INTERNATIONAL BLACK SEA UNIVERSITY
FACULTY of COMPUTER TECHNOLOGIES and ENGINEERING
Computer Sciences, Doctoral PROGRAM

Human Daily Physical Activity Recognition Using Body Mounted Sensors

MARIAM DEDABRISHVILI

Extended Abstract of Doctoral Dissertation of Engineering in Informatics

Tbilisi, 2018

Scientific Supervisor: Irakli Rodonaia

Professor, Doctor, at International Black Sea University

(supervisor's signature)

Scientific Co-supervisor: Alexander Milnikov

Doctor

(co-supervisor's signature)

Experts (full name & academic title):

1. Prof. Dr. Nodar Momtselidze
2. Prof. Dr. Besik Dundua
3. Assoc. Prof. Dr. George Mandaria

Opponents (full name & academic title):

1. Prof. Dr. Diane Cook
2. Assist. Prof. Dr. Daniela Micucci
3. Assoc. Prof. Dr. Khatuna Bardavelidze

Introduction

This work will deal with the classification of daily living human activities using wearable inertial sensors. Walking, Lying, Standing up, etc. are examples of these activities. In this study, a dataset including 12 activities namely: stair descent, standing, sitting down, sitting, from sitting to sitting on the ground, sitting on the ground, lying down, lying, from lying to sitting on the ground, standing up, walking and stair ascent, is created using three inertial sensors. Four supervised classification techniques namely, k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), Supervised Learning Gaussian Mixture Models (SLGMM) and Random Forest (RF) as well as three unsupervised classification techniques namely, k-Means, Gaussian Mixture Models (GMM) and Hidden Markov Model (HMM), will be compared in terms of correct classification rate, F-measure, recall, precision, and specificity. Raw data and extracted features will be used separately as inputs of each classifier. The inertial sensor units worn by different healthy subjects are placed at key points of upper/lower body limbs (chest, right thigh and left ankle). The activity recognition process includes three main steps: sensors' placement, data pre-processing and data classification. In this study, only acceleration data is used, as a modality for estimating the activities (Altun, Barshan, & Tunçel, 2010; Chamroukhi, Mohammed, Trabelsi, Oukhellou, & Amirat, 2013). The results obtained with supervised and unsupervised classification algorithms will be provided and analyzed. Furthermore, there will elaborated hybrid techniques to enhance the performance evaluation of the classification algorithms.

Problem Statement

Development of reliable and precise methods of Human Activity Recognition (HAR) are highly important, since wrong or inaccurate recognition can cause harmful consequences for human health.

Scientists working in the field try to find the ways to enhance achievements for recognition accuracy. Taking this into account, it is vital to choose classifiers which make classification of activities with reliable rates. However, limitations of the currently existing algorithms and inherent lack of precision level can put their applicability to the field under the question. In particular, Hidden Markov Model has certain restrictions, which is caused by the principle of random selection of parameters and it is problematic to discriminate between the classes with high accuracy.

Other well-known methods also have some drawbacks due to their nature. Therefore, to solve these problems and to ensure the required results, as well as taking into account the recommendations of well-known researchers from Lissi Lab, Paris, France, when developing the dissertation methodology it was decided to create a hybrid complex of classifiers that provide the improved accuracy and adequacy of models of the study area.

As a result, in the thesis a set of methods and algorithms was proposed and developed, which successfully solves the problems posed. In particular, the developed hybrid of the algorithms is more natural and applicable to the HAR framework as compared to other existing so far methods. Results and achievements of the part of the thesis were published in the scientific journal "Sensors", which has a high impact factor. The paper deserved great interest of scientific society which can be approved by the large number of the citations. Taking into account the interest caused by the article, it was decided to continue the development of the proposed methodology with the purpose of justifying the choice of a combination of known classification methods and aimed at a significant increase in the adequacy and accuracy of selected recognizable human daily actions.

Goal Statement

The main idea of the research is to find out the most suitable classifiers for Human Daily Physical Activity Recognition in supervised and unsupervised learning environments.

Twelve basic activities of daily lives of human will be selected to fulfill research goals. Data will be collected using wearable technologies, particularly using wearable inertial sensors that are mounted on human body.

Three main steps that describe the activity recognition process: sensors' placement, data pre-processing and data classification will be investigated. Four supervised classification techniques namely, k-Nearest Neighbor (k-NN), Support Vector Machines (SVM), Gaussian Mixture Models (GMM), and Random Forest (RF) as well as three unsupervised classification techniques namely, k-Means, Gaussian mixture models (GMM) and Hidden Markov Model (HMM), will be compared in terms of correct classification rate, F-measure, recall, precision, and specificity. Raw data and extracted features will be used separately as inputs of each classifier. In addition, data will be studied through the combination of supervised and unsupervised classification techniques.

Originality Perspectives

- ✓ The way of the sensor placement – in other studies there are different combinations
- ✓ Selection of set of activities and their order
- ✓ The way data has been filtered and reduced
- ✓ Number of features found
- ✓ Selection of most popular algorithms in HAR based on other studies' achievements

- ✓ Finding out best classifier among selected algorithms in Supervised Learning Environment according to accuracy rate
- ✓ Finding out best classifier among selected algorithms in Unsupervised Learning Environment according to accuracy rate
- ✓ Outperforming overall accuracy rate in the special environment of selected sensor placement, selected activities and employed algorithms by 9% for instance, in case of NN-HMM hybrid, which is important improvement
- ✓ Application of combinations of algorithms having different natures that have not been applied to the field before to our knowledge

Research Objectives

Research objectives of this thesis is to ascertain best classifier for HAR which will outperform other achivements in the same field and will enable users to recognize human dayly activities with higher accuracy.

