This is a pre-print version of the article that has been accepted in the Applied Soft Computing, Special Issue on Soft Computing Methods for Remote and Mobile Healthcare Applications

The accepted article can be accessed on the following link:
http://www.sciencedirect.com/science/article/pii/S1568494615003014
doi:10.1016/j.asoc.2015.05.001

Context-based Ensemble Method for Human Energy Expenditure Estimation

Hristijan Gjoreski[1,2], Boštjan Kaluža[1,2], Matjaž Gams[1,2], Radoje Milić[3], Mitja Luštrek[1,2]

Affiliation of authors:

[1] Jožef Stefan Institute, Department of Intelligent Systems, Jamova cesta 39, 1000 Ljubljana, Slovenia

[2] Jožef Stefan International Postgraduate School, Jamova cesta 39, 1000 Ljubljana, Slovenia

[3] University of Ljubljana, Faculty of Sport – Institute of Sport, Gortanova 22, 1000 Ljubljana, Slovenia

*Corresponding author: Hristijan Gjoreski
Address: Jamova cesta 39, 1000, Ljubljana, Slovenia
Tel: +386 1 477 3812
E-mail: hristijan.gjoreski@ijs.si

ABSTRACT

Monitoring human energy expenditure (EE) is important in many health and sports applications, since the energy expenditure directly reflects the intensity of physical activity. The actual energy expenditure is unpractical to measure; therefore, it is often estimated from the physical activity measured with accelerometers and other sensors. Previous studies have demonstrated that using a person's activity as the context in which the EE is estimated, and using multiple sensors, improves the estimation. In this study, we go a step further by proposing a context-based reasoning method that uses multiple contexts provided by multiple sensors. The proposed Multiple Contexts Ensemble (MCE) approach first extracts multiple features from the sensor data. Each feature is used as a context for which multiple regression models are built using the remaining features as training data: for each value of the context feature, a regression model is trained on a subset of the dataset with that value. When evaluating a data sample, the models corresponding to the context (feature) values in the evaluated sample are assembled into an ensemble of regression models that estimates the EE of the user. Experiments showed that the MCE method outperforms (in terms of lower root means squared error and lower mean absolute error): (i) five single-regression approaches (linear and non-linear); (ii) two ensemble approaches: Bagging and Random subspace; (iii) an approach that uses artificial neural networks trained on accelerometer-data only; and (iv) BodyMedia (a state-of-the-art commercial EE-estimation device).

Keywords: human energy expenditure estimation; machine learning; regression; ensembles; context; wearable sensors.

1. Introduction

Human energy expenditure (EE) directly reflects the intensity of physical activity, which makes it important for sports training, weight control, management of metabolic disorders (e.g., diabetes), and other health goals. There are various approaches that can reliably estimate the EE. Direct calorimetry measures the total heat output of a person in an accurate way, but is only usable in laboratory conditions. The slightly less accurate indirect calorimetry analyzes the respiratory gases and requires wearing a breathing mask, making it impractical for everyday use. Doubly labeled water is both accurate and convenient, but can measure only long-term EE. Finally, self-reporting is highly unreliable. Therefore, if both accuracy and convenience are required, a different approach is needed.

With the increasing accessibility and miniaturization of sensors and microprocessors, ubiquitous monitoring systems are becoming a practical solution for measuring the EE. Such systems primarily measure the physical activity with accelerometers, but can include additional sensors that indirectly measure the metabolic activity, such as a heart rate monitor or thermometer. The main challenge is how to estimate the EE from wearable sensor outputs accurately, irrespectively of the participant's activity, ambient conditions and other circumstances, i.e., contexts.

Recent studies in the EE-estimation field showed that machine learning (ML) techniques applied on sensor data can accurately estimate the EE [1][2][3]. In these studies, the EE estimation is defined as a process of transforming the sensor data into METs, where one MET is defined as the energy expended at rest. MET values usually range from 0.9 (sleeping) to over 20 (extreme exertion). Researchers usually use an indirect calorimeter to estimate the actual EE in METs with a high accuracy, which is later used as the ground truth during the ML phase.

In this study, we propose a novel, Multiple Contexts Ensemble (MCE) approach, which is applied to the task of EE estimation. The MCE approach uses multiple contexts extracted from sensor data and performs context-based reasoning in order to estimate the EE. In general, context is any information that characterizes the circumstances in which an event/situation occurs [4]. In our application, the context information is represented by the eight features extracted from the sensor data: activity, heart rate (HR), breath rate (BR), acceleration counts, chest skin temperature, galvanic skin response (GSR), arm skin temperature and near-body ambient temperature. Each of these features is used as a context in which ML models are built using the remaining features as training data. More precisely, for each value of each context feature, a regression model is trained using the subset of the dataset that corresponds to that particular context (feature) value. For example, for the activity of the user, a regression model is trained for each activity (sitting, walking, running, etc.) using the rest of the features as training data (HR, BR, body temperature, etc.). When evaluating a data sample, a custom ensemble of regression models is assembled from the previously constructed set of models, i.e., the models that correspond to the context (feature) values in the evaluated sample. The final estimation is provided by aggregating the outputs of the assembled models. This way, context-based reasoning is performed, which provides the benefit of combining multiple "viewpoints" when estimating the EE, resulting improved accuracy compared to previous approaches.

The remainder of this paper is organized as follows: Section 2 presents the background of the study and reviews the related relevant methods; Section 3 describes the proposed MCE approach; Section 4 presents the experimental setup, including the description of the activity scenario, the sensor equipment, the evaluation technique, and the description of the competing approaches; Section 5 presents the experimental results and a discussion; and lastly, Section 6 offers concluding remarks.

2. Background and Related Work

The first automatic methods for EE estimation included supervised ML, i.e., regression learning techniques. In particular, linear regression was used to map a single accelerometer output to EE [5][6][7][8]). The accelerometer output was often expressed in “counts”, an aggregate acceleration measure reported by devices such as ActiGraph. To estimate the EE, investigators used these "counts" to develop linear regression models. Although numerous studies showed reasonably good correlation between the counts and the EE [5][9], the estimation accuracy of accelerometer count-based linear regression was shown to contain systematic errors and vary with the type of activities, resulting in overestimations for the walking activity and underestimations during the moderate intensity lifestyle activities [10]. This limitation is probably due to the insufficient information provided by the counts and the simplicity of the linear model. Efforts were made to improve the estimation accuracy by using a richer representation of the accelerometer output consisting of multiple features [11][12], as well as non-linear regression methods such as artificial neural networks (ANNs) [1][13][14] or support vector machine for regression (SVR) [15][16]. These approaches were experimentally shown to substantially improve the accuracy compared to earlier work [17].

