | \text{obs}_{1:t})$  Forward Backward ალგორითმის გამოყენებით.

*ცხრილი 11. ცალკეულად და კომბინირებულად (NN-HMM) გამოყენებული ალგორითმების შედეგები პირველად მონაცემებზე.*

Class	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12	Average %
HMM	75.334	96.325	69.458	87.312	73.819	91.836	69.754	94.387	73.275	89.686	62.195	80.034	80.28458
NN	88.017	93.738	75.014	89.577	85.817	86.959	77.055	92.123	87.601	88.677	92.038	94.599	87.60125
HMMNN	91.023	95.028	85.012	93.031	87.702	88.021	86.055	90.327	85.706	86.857	90.038	92.001	<b>89.23342</b>

ექსპერიმენტები ცხადყოფს, რომ ზოგადად კონტროლირებადი დასწავლის ალგორითმები უსწრებს არაკონტროლირებადი დასწავლის მიდგომებს, თუმცა ამ უკანასკნელის შედეგები საიმედოა აღწერილ სასწავლო პირობებში. აღსანიშნავია, ისიც, რომ სწავლის ფარგლებში ჰიბრიდული ალგორითმები უზრუნველყოფენ უკეთეს სიზუსტის მაჩვენებელს ვიდრე ცალკეული ტექნიკები, რაც შეიძლება აიხსნას  $k$ -NN-NB კომბინაციით მიღწეული გაუმჯობესება 0.3%-ით და NN-HMM ჰიბრიდის შემთხვევაში კი 9%-ით.

## დასკვნა

ძირითადი მეთოდური და ღირებული მიღწევები დისერტაციაში:

- სწავლის ფარგლებში გამოყენებული ცნობილი კონტროლირებადი ალგორითმებიდან HAR-ისთვის ერთ-ერთი საუკეთესო კლასიფიკატორის დადგენა.

- სწავლის ფარგლებში გამოყენებული ცნობილი არაკონტროლირებადი ალგორითმებიდან HAR-ისთვის ერთ-ერთი საუკეთესო კლასიფიკატორის დადგენა.
- შემუშავდა ალგორითმების ახალი ჰიბრიდები მონაცემთა კლასიფიკაციისათვის, რომლებიც აუმჯობესებენ ცალკეულად გამოყენებულ კლასიფიკატორთა შედეგებს HAR-ის სფეროში.
  - შემუშავდა Instance-based ალგორითმისა და Naïve Bayes-ის კომბინირებული კლასიფიკატორი.
  - შემუშავდა Hidden Markov Model-ისა და Artificial Neural Networks – ის კომბინირებული კლასიფიკატორი
- კვლევებმა ცხადყო, რომ ერთი ალგორითმის სიძლიერე შეიძლება გამოყენებულ იქნას სხვა ალგორითმის სუსტი მხარის ჩასანაცვლებლად.
- კვლევის შედეგებმა აჩვენა, რომ სხვადასხვა კლასიფიკაციის ალგორითმებში დატრენინგებული მონაცემების გაერთიანებით იზრდება სიზუსტის მაჩვენებელი.
- სამომავლოდ შემუშავდა ჰიბრიდული კლასიფიკატორის ფუნქციის შექმნის გზა სხვა მომხმარებლებისთვის მომიჯნავე სფეროებში გამოსაყენებლად.

### **შესაძლო მიმართულებები და რეკომენდაციები შემდგომი სწავლისთვის**

- სწავლის განვრცობა შესაძლებელია სხვადასხვა მანქანური სწავლების მეთოდით როგორც კონტროლირებად ისე არაკონტროლირებად კონტექსტებში ფოკუს ჯგუფის ფართო მაჩვენებელზე, რაც ნიშნავს მონაცემების გაზრდას სხვა მონაწილეების დამატებით ზოგადად და კონკრეტულად ხანდაზმული სუბიექტების ჩართულობით.
- სენსორების მოსახერხებელი გამოყენებისათვის (ხელმისაწვდომობისთვის) მოგვიანებით ჩატარებულ კვლევებში შეიძლება ჩაირთას ჭკვიანი სატელეფონო შესაძლებლობები მომხმარებლის მოძრაობის შესწავლის მიზნით.
- ექსპერიმენტები შესაძლებელია გაგრძელდეს კლასიფიკაციის სხვა ალგორითმებით და მათი კომბინაციებით როგორც პირველად ისე მოპოვებული მახასიათებლების მქონე მონაცემებზე.

- შესაძლებელია ჰიბრიდული კლასიფიკატორების ფუნქციების შექმნა Matlab-ში, რომლებიც გაუადვილებს მომხმარებლებს კლასიფიკაციას.
- დისტანციური მონიტორინგის პროცესში შეიძლება ჩართულ იქნას მობილური აპლიკაციები არანორმალური სიტუაციების დასადგენად და გადაუდებელი დახმარების აღმოსაჩენად.

### გამოქვეყნებული სტატიები

1. Dedabrishvili, M. (2017). Study of Activity Recognition Dataset Using Combined Probabilistic and Instance Based Algorithms. *Journal of Technical Science & Technologies; Volume 6, Issue 2, 27-33.*
2. Attal, F., Mohammed, S., Dedabrishvili, M., & Chamroukhi, F. (2015). Physical Human Activity Recognition Using Wearable Sensors. *Sensors*, 31314-31338.
3. Dedabrishvili, M., Rodonaia, I. (2018). Hybrid of Supervised and Unsupervised Learning Algorithms for Human Activity Recognition. *Transactions Automated Control Systems, 1(25)*, 31-40.