Novelty and Actuality

Based on the investigation and research achievements novelty and actuality of the thesis can be explained by given factors:

- ✓ Has been found out the one of the most suitable classifier from study employed popular supervised learning algorithms in HAR.
- ✓ Has been found out the one of the most suitable classifier from study employed popular unsupervised learning algorithms in HAR.
- ✓ Has been developed new hybrids of classifiers for data classification which outperform solely application of the algorithms in the field of HAR.
    a. Has been proposed original way of combination of Instance-based algorithms with Naïve Bayes approach.
    b. Has been proposed a way of combination of Hidden Markov Model with Artificial Neural Networks.
- ✓ Has been shown that strengths of the single algorithm can be used as a part of another algorithm with weaknesses at the similar point.
- ✓ Has been shown that by concatenation of trained data from different classification algorithms, the overall accuracy rate can be increased.

- ✓ Has been developed a way to create a Function in the future of the hybrid classifiers for other users to apply in neighboring areas of the field straightforwardly.

Significance of the Problem

The problem of recognition of human activities with high accuracy rate is highly important due to sensitivity of consequences, which can couse harmful effects for human health.

Practical and Theoretical Value

Selection of the algorithms for data classification and creation of hybrid classifiers which increases the accurary rate of the recognition process of the human activities leads to the possibility to prolong the authonomy and well being of patiences suffering from different deseseas as well as elderly population living alone whose number nowadays is considerably raised.

Research Methods

In the conducted research, human activities are assessed utilizing the Xbus Kit from Xsens (Enschede, Netherlands) which empowers ambulatory measurement of the human movement. It comprises of a portable framework that incorporates a Xbus Master and three MTx inertial units that are attached on the chest, the right thigh and the left ankle of the subject.

Structure of the Dissertation

Dissertation consists of Introduction, Review of the Literature, Theoretical part, and Practical implementation, Conclusion, 39 Figures, 21 Tables and 181 References.

Basic content of research

Chapter 1. Research Goals and Literature Review

In Chapter 1, critical analysis of human activity recognition methods are conducted. The investigation revealed that comparing algorithm performance across different studies is a difficult task for many reasons. This difficulty is mainly related to: (i) the variability in the experimental protocols (the number of recruited subjects, the nature and the number of the recognized activities—ambulation, transportation, daily activities, exercise/fitness—the duration and the order of different activities, etc.); (ii) the applicative objectives behind the human activity recognition (monitoring, fall detection, home-based rehabilitation, etc.); (iii) the type of sensors used (accelerometers, plantar pressure sensors, gyroscopes) and their attachment to the body (wrist, chest, hip, thigh, necklace); (iv) the sort of data mining techniques (feature extraction, selection, data reduction, classification) used in the studies; (v) the performance evaluation criteria (accuracy, F-measure, recall, precision, specificity, etc.); the validation procedure (P-fold, leave one out, repeated random sub-sampling, bootstrap, etc.). By analyzing the circumstances and limitations of the current strategies for human activity recognition, and taking these issues into consideration by the aim to find the optimal way to reach high accuracy levels, the following research objectives have been established:

Study Objectives

- To find out the finest classifier from popular supervised learning algorithms in HAR.
- To find out the finest classifier from popular unsupervised learning algorithms in HAR.
- To develop new hybrids of classifiers for data classification which outperform solely application of the algorithms in the field of HAR.
    - To propose original way of combination of Instance-based algorithms with Naïve Bayes approach.
    - To propose a way of combination of Hidden Markov Model with Artificial Neural Networks.
- To show that strengths of the single algorithm can be used as a part of another algorithm with weaknesses at the similar point.
- To show that by concatenation of trained data from different classification algorithms, the overall accuracy rate can be increased.
- To develop a way to create a Function in the future of the hybrid classifiers for other users to apply in neighboring areas of the field straightforwardly.

*Figure 1*summarizes the different steps of the first part of the adopted approach which will be computed and evaluated step-by-step during the study process.