Researchers soon realized that single-regression approaches cannot accurately estimate the intensity of physical activity across a range of activities, and that different activities require different EE equations. Crouter et al. [18] used the acceleration counts in order to divide the activities into three categories and assigned the following EE estimations: 1 MET to inactivity and two regression equations for light and intense activity, thus achieving a better estimate than previous single-regression methods. The advances in the accelerometer-based recognition of activity type allowed finer-grained activities as the context for EE estimation [2][20][19]. Lester et al. [21] used a Naive Bayes classification model to first recognize three activities (rest, walking and running) from the accelerometer's data, and then to apply the appropriate regression equations in order to estimate the EE. They also considered GPS and barometer information to estimate the slope of walking/running, and showed that additional sensor information improves the EE estimation. However, even with these three types of sensors (accelerometer, GPS, and barometer) they still encountered two problems: (i) EE underestimation of activities that are not characterized by acceleration, but are still energy demanding, e.g., carrying a box and (ii) EE underestimation of activities that follow an intense activity, i.e., the "cool-down" effect (sitting after intense running). Both problems can be solved by sensing other physiological parameters such as the HR and BR. In our previous work [22], we showed that by using data from multiple sensors one can more accurately estimate the EE. This may seem as an additional burden to the user, because it requires additional sensors attached to him/her. However, today's commercial wearable

devices already provide multiple sensors packed in a single enclosure, e.g., BodyMedia, Basis, Empatica wristband, etc.

The BodyMedia armband sensor uses both multiple sensors and multiple regression models. Vyas et al., [3], the research team of the BodyMedia, proposed a method that uses an activity recognition model that recognizes dozens of activities which are used as the context, and then it combines multiple regression models according to the probabilities for the recognized activities. They showed that by using multiple sensors: an accelerometer, two thermometers, GSR and heat-flux sensors, the estimation of the EE significantly improves. Additionally, a recent review showed that it is the most accurate EE estimation consumer device [23].

The aforementioned studies showed that: (i) using multiple regression models for different user's activities (i.e., context) outperforms single-regression approaches, and (ii) using multiple features extracted from multiple sensor data provides more accurate EE estimation than using only acceleration data (even when multiple acceleration features are extracted). In this work we improve upon these findings and propose a method that uses multiple features extracted from multiple sensor data, and uses not only the activity as the context, but multiple contexts, so that each measurement can be placed in multiple contexts simultaneously (e.g., activity = running, HR = high, BR = moderate, etc.).

3. Multiple Context Ensemble Approach

The proposed MCE is a general approach which can be applied to various reasoning tasks about the user's health and condition. The only requirement is that the data to reason about can be represented by multiple context features. This is usually the case when the reasoning task includes multiple sources of information, for example data streams provided by multiple sensors.

The application of the MCE approach to the task of EE estimation is shown in Figure 1. It consists of three phases: context extraction, context modeling and context evaluation. In the first phase the data provided from multiple sensors is used in order to extract eight features. In the second phase, each of these eight features is individually exploited as the user's context, and the other seven features to model the EE in the context of the first feature. That is, for each value of each feature a regression model is trained on the subset of the dataset that corresponds to that particular value using a regression learning method. In the evaluation phase a custom ensemble of regression models is assembled from the previously constructed set of models, i.e., the models that correspond to the context (feature) values in the evaluated data sample. The final EE estimation is provided by aggregating the outputs of the ensemble models by using an aggregation technique such as averaging, median or stacking. Each of the phases is described in the following three subsections.