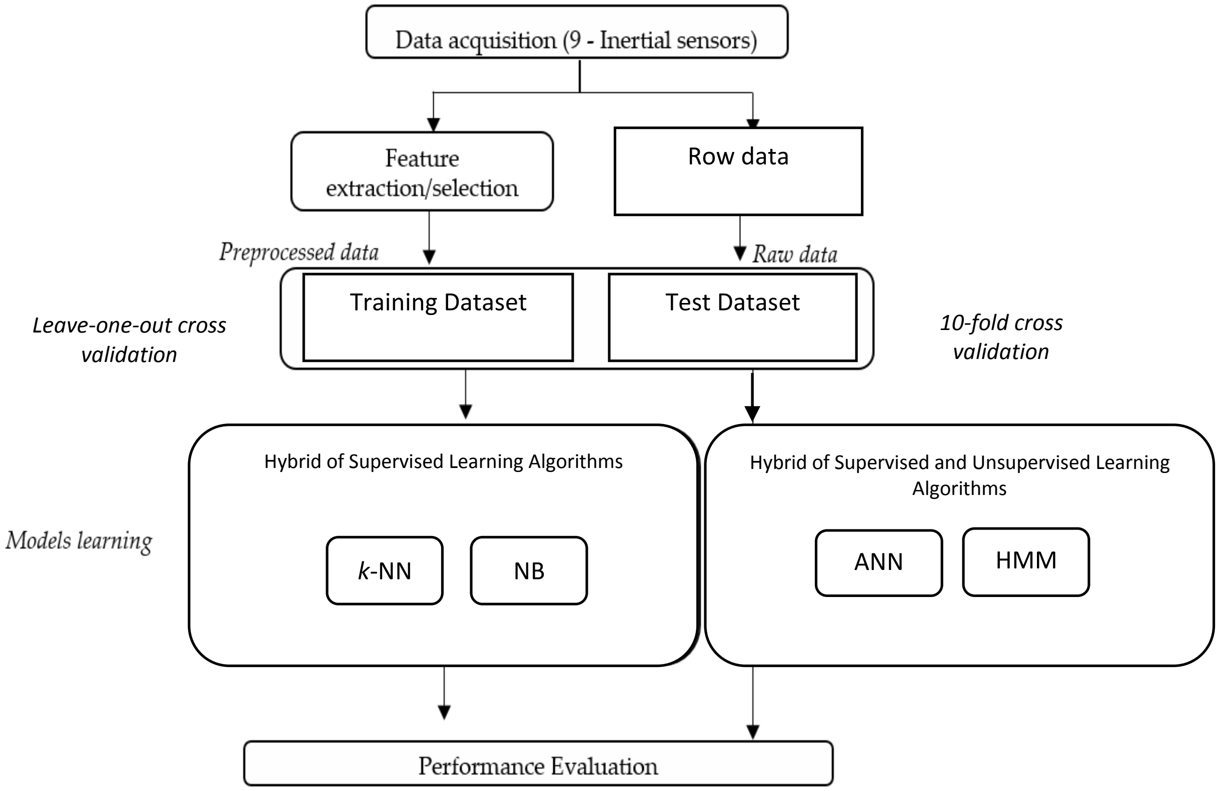

Figure 1. Steps of human activity recognition using different algorithms.

Chapter 2. Theoretical Models and Problem Solving Solutions

In the second chapter, the classification techniques used in this study for human activity recognition (GMMs, k-Nearest Neighbors (k-NN), Naïve Bayes, SVMs, Random Forests (RFs), K-means, ANN and HMMs) are briefly described and analyzed in terms of their characteristics for practical implementation which leads to the decision to combine different algorithms for performance enhancement purposes.

General issues of Supervised and Unsupervised learning algorithms

Predictive data mining is the most significant application for Machine Learning (ML) field. Instances that are used by the ML algorithms are represented by means of features. Extracted/selected features from the raw sensor data are used as inputs of the classification algorithms. In case of human activity recognition, the patterns of input data are associated with the activities (classes) under consideration. In general, the classification task requires learning a decision rule or a function associating the inputs data to the classes (Duda, Hart, & Stork, 1999; Webb, 2003; Theodoridis, Pikrakis, Koutroumbas, & Cavouras, 2010). Categorization of instances can be explained by two ways: the one, which is labeled or having corresponding correct outputs under the supervision of subjects and is known as Supervised, and another, Unsupervised learning, where instances are unlabeled and classification is carried out using different predictive methods (Jain AK, 1999).

In supervised learning, based on labeled data the algorithms are trained with predefined concepts and functions (Zoila Ruiz, 2017). Supervised learning methods try to discover the relationship between input attributes (i.e. independent variables) and a target attribute (i.e. dependent variable). The obtained relationship represents the structure which is denoted to as a model. Discovered models generally define and explain phenomena, which are out of sight in the dataset and can be useful for prediction of the value of the dependent variable while the independent variable values are known (Maimon & Rokach, 2005). The process of learning a set of rules from instances (examples in a training set) is known to as inductive machine learning, or in other words, building a classifier that can be utilized to generalize from new instances (Kotsiantis, Zaharakis, & Pintelas, 2007).

In unsupervised learning, algorithms have to find out interesting properties of the given a set of instances (Attal, Mohammed, Dedabrishvili, & Chamroukhi, 2015). Unsupervised learning task is to find out how systems can learn to show particular input patterns in a way that replicates the statistical structure of the whole set of input patterns. Compared to Supervised Learning or Reinforcement Learning, there are no obvious target outputs nor environmental evaluations related with each input; rather the unsupervised learner holds prior biases as to what characteristics of the structure of the input ought to be captured in the output. Unsupervised learning is more common to human brain structure and thus it is important.

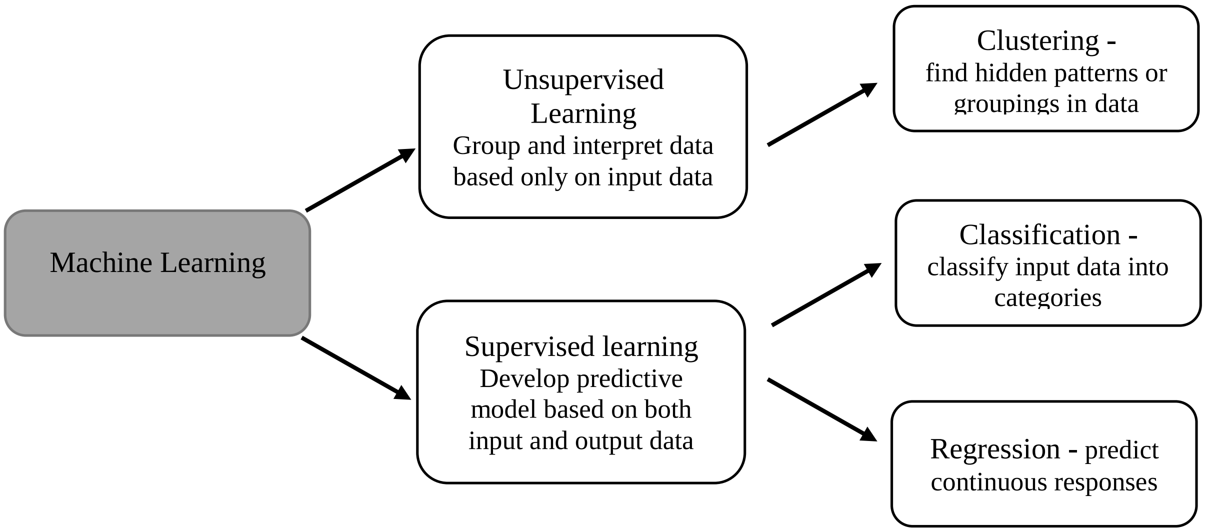

Figure 2. Machine Learning Techniques.

For instance, Artificial Neural Networks (Webb, 2003) and Support Vector Machines (SVM) (Vapnik, 2000), represent supervised learning approaches for classification and require entirely labeled activity data. Whereas The unsupervised learning approaches, such as those based on Gaussian Mixture Models (GMMs) and Hidden Markov Models (HMMs) (Rabiner, 1989) allow to infer automatically the labels from the data.

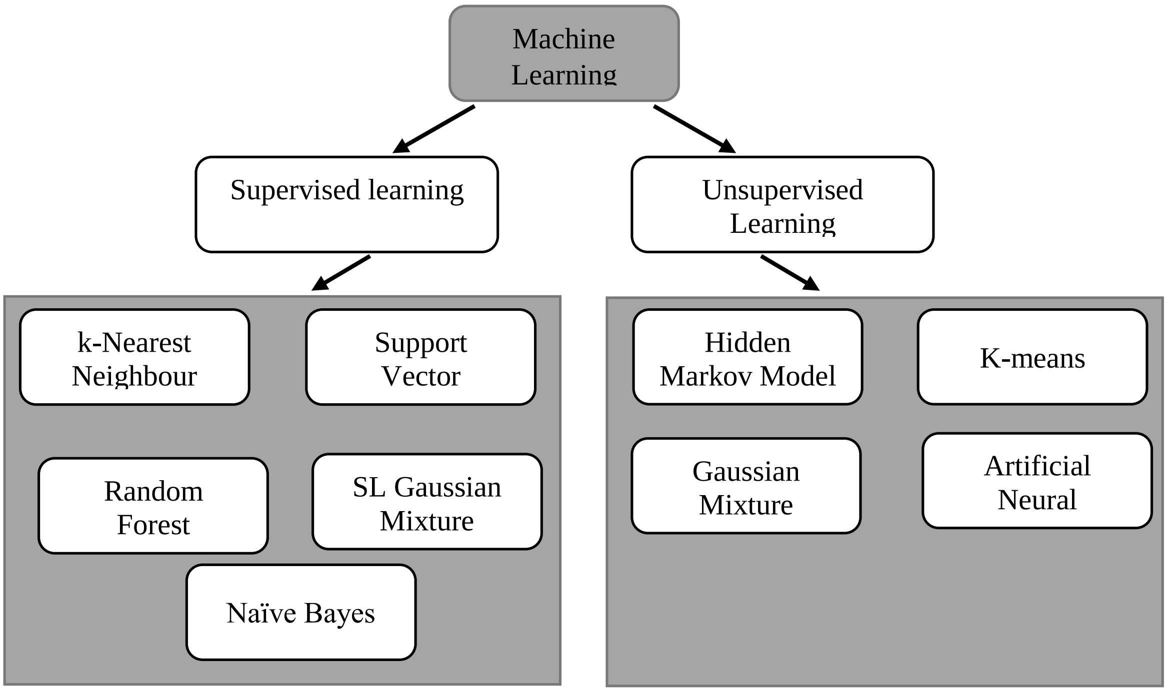

Figure 3. List of the selected algorithms in the study.

In overall, this chapter of the thesis presents the overview of the classification techniques for human activity recognition on theoretical bases, which are lately deployed in the study. General issues of the supervised and unsupervised learning techniques together with separate algorithm description are discussed and summarized. As far as algorithms are having weaknesses and strengths over others, combination of the instance-based algorithms with Naïve Bayes algorithm is proposed on one hand, and on the other hand – ANN with HMM hybrid is selected for the recognition of human activities. Application of the different algorithms to the dataset shows that in given case, separate usage of the algorithms give lower performance than hybrid of the those techniques. Combination of the algorithms which coincides strengths of the one algorithms to overcome weakness of the other one, leads to the better achievements.

Problem Statement 1: Supervised Algorithm Combination Methodology

There are various methods suggested for the creation of ensemble of classifiers (Sergey Tulyakov, 2008). Even though one can find number of proposed techniques of ensemble creation, there is as yet no clear picture of which technique is finest (Villada & Drissi, 2002). Consequently, construction of good combination of classifiers is an active area of research in supervised learning. There are three main methodologies to build an ensemble of classifiers: (i) by means of different subsets of training data with a particular learning technique, (ii) by means of different training parameters with a particular training technique (e.g., using different initial weights) and (iii) by means of different learning techniques. While combining classifiers complementary information can be gained by fusing the different sources. All those described combinations can produce appreciable improvements (Lazkano & Sierra, 2003; Sierra, Lazkano, Martinez-Otzeta, & Astigarraga, 2003).

The statistical based algorithm Naïve Bayes Classifier and distance-based algorithm K-Nearest Neighbor are often used in prediction problem (Ferdousy, Islam, & Matin, 2013). With regard to Naïve Bayes algorithm one of the key factors is to work with numerical attributes. It is understandable since in the algorithm one must define the conditional probability for each possible value of all attributes. To solve this problem, numerical attributes have to be discretized into numerous classes by adopting a discretization method from a variety of options available. Thus, the technique selected for discretization plays an essential role over the accuracy of the method. Several attempts have been carried out with the aim to increase the accuracy of the Naïve Bayes algorithm by adopting new discretization structure (Yang & Webb, 2002).