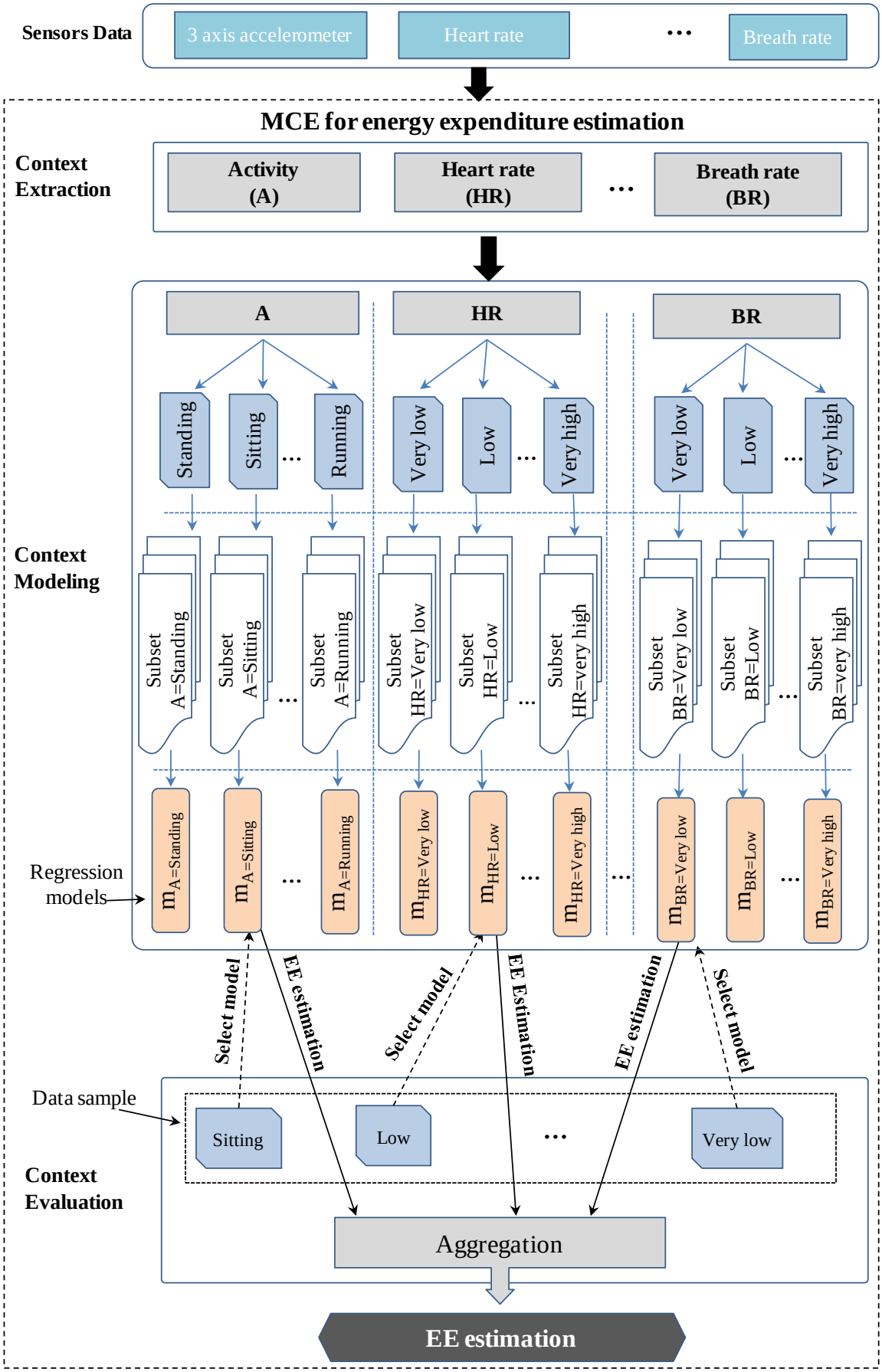


Figure 1. Multiple Contexts Ensembles (MCE) EE estimation algorithm.

3.1. Context Extraction

In the context extraction phase, raw sensor data are acquired and the multiple contexts are extracted. This phase is similar to the feature-extraction phase commonly used in ML tasks. We refer to features as contexts, because in our approach each feature is individually used as a context in order to train multiple regression models. In this study, each feature represents unique information about the user: activity, acceleration peak counts (similar to a pedometer), HR, BR, chest body temperature, arm body temperature, GSR, and near-body temperature. More details about the sensor equipment and the raw sensor data are provided in Subsection 4.2.

Similar to feature values in ML, each context has values (context values), e.g., "sitting" for the activity context (see Figure 1). In our case most of the extracted features were in numerical format, i.e. numbers that represent the user's heart-rate, temperature, etc. In order to train a reasonable number of models for different context values, a discretization procedure was performed. The discretization process allows us to group data samples with similar context values, e.g., data samples with "very low" HR value.

Each numerical feature was discretized using the split criterion proposed by Yong et al. [24]. It is the most commonly used supervised discretization technique in the ML community, which finds such splits for a given feature that the standard deviation of the class (EE) value in each interval of the feature is minimized. The standard deviation reduction (SDR) achieved by a given split is calculated by the formula:

$$SDR = sd(T) - \left[\frac{|T_1|}{|T|} \times sd(T_1) + \frac{|T_2|}{|T|} \times sd(T_2)\right] \quad [1]$$

where T is the set of data samples before the split, T_1 and T_2 are the sets that result from the binary split, $|T|$ is the number of data samples in the set T, and $sd(T)$ is the standard deviation of the class value. The discretization procedure tests all the possible splits and selects the one with the highest SDR. This is repeated as long as at least 10% of the data samples remain in each interval. This resulted in 46 discrete context values – around 6 per context on average.

3.2. Context Modeling

In the second phase, the context modeling is performed by first partitioning the data into multiple subsets using each feature as a context, and then learning a regression model using the subsets.

In order to explain the context modeling phase, consider the dataset shown in Figure 2. It consists of three features (activity, HR and BR) and a target feature (EE in MET). A conventional ML approach would apply a single ML algorithm to learn a single regression model. More advanced approaches may use an ensemble-based approach and learn multiple models on multiple subsets of the dataset. These subsets are usually created by using techniques such as bootstrapping (sampling with replacement) [25], and modifying the dataset in the feature space (e.g., Random Subspace method, which randomly chooses a subset of features multiple times) [26]. Even though these techniques have proven to be successful in numerous ML applications [27], they do not take the nature of the domain into account. In EE estimation and other tasks dealing with a human in an environment, the context is known to be very important [23] and is also close to how humans reason about

such situations. Because of that, MCE uses each of the features as a context, and the reasoning is performed using multiple contextual views of the data.

An example of such context-based data partitioning is shown in Figure 2, where the BR feature is used as the context. Therefore, the dataset is partitioned according to each feature value (low, medium and high), resulting in three subsets. This way, each subset contains data samples with similar values for the chosen context value. In the next step, for each of the subsets a regression model is trained. The same procedure is performed for the other two features, i.e., activity and HR, resulting in nine regression models (a model for each feature value). When evaluating a data sample, only the models that correspond to the particular feature values are invoked in order to evaluate it. This way each data sample is evaluated by an ensemble constructed of three models that correspond to the three contexts.

Figure 2 also shows the advantage of using intervals (discrete values) instead of numerical values. In particular, the discrete values for BR are used to create the subset for the BR feature.