The situation is quite contrasting in case of K Nearest Neighbor algorithm. Here the issue is about categorical attributes. Since the algorithm chooses a segment from the training data according to the distance, a distance measurement scheme has be obtained for the categorical data. Usually, it is conducted through different similarity measurement techniques. The algorithm used in the study combines these two classifiers in the way that both involved issues can be resolved. Particularly, in the algorithm there is no more requirement to discretize the continuous variables and, in the meantime, does not need to measure the distances among categorical attributes. The combination is supposed to enhance the algorithm performance and will be reliable in nature.

This technique goes under third (iii) methodology mentioned above. Hybrid classifier of Bayesian Network (special case of BN - Naïve Bayes) and Nearest Neighbor distance based algorithms are applied to the dataset. The Bayesian Network structure is obtained from the data and the Nearest Neighbor algorithm is used in combination with the Bayesian Network (Dedabrishvili, 2017).

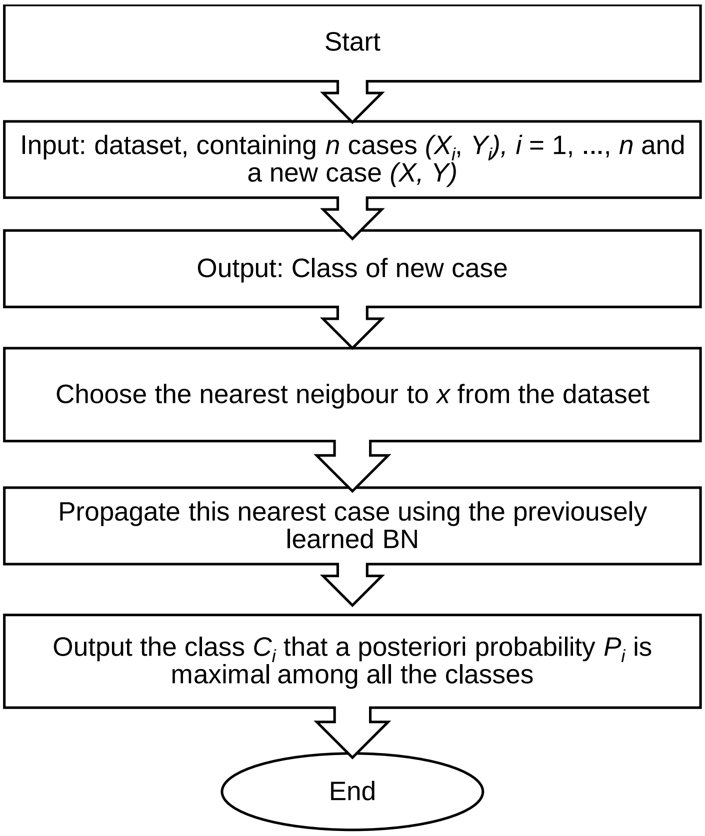

Figure 4. The pseudo-code of the kNN-NB Hybrid Algorithm.

Steps of the combined algorithm is presented in *Figure 4.* New cases in the training dataset are classified according to the nearest case and final decision is made by propagating the evidence of this nearest case in the previously learned Bayesian Network. The schema of the new case classification is given in *Figure 5.*

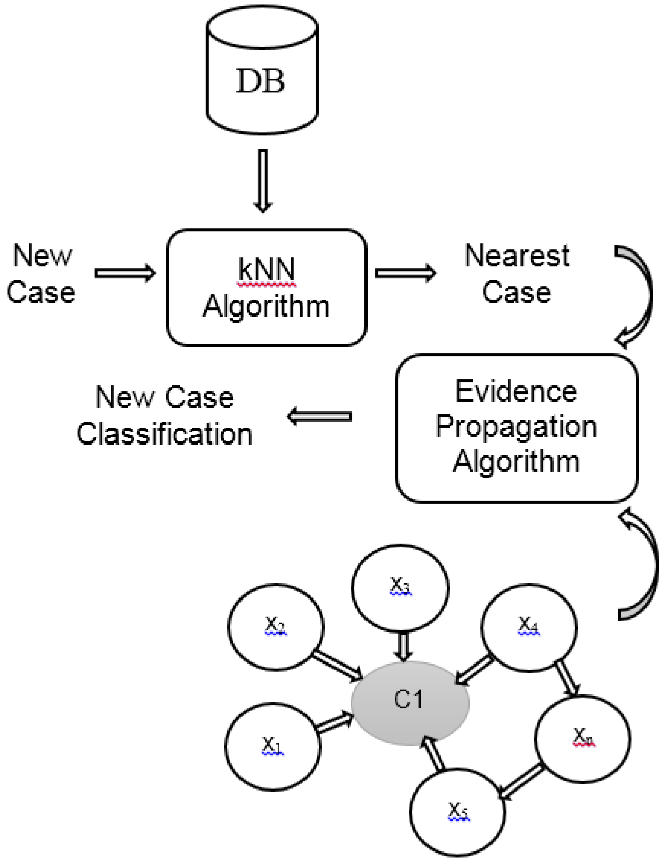

Figure 5. The scheme of new case classification.

To conclude, the classification of a new object happens through the following way: first KNN algorithm is applied to find the K Nearest Neighbor from the training dataset. During implementation of the KNN, the categorical attributes are not involved. The distance measurement is accomplished only by the numerical attributes. After choosing the K nearest object, then a model utilizing the Naïve Bayes algorithm is constructed, but this time only the categorical attributes are taken into the consideration. Only combined model is responsible for the classification of the new object. Accordingly, this is a two-step process, in which the first step coincides only the numerical attributes to select the closest data of the new object. This is logical, as numerically close objects should have the same characteristics. Providing the K nearest object of the new object, as a next step, rather than, taking basic voting scheme as it happens in *k*NN, the features of the categorical data and their relation to the class is discovered by the Naïve Bayes classification technique. Thus, both the numerical and categorical attributes are involved for the classification of a new object without touching to the data, and what is important, no discretization or complex similarity measurement is no more required.

Problem Statement 2: Supervised and Unsupervised Algorithm Combination Methodology

It is known fact that HMMs suffer from intrinsic limitations, mostly because of their arbitrary parametric assumption (Trentin & Gori, 2003). With this respect Artificial Neural Networks appear to be a promising alternative. The combination of the algorithms used in the study is grounded on a gradient-ascent method for global training of a hybrid ANN-HMM system, where the ANN is trained for estimating the emission probabilities of HMM states. The approach is associated to the major hybrid systems developed by Bourlard, Morgan and Bengio, with the goal of combining algorithm benefits within a united framework in order to overcome their constraints. The applied method contains several functions (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)
```

The structure of the algorithm application is given in the *Figure 6.*

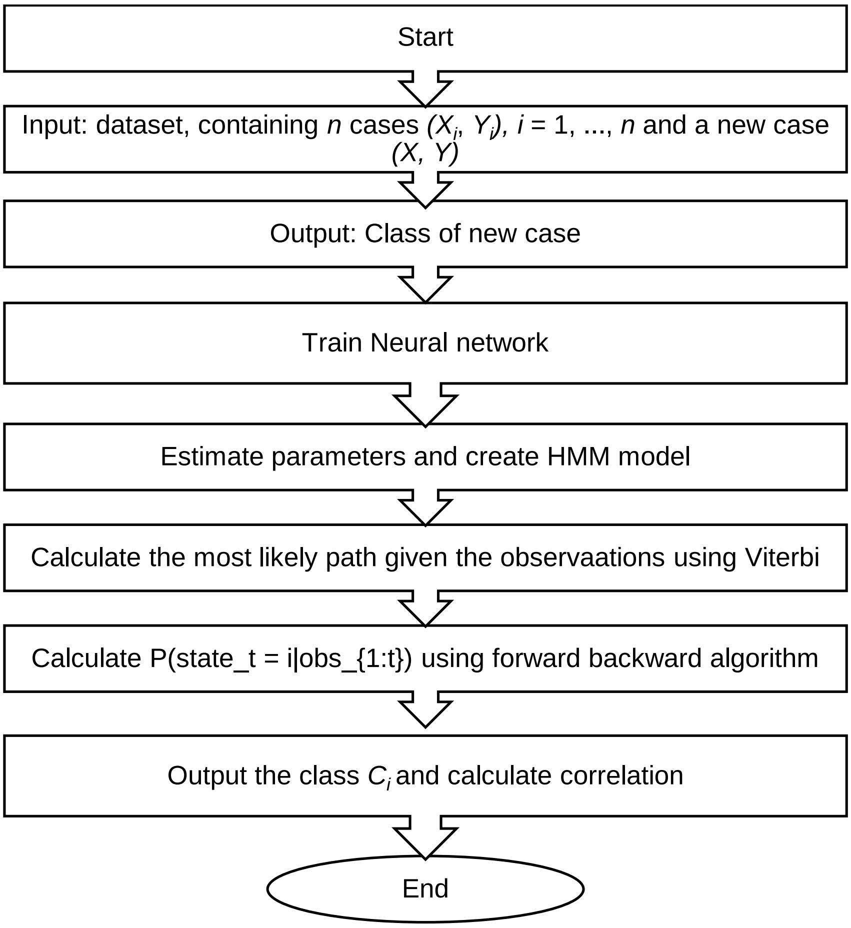

Figure 6. The pseudo-code of the NN-HMM Hybrid Algorithm.

Chapter 3. Proposed and Developed HAR Techniques

In this chapter, proposed methodology is presented, which includs data collection, the classifiers usage and the performance evaluation using the 10-fold cross validation method. *Figure 1*summarizes the different steps of the adopted approach using seperate algorithms (part one), whereas, *Figure 7* represents the steps of the approach by employing the combinations of different algorithms (part two).

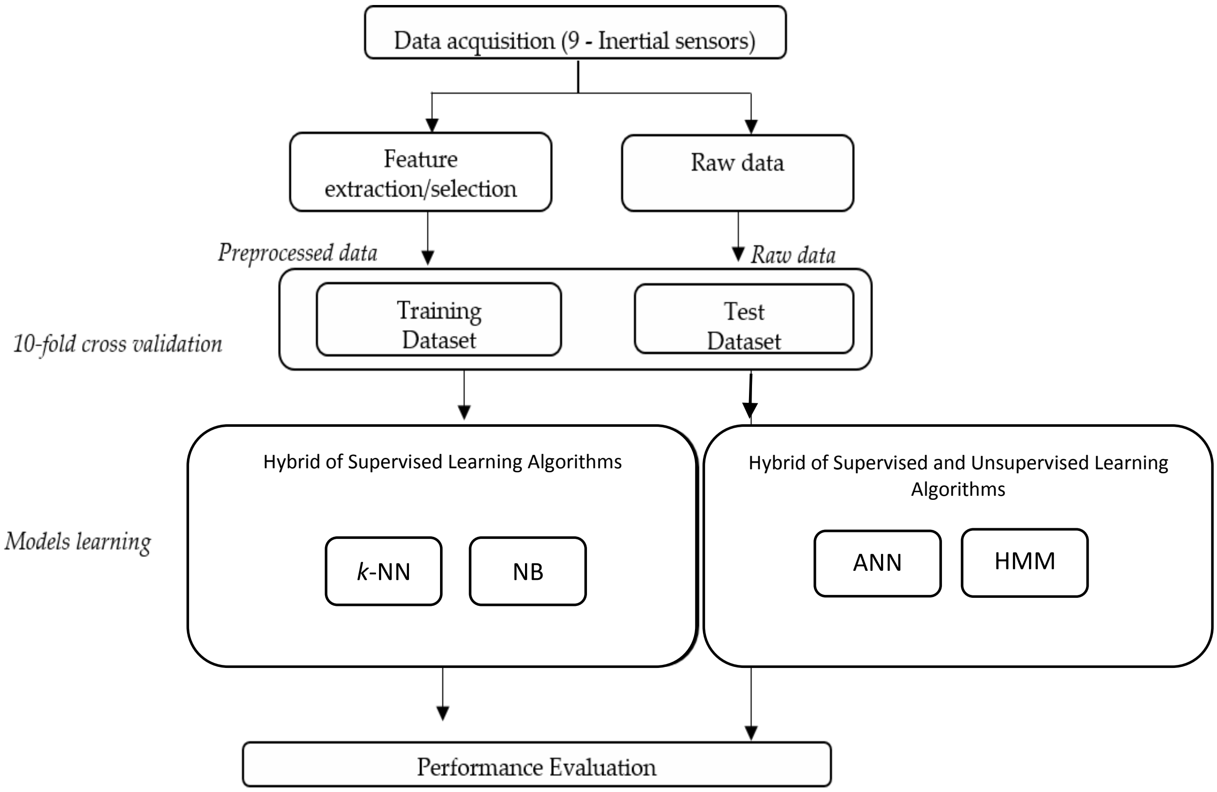

Figure 7. Steps of human activity recognition using hybrid of algorithms.

In this study, human activities are estimated using the Xbus Kit from Xsens (Enschede, Netherlands) which enables ambulatory measurement of the human motion. It consists of a portable system that incorporates an Xbus Master and three MTx inertial units that are placed on the chest, the right thigh and the left ankle of the subject *Figure 8*.

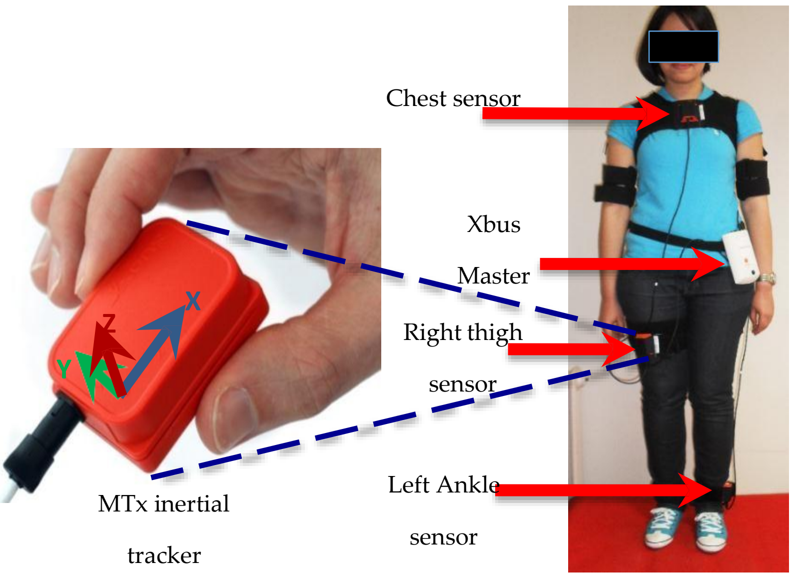

Figure 8. MTx-Xbus inertial tracker and sensors placement.

Data were collected at the LISSI Lab/University of Paris-Est Creteil (UPEC). Six healthy subjects with different profiles (mean age: 26 years old, mean weight: 65 kg) participated in the experiments. The subjects were given instructions to perform activities in their own way without specific constraints. Each subject conducted a total of twelve activities. The data acquisition was performed in the office environment over a period of about 30 min. The different activities and their descriptions are given in *Table 1*. The acquired data were manually labeled by an independent operator (Attal, Mohammed, Dedabrishvili, & Chamroukhi, 2015).

Table 1. List of the selected activities (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

Experimental Results

In this section, the performances of the standard supervised and unsupervised ML approaches which were used to recognize the daily living activities are reviewed and compared. This comparison highlights the different algorithm performances in terms of average accuracy rate (R) and its standard deviation (std), F-measure, recall, precision and specificity. In this comparative study, several cases are considered:

Case 1: Raw Data

The results obtained in the case of raw data are given in *Tables 2 and 3*. *Table 2* summarizes the performance results obtained when using the supervised approaches. It can be observed that the correct classification rates obtained with different techniques are all higher than 84%.

The *k*-NN algorithm gives the best results in terms of global correct classification rate, F-measure, recall, and precision, followed by RF, then SVM and at finally the SLGMM algorithm gives relatively the worst results.

Table 3 summarizes the results obtained when using the different unsupervised learning approaches.

Table 2. Performances of the supervised algorithms using raw data.

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