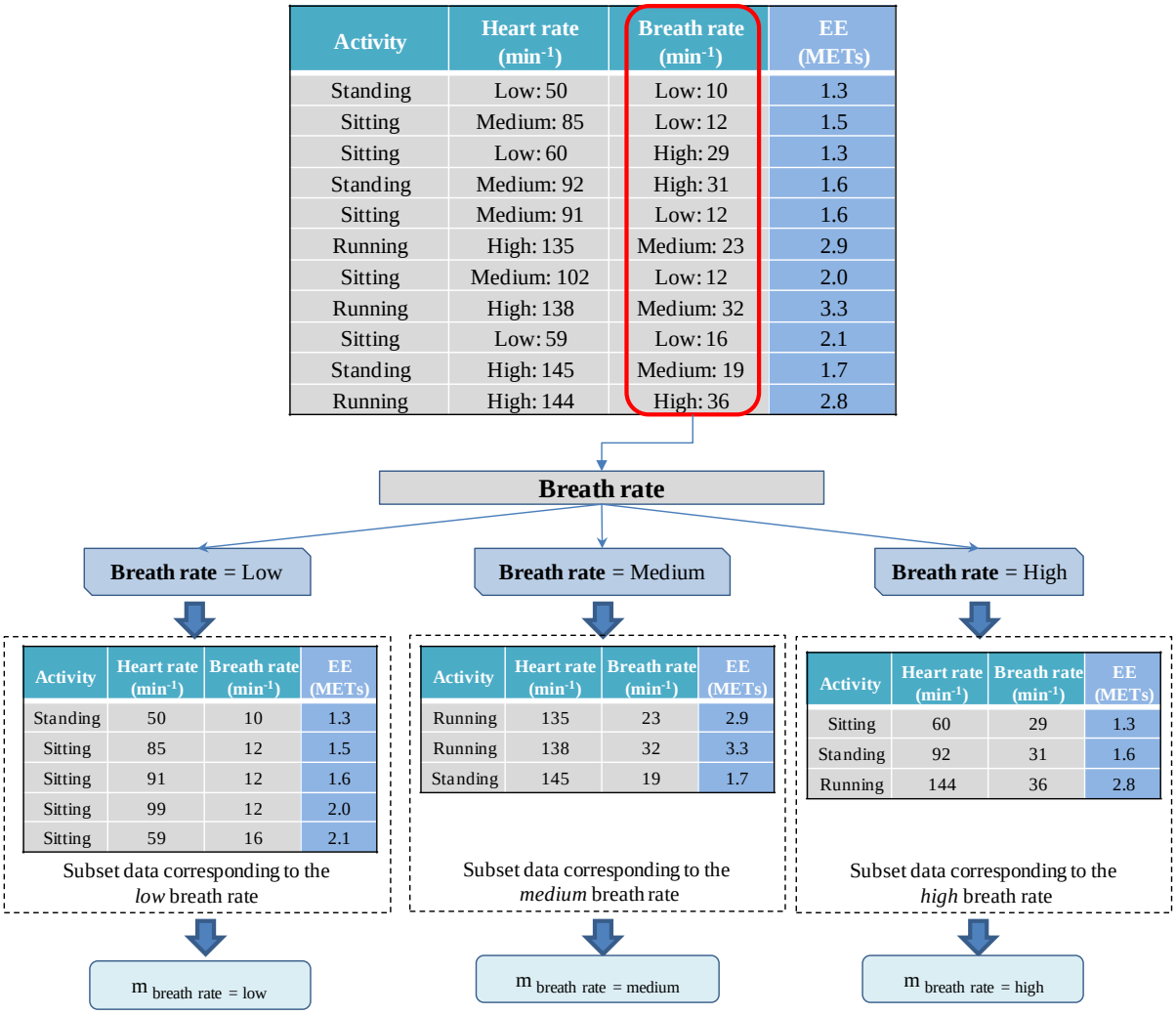


Figure 2. Context-based data partitioning.

For the application of MCE to EE estimation, 46 regression models were constructed (for each discrete value of each context). The MCE method does not restrict the choice of the regression learning algorithm; therefore we used and later compared the results achieved by

five linear and non-linear regression learning methods as implemented in the WEKA ML toolkit [28]. More details about these methods are provided in Subsection 4.4.

By constructing multiple models corresponding to different contexts of the user, the MCE considers multiple views on the reasoning problem. This way, the MCE not only exploits the complementarity of multiple models like most other ensemble approaches, but also contains models that tend to be more accurate for a particular context than those trained on the whole training set. The reason for that is that each model is trained on a subset of the training set that is more homogeneous than the whole set, and used in the context of this subset, i.e., to reason about data samples similar to the ones in the subset. In other words, in our approach we try to semantically split the domain (dataset) into meaningful viewpoints and not on some statistics about the data (as most of the ensemble-based approaches).

3.3. Context Evaluation

When evaluating a data sample, a custom ensemble is constructed from the models that correspond to the context (feature) values of that sample. In the context evaluation phase, the estimations from each of the context models are combined in order to provide the final estimation of the EE. For example, consider the scenario shown in Figure 1 using three contexts: a user is *sitting* with the HR of *50* min^{-1} and BR of *10* min^{-1}. Suppose that the HR value falls in the second HR interval (*low*), and the BR value into the first BR interval (*very low*). The data sample will thus be evaluated by the models $m_{\text{A=sitting}}$, $m_{\text{HR=low}}$ and $m_{\text{BR=very low}}$, whose outputs will be aggregated by an aggregation method in order to estimate the final EE. The related literature suggests various aggregation methods, including averaging, median, aggregating with learning − stacking, and similar. In this study, we tested the first two (averaging and median), which are the simplest for implementation but still enabled the MCE approach to achieve significantly better performance than other ensemble methods that use the same aggregation technique (i.e., MCE with averaging outperformed Bagging and Random subspaces, which also use averaging). We plan to experiment with more advanced techniques in future work.

4. Experimental Setup

4.1. Participants and Activity Scenario

A total of ten healthy participants (age 27.2 years (SD = 3.1); BMI 24.1 $\text{kg}\cdot\text{m}^{-2}$ (SD = 2.3); weight 78.2 kg (SD = 10.9)) completed a two-week study (each participant was recorded during one day for approximately eight hours). Before testing, height and weight (one layer of clothes, no shoes) were measured via the InBody-720 body-composition analyzer. Prior to participation, informed consent was obtained from the participants. Each participant was observed by a medical supervisor during the execution of a comprehensive pre-defined activity scenario. The activity scenario included 15 different atomic activities, which were categorized into seven activity types according to the intensity and the type, as presented in Table 1.

Table 1. Activity scenario.

Activity type	Atomic activities	METs
Sedentary	Lying, sitting, standing, on all fours, kneeling	1.0−1.5
Light household activities and exercise	Washing dishes, working on a PC, lying and doing light exercise, walking doing light chores	1.5−2.5
Moderate to vigorous household activities	Scrubbing the floor, shoveling snow – digging	2.5−3.5
Walking	Walking on a treadmill with 4 km/h	4.0
Running	Running on a treadmill with 8 km/h	8.0
Light cycling	Light stationary cycling: 1 W/kg of body mass, 65 RPM	5.0
Vigorous cycling	Vigorous stationary cycling: 2 W/kg of body mass, 65 RPM	7.5

4.2. Sensors

An example of a person wearing the sensor equipment and walking on a treadmill is shown in Figure 3. The equipment consists of the following wearable sensors: two 3-axis accelerometers, a Zephyr sensor, and a BodyMedia sensor. Each of the sensors used in this study provided different information about the user's EE. Because the BodyMedia sensor is the state of the art EE estimation sensor, its MET output was used for comparison. Additionally, a Cosmed indirect calorimeter was used to provide the ground truth for the EE estimation.