Table 3. Performance results of the unsupervised algorithms using raw data.

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

In order to identify the patterns that are difficult to recognize, the global confusion matrix are given in *Tables 4* and *5* in the case of *k*-NN and HMM, respectively. One can observe that confusions in most cases, occur between transition activities such as (A9, A7) and dynamic activities such as (A1, A11), (A1, A12) and (A11, A12). These confusions are more important in the case of HMM. One can also observe that the basic activities such as A2, A4, A8 are easier to recognize than transition activities such as A3, A5 and A7.

Table 4. Global confusion matrix obtained with k-NN using raw 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

Table 5. Global confusion matrix obtained with HMM using raw data.

		Obtained Classes											
		A1	**A2**	**A3**	**A4**	**A5**	**A6**	**A7**	**A8**	**A9**	**A10**	**A11**	**A12**
	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
True	**A5**	0.67	0	7.20	0.29	73.61	0.10	1.06	0	1.44	15.55	0	0.10
Classes	**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

Case 2: Feature Set Extracted/Selected from Raw Data

In order to improve the results presented above a preprocessing step consisting of features extraction and selection is performed. Nine accelerometrics signals are acquired from three MTx IMUs and for each signal; the following time and frequency domain features are calculated:

- Eleven time-domain features are extracted, namely: mean, variance, median, interquartile rang, skewedness, kurtosis, root mean square, zero crossing, peak to peak, crest factor and rang.
- Six frequency-domain features are extracted, namely: DC component in FFT spectrum, energy spectrum, entropy spectrum, sum of the wavelet coefficients, squared sum of the wavelet coefficients and energy of the wavelet coefficients.

Table 6. Performances of the supervised algorithms using extracted features.

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

Table 7. Performances of the unsupervised algorithms using extracted features.

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

Tables 8 and 9 represent confusion matrix obtained with k-NN and HMM using selected features.

Table 8. Global confusion matrix obtained with k-NN using selected features.

		Obtained Classes											
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
	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
True	A5	0	0	0	0.23	99.38	0.23	0	0	0.08	0.08	0	0
Classes	A6	0	0	0	0	0.07	99.78	0.07		0.03	0.05	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
	A9	0	0	0	0	0.08	0.17		0.33	99.42	0	0	0
	A10	0.35	0.18	0	0	0.09	0.09	0	0	0	99.20	0.09	
	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 obtained with HMM using selected features.

		Obtained Classes											
		A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	A12
	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
True	A5	2.05	0	1.80	0.66	71.62	4.35	2.21	0	5.50	11.48	0	0.33
Classes	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
	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

Case 3: Experimental Results Using Combination of Supervised Learning Algorithms

In this section, there is a review and comparison of experimental results on the dataset of human activity recognition using multi-classifier or hybrid classifier of NB and *k*-NN.

While learning the dataset, new cases were classified according to the following process:

(i) Firstly by looking for the nearest neighbor case in the training database affording to the *k*-NN algorithm, where, *K*i represented the nearest case,

(ii) Then, by propagating the *K*i case in the learned BN as if it was the new case,

(iii) And finally, after propagation according to the posteriori higher probability (which is done by achieving two sub-goals of the Bayesian network approach: fixing the network structure and establishing the values of the probability tables for each node) by marking the new case with class label (Dedabrishvili, 2017).