Figure 3. A participant wearing the sensor equipment while walking on a treadmill.

Indirect calorimeter − Oxygen uptake (VO_2) during each activity was measured breath-by-breath and averaged every ten seconds using the Cosmed K4b2, a light-weight portable indirect calorimeter. Prior to each test, the Cosmed unit was calibrated according to the manufacturer's guidelines. This sensor data were used as the ground truth for the regression learning algorithms and for the evaluation of the performance of the tested EE-estimation methods.

Accelerometers − The Shimmer sensor platform was used to measure the accelerations of the user. The chosen platform contains a 3-axis accelerometer and uses Bluetooth communication for sending data in real-time. To record the participants' acceleration, 50 Hz data-sampling frequency was used. Each participant wore two accelerometers while performing the activities. They were attached to the participants' chest and thigh using elastic Velcro straps. The placement was chosen as a trade-off between the physical intrusiveness and the performance achieved for the activity recognition in prior studies of activity recognition [29][30]. The results achieved in those studies [29][30] showed that from the 6 placements analyzed (chest, waist, right and left thigh, right and left ankles), the combination of the chest and the thigh is the most suitable (tradeoff between the number of sensors and the accuracy) for the activity recognition task, achieving the accuracy of 93%.

Zephyr − The Zephyr BioHarness sensor is a commercial sports strap worn on the chest with direct contact to the skin. It measured the participants' HR, BR, and chest skin temperature, which were used as contexts.

BodyMedia − The BodyMedia sensor is a state-of-the-art commercial sensor for EE estimation, which was worn on the left upper arm as suggested by the manufacturer. It served as a benchmark for EE estimation, and additionally provided the data for GSR, ambient temperature and arm skin temperature used as features for EE estimation. The ambient temperature is an estimation of the ambient temperature near the arm.

4.3. Data Preprocessing

A custom PC application was created to record, preprocess and synchronize the multiple sensor data. During the recordings, the accelerometers' data were acquired on a laptop using Bluetooth and were manually labeled with the corresponding activity, which was later used for the training of the activity-recognition classification model. The data provided from the other sensors was labeled with the appropriate timestamp and stored locally in the sensor's internal memory. Afterwards, they were transferred into a database for offline analysis together with the accelerometer data and the activity labels. Once the multiple sensor data were saved into the database, they were synchronized using the unique timestamp for each data sample. In the next step, they were segmented using a sliding window of ten seconds. In each data window, eight context features were extracted from the sensor data, as shown in Table 2.

Table 2. The sensors and the appropriate contexts extracted.

Sensor	Contexts
Shimmer accelerometers	Activity, acceleration peaks count
Zephyr	Heart rate, breath rate, chest skin temperature
BodyMedia	Galvanic skin response, arm skin temperature, ambient temperature

Except for the activity and the acceleration peak counts, all other features are provided directly by the sensors (Zephyr, BodyMedia or Cosmed) and are computed by averaging the raw sensor data in the ten-second intervals. The physiological signals provided by the Zephyr and BodyMedia (HR, BR, etc.) differ from participant to participant and were additionally normalized. After empirical analysis of the data, we used the 15-minutes lying activity data recorded at the beginning of the activity trials in order to calculate the average resting value for each sensor, which was subtracted from each subsequent sensor value.

To extract the acceleration peak counts and the activity of the user from the acceleration data, they were first filtered using a band-pass filter [31]. The acceleration peak counts is the number of times the length of the acceleration vector stops increasing and starts decreasing or vice versa in the 10-second interval. For the activity recognition we used a previously developed classification method based on ML [29][32]. The method uses the data from the two accelerometers (chest and thigh), extracts 128 features and applies a Random forest classification model to recognize the atomic activities of the user: lying, sitting, standing, walking, running, cycling, bending, on all fours and kneeling. It achieved 93% accuracy on a one-second recognition interval. For the EE, the majority activity value was chosen for each ten-second window interval. A similar implementation of our activity recognition method achieved the best recognition performance at the international competition in activity recognition − EvAAL-2013 [33][34].

4.4. Evaluation Techniques

The evaluation of the method was performed using the leave-one-person-out cross-validation technique [35]: models were trained on the data of nine people and tested on the remaining person. This procedure was repeated ten times, once for each person. The same procedure was performed for the activity recognition classifier, whose output was used as a feature in the EE estimation. This evaluation technique is the most commonly used in the ML community if the model is intended to be used on a user different from the ones used for training, which is the case in the EE estimation [14]. This method yields an estimate of how well the model would do if it were applied to a population on which it was not trained. As for the evaluation metrics, Root Mean Squared Error (RMSE) and Mean Absolute Error (MAE) were used, since they are the most commonly used metric in the EE estimation domain. They are defined as follows:

$$RMSE = \sqrt{\frac{1}{q}\sum_{1}^{q}(EE_{estimated} - EE_{true})^2} \quad (1)$$

$$MAE = \frac{1}{q}\sum_{1}^{q}|EE_{estimated} - EE_{true}| \quad (2)$$

where q is the number of data samples, $EE_{\text{estimated}}$ is the estimated EE and EE_{true} is the ground-truth EE measured by the Cosmed device.

For each comparison, tests to confirm the statistical significance of the MAE and RMSE results were performed using paired Student's T-test with a significance level of 5%.

4.5. Competing Approaches

This subsection briefly presents the methods to which we compared the MCE approach. In particular, two types of comparisons were made: relative and absolute. The relative comparison compares the MCE to conventional ML approaches: single regression and ensembles. That is, the same dataset with the eight extracted features was used to evaluate both the MCE and the conventional ML methods. This comparison was proposed in order to confirm the hypothesis that context-based partitioning of the data improves the accuracy compared to conventional ML methods. On the other hand, the absolute comparison illustrates the accuracy of the MCE on an absolute scale, i.e., compared to estimations provided by the BodyMedia state-of-the-art commercial device and to an algorithm that is established as state-of-the-art accelerometer-based EE estimation commonly used in medical applications.

First, we compared the MCE to **single-regression learning methods**. This means that we trained single models on all the eight features. We tested five regression learning methods that are also commonly used in the EE literature: multiple linear regression (MLR) [36], support vector machine for regression (SVR) [37], Gaussian processes for regression (GPR) [38], model trees (M5P) [24], and multilayer perceptron feedforward artificial neural network (ANN) [39]. MLR is the simplest approach: it finds a linear function using all of the features in that matches the target variable, i.e., the EE [36]. The next method, SVR, is an extension of the classical SVM (commonly used for classification) adapted to regression, i.e., numeric output. The main characteristic of SVR is that the kernel function used for building the model ignores the training data samples close to the model hyperplane (within a pre-defined threshold) [37]. Similar to SVR, GPR is another non-linear, kernel-based method; however the theory behind the GPR is different compared to the SVR and is based on the Bayesian probability theory and assuming that the target variable follows a multivariate Gaussian distribution [38]. The next method, M5P is a model tree. The difference compared to standard decision/regression tree is that each leaf contains a linear function instead of a single value [24]. The last method is the ANN which is a popular non-linear regression learning method [39]. ANNs are composed of interconnecting artificial neurons, which are arranged in layers where each unit receives inputs from its immediately preceding layers. Each neuron computes a summation of its inputs weighted by a weight vector, and then applies an

activation function, which can be a logistic or linear function. In our case we used an ANN with one hidden layer and logistic learning function. Each of the five algorithms was used as implemented in the WEKA ML toolkit [28]. In addition to using these algorithms to build single-regression models as a baseline for comparison, we also used them to train base learners in the ensemble schemes constructed by our MCE algorithm.

Next, because MCE is an ensemble of regression models, we also compared it to two commonly used **ensemble learning methods:** Bagging [25] and Random subspaces [26]. Bagging is an approach that is based on bootstrapping, i.e., training multiple models on different subsets of the whole training dataset, constructed by sampling the whole dataset with replacement, and then aggregating the outputs from each model by averaging (regression) or voting (classification). Random subspace is an ensemble method proposed by Ho [6], which also modifies the training data; however, this modification is performed in the feature space. That is, a pre-defined number of features are selected randomly from the whole feature set. This procedure is repeated multiple times, creating a different training set for each selection. Then, for each training set, a regression model is built. Similar to Bagging, the final output is provided by aggregating the outputs from each model by averaging (regression) or voting (classification). Please note that for both ensemble techniques, the same five base machine learning algorithms were compared as in the single-regression learning.

Because the final goal of the approach is to be used in real-life applications, we also compared it to a commercial device for EE estimation. A recent review showed that **BodyMedia** commercial sensor is the most accurate EE estimation consumer device [23]. Therefore, we compared the MCE's estimated MET to the MET output of the **BodyMedia** commercial sensor (it should be noted that the BodyMedia sensor averaged the MET estimation over 1-minute interval, while our methods over 10-second interval). This comparison illustrates the accuracy of our method on an absolute scale (compared to a device which is already used in real life).

Finally, we re-implemented and compared the results to an approach that is widely accepted in the medical and sports research community [1][14]. It is an approach that uses an ANN trained with 6 features extracted from the chest accelerometer data only: 10^{th}, 25^{th}, 50^{th}, 75^{th} and 90^{th} percentiles of the acceleration peak counts, and the lag one autocorrelation. The approach was first introduced by Staudenmayer et al. [14] and further improved by Trost et al. [1], and is a state-of-the-art approach if a single accelerometer data is used to estimate the EE. From this point on, we refer to this method as **ANN-Acc** (ANN trained only on accelerometer data).

5. Experimental Results and Discussions

Table 3 compares four approaches in terms of RMSE and MAE: single regression, Random subspace, Bagging and our MCE. The five base learners explained in Subsection 4.5 were tested for each of the approaches. The best performing base learner is marked with bold. Additionally, the best performing approach for each base learner is marked with a gray background.

The results achieved by the single-regression methods show that in general the methods that use simple learning functions, e.g., linear or polynomial (SVR, GPR and MLR) are better

compared to the more complex ones such as ANN an M5P. This is in a way expected since ANNs and M5P are more susceptible to overfitting, and this problem is particularly likely to harm the performance when the testing data are from a person that is not used in the training data. When these same methods are used as base learners in the two ensemble schemes, i.e., Random spaces and Bagging, the results are similar to the single regression, except for the ANN and M5P, for which a slight improvement is achieved. However, when our ensemble method (MCE) uses the same base learners, the achieved RMSE and MAE are significantly better (lower) compared to the other three approaches. The difference ranges from 0.08 METs to 0.24 MET for the RMSE, and from 0.05 to 0.21 MET for the MAE.

The improvements of our MCE compared to the single-regression approach confirms the general rule in ensemble learning, i.e., ensembles tend to train multiple weak learners, and by combining the learner's outputs they create a stronger and more robust model [40]. The further comparison to the two standard ensemble approaches (Random subspace and Bagging) shows the advantage of using the nature of the domain (context) to resample the training data instead of using bootstrapping (Bagging) or randomly selecting features (Random subspaces).

The MCE approach is general and can use different techniques for aggregating the outputs of each model into a final one. In the tests shown in Table 3 we used the simplest aggregation technique, i.e., averaging. This technique was also used by the other two ensemble approaches: Bagging and Random subspaces, making the results more comparable. We additionally tested the performance achieved by the median technique.

Table 4 shows the comparison of the RMSE and MAE achieved by averaging compared to the median technique. The results show that the RMSE and MAE achieved by the median are almost always better (lower), except for the RMSE achieved by the M5P. The rationale why the median should work better is that by choosing the median value the models that are not accurate for some situations are not taken in consideration, which is not the case if the average is chosen. Because SVR by using the median achieved the best overall results (0.825 RMSE and 0.601 MAE), it is used in all further analysis.