The results of the experiments are given in *Table 10*. As described in above (Attal, Mohammed, Dedabrishvili, & Chamroukhi, 2015) dataset has passed the preprocessing phase and its' dimensionality is reduced using Principal Component Analysis.

Table 10. Performances of the algorithms separately and in hybrid manner (k-NN-NB) using extracted features.

(%)	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	**0.99526**	0.00474	0.99526	0.99527

Case 4: Experimental Results Using Combination of Supervised and Unsupervised Learning Algorithms

Here experimental results of the combination of supervised and unsupervised algorithms are presented. The results are achieved through the application of ANN and HMM algorithms to the HAR dataset.

During the learning process of the dataset, new cases were classified according to the following process:

(i) First of all, by training the Neural Networks,

(ii) Then, by estimating the parameters (Transition Matrix, Emission Matrix and PI) and by creating HMM model,

(iii) After that, by calculating the most likely path given the observaations using Viterbi algorithm,

(iv) And finally, by Calculating the probability, P(state_t = i|obs_{1:t}) using Forward Backward algorithm.

Table 11. Performances of the algorithms separately and in hybrid manner (NN-HMM) using row data.

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	**89.23342**

Experiments show, that Supervised learning algorithms outperform Unsupervised learning approaches in general. But obtained resuls are encouraging in latter learning environment. It is also remarkable that in the study context hybrid algorithms provide better accuracy lavels than seperate algorithms, which can be explained for *k*-NN-NB combination accuracy rate increament by 0.3%, while NN-HMM hybrid provided 9% better classification.

Conclusion

The core methodical and valuable achievements obtained in the dissertation:

- Has been found out the one of the most suitable classifier from study employed popular supervised learning algorithms in HAR.
- Has been found out the one of the most suitable classifier from study employed popular unsupervised learning algorithms in HAR.
- Has been developed new hybrids of classifiers for data classification which outperform solely application of the algorithms in the field of HAR.
    - Has been proposed original way of combination of Instance-based algorithms with Naïve Bayes approach.
    - Has been proposed a way of combination of Hidden Markov Model with Artificial Neural Networks.
- Has been shown that strengths of the single algorithm can be used as a pa65rt of another algorithm with weaknesses at the similar point.
- Has been shown that by concatenation of trained data from different classification algorithms, the overall accuracy rate can be increased.
- Has been developed a way to create a Function in the future of the hybrid classifiers for other users to apply in neighboring areas of the field straightforwardly.

Possible Directions and Recommendations for Further Study

- Study can be further extended by making experiments using different machine learning methods in both supervised and unsupervised contexts on increased lavel of focus group, which means expansion of the dataset by adding further participants in general and, in particular, elderly subjects.
- For the convenient usage of the sensors, later studies can involve smart phone capabilities to control the motion of the user.
- Experiments can be continued using other classification algorithms and combinations of them for both row and feature extracted data.

- Can be created Matlab Functions of hybrid classifiers to assist other users perform classification easier.
- Mobile applications can be involved in the remote monitoring process to catch abnormal situations and let caregivers act without delay.

Published articles

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.