Since the MCE consists of eight contexts, we additionally show the results if only a single context is used. In Figure 4, one can see that the MCE's EE estimation is better than the estimations provided by each of the base learners individually using RMSE metrics. This shows the advantage of using an aggregation function, i.e., by combining the individual models outputs using a median, the ensemble outperformed the individual models. This result is in accordance with the hypothesis presented by Dietterich [40], who studied the process of combining (aggregation) of the decisions provided by multiple models. He showed that it is better to find a good aggregation function instead of choosing the best single model, which also leads to stronger generalization.

Of particular interest is the comparison with the first regression model, which uses the activity as the only context: it only uses different regression models for different activities, like in several pieces of related work [18][21][3]. The results show that by combining multiple contexts one should expect better performance than by only using the activity as the context, i.e., a decrease of the RMSE by 23%.

Table 3. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for the MCE's MET estimation compared to single regression, Random subspace and Bagging using 5 base learners: artificial neural network (ANN), support vector regression (SVR), multiple linear regression (MLR), Gaussian processes for regression (GPR), and model tree (M5P). The best performance achieved by each of the aggregation techniques for each base learner is marked with bold. The overall best performance for each aggregation technique is marked with a gray background.

	Base learner	Single regression	Random subspace	Bagging	MCE
RMSE	**ANN**	1.094	1.059	1.054	**0.850**
	SVR	**0.962**	1.033	**0.965**	0.851
	MLR	0.967	1.033	0.969	0.854
	GPR	0.967	1.081	0.968	0.883
	M5P	1.113	**0.991**	0.966	0.887
MAE	**ANN**	0.820	0.770	0.740	**0.613**
	SVR	**0.703**	0.749	0.705	**0.613**
	MLR	0.713	0.766	0.715	0.622
	GPR	0.714	0.818	0.715	0.645
	M5P	0.787	**0.734**	**0.688**	0.637

Table 4. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) achieved by the MCE using two aggregation techniques: average and median; and five base learners: artificial neural network (ANN), support vector regression (SVR), multiple linear regression (MLR), Gaussian processes for regression (GPR), and model tree (M5P). The best performance achieved by each of the approaches for each base learner is marked with bold. The overall best performance for each base learner is marked with a gray background.

	RMSE		MAE	
Base learner	Average	Median	Average	Median
SVR	0.851	**0.825**	0.613	**0.601**
ANN	0.850	**0.840**	0.613	**0.594**
MLR	0.854	**0.830**	0.622	**0.610**
GPR	0.883	**0.872**	0.645	**0.637**
M5P	**0.887**	0.893	0.637	**0.633**

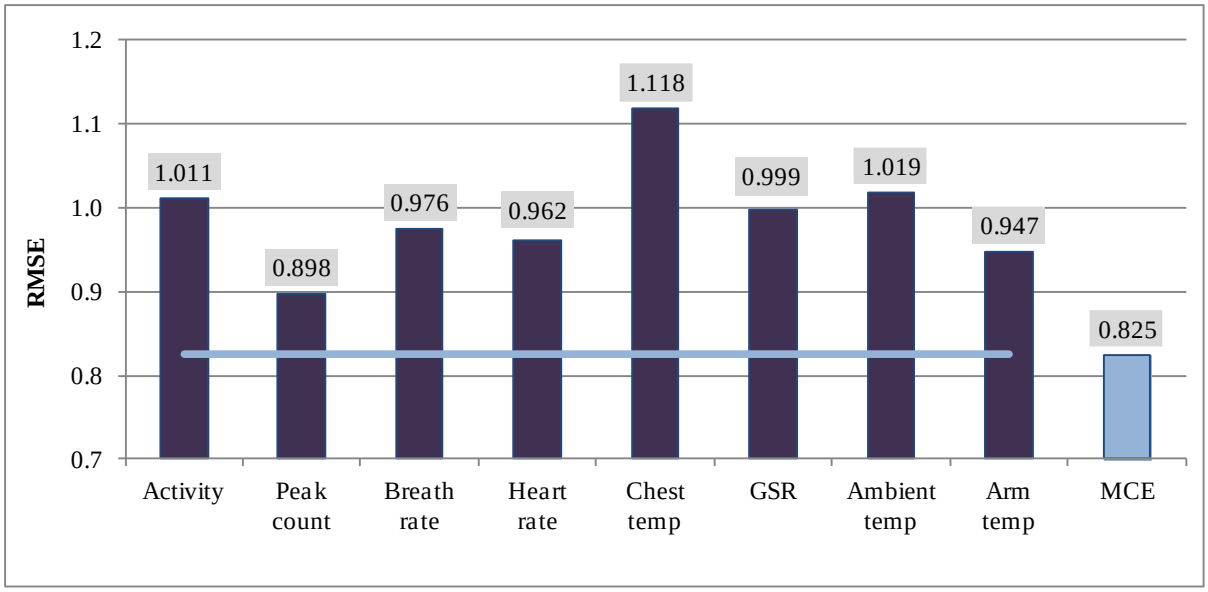


Figure 4. Root Mean Square Error (RMSE) for the MCE's EE estimation compared to each of the context models (base learners) used individually.

Figure 5 shows a scatter plot comparing the ground-truth and estimated MET values for different activities. Three approaches are compared: our MCE approach, the EE output of the BodyMedia sensor, and the ANN trained on chest-accelerometer data only (ANN-Acc). The results show that in general, the estimations of the MCE better match the true Cosmed values (the diagonal line in Figure 5) for almost all of the activities. The BodyMedia sensor has comparably good performance for the sedentary activities and for the more dynamic, exercise activities (walking, cycling, running), which is probably because the device is intended for physically active users. On the other side, for everyday light and moderate household activities the performance is significantly worse than the MCE's estimations. In addition, the results in Figure 5 show that the ANN-Acc approach largely underestimates the METs for the dynamic activities, especially for the cycling activity. This was in a way expected, because this method uses only the torso acceleration, while the cycling activity is an activity that does not include a lot of torso movement, but has a relatively high MET value.

The results achieved by the ANN-Acc approach show that acceleration information alone is not sufficient for accurate EE estimation. This is especially notable during the running activity and activities that are not characterized by high accelerations, but are still EE demanding, e.g., cycling (light and intense) and moderate-to-vigorous intensity household activities (digging, scrubbing the floor, etc.). These results are in accordance with the findings of Trost et al. [1], who also showed the highest RMSE values during these activities, except for the cycling activity which was completely omitted in their scenario. Staudenmayer et al. [14] also investigated and showed that worse results should be expected during the cycling activity when only torso-placed accelerometer is used to estimate the EE.

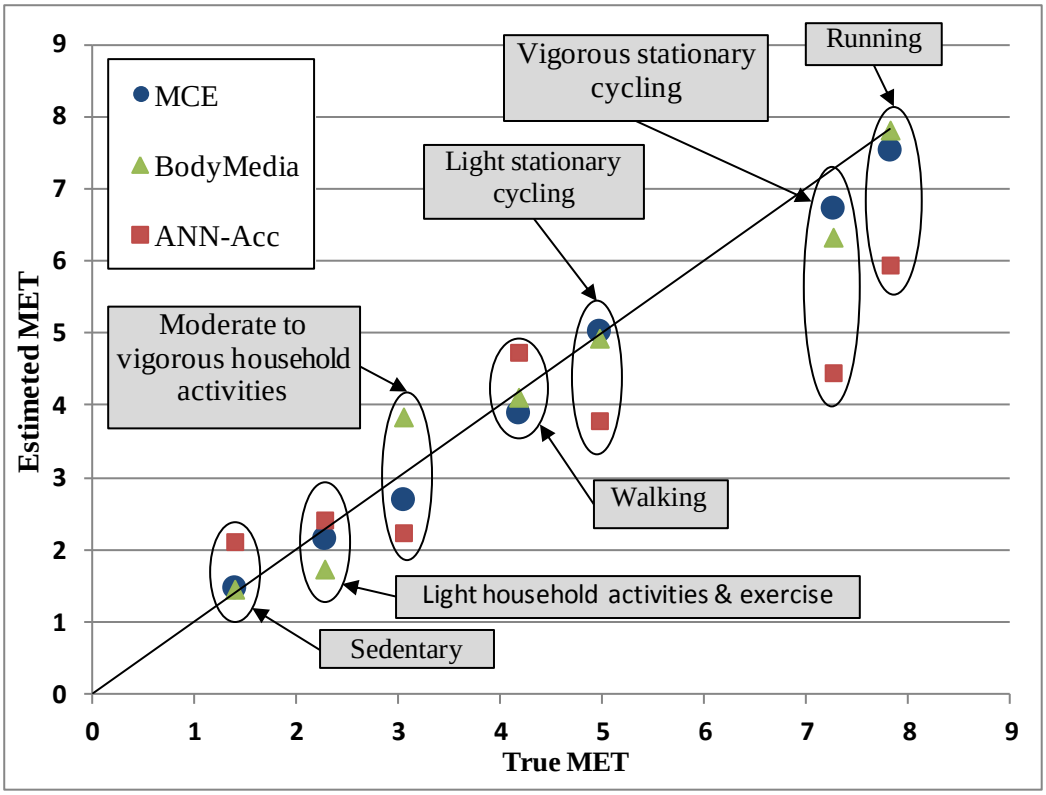


Figure 5. Measured and estimated METs for different types of activities using the MCE approach, EE output of the BodyMedia sensor and ANN trained on chest-accelerometer data only (ANN-Acc).

The comparison in Figure 5 has a drawback because it averages the estimated EE over one type of activity, and thus allows the errors of the methods (underestimations and overestimations) to cancel each other. For this reason we further analyzed the performance using the MAE and RMSE.

Table 5 presents the results for the MCE, BodyMedia and ANN-Acc EE estimations for all the activities and for different activity types individually. When calculated for all the activities, the MCE has significantly lower MAE and RMSE compared to the BodyMedia and ANN-Acc. Per-activity analysis shows that the MCE also has significantly lower RMSE and MAE for all activity types compared to the BodyMedia and ANN-Acc approach: on average 0.45 (RMSE) / 0.25 (MAE) lower than the BodyMedia, and 0.94 (RMSE) / 0.67 (MAE) lower than the ANN-Acc. The averaging issue (canceling the errors), which is present in Figure 5, is confirmed with the running activity in Table 5. That is, according to Figure 5 BodyMedia better match the EE estimation, however the RMSE and MAE statistics in Table 5 showed that MCE achieves significantly lower errors: 1.27 difference in RMSE and 0.79 difference in MAE.

The difference in performance of the ANN-Acc compared to the other two (MCE and BodyMedia) additionally confirms that acceleration information is not sufficient for accurate EE estimation. These results are in accordance with the findings of Lester et al. [21], Liu et al. [17], and Vyas et al. [3], who showed that by using multiple sensors one can overcome this problem.

Table 5. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) for the MCE's EE estimation compared to the BodyMedia sensor and the ANN-Acc regression model. The best performance for each activity type is marked with bold.

	Activities	MCE	BodyMedia	ANN-Acc
RMSE	**Overall activities**	**0.825**	1.326	1.763
	Sedentary	**0.571**	0.957	1.374
	Light HH & exercise	**0.807**	1.248	1.236
	Mod-Vig HH & sports	**1.094**	1.938	1.519
	Walking	**0.992**	1.165	1.203
	Cycling light	**0.932**	1.250	2.004
	Cycling vigorous	**1.205**	1.778	3.988
	Running	**1.192**	2.458	3.351
MAE	**Overall activities**	**0.601**	0.848	1.266
	Sedentary	**0.410**	0.490	0.950
	Light HH & exercise	**0.630**	1.000	0.960
	Mod-Vig HH & sports	**0.880**	1.560	1.140
	Walking	**0.770**	0.830	0.960
	Cycling light	**0.670**	1.010	1.730
	Cycling vigorous	**0.940**	1.290	3.800
	Running	**0.970**	1.760	3.110

6. Conclusions

This study presented a novel context-based ensemble method called MCE. The method was applied to the task of human EE estimation using multiple sensors. It builds upon the work of Crouter et al. [18], Lester et al. [21], and Vyas et al. [3], who showed that single-regression models cannot accurately estimate the EE over a range of activities, and that using multiple models based on the context (in their case the activity) significantly improves the EE estimation. It goes a step further in that it uses not only the activity as a context, but multiple contexts (HR, BR, GSR, etc.), resulting in an ensemble of models invoked for the contexts in which the participant is at a particular moment.

The presented MCE approach has a number of strengths. First, the novel reasoning with the use of multiple contexts enables more accurate and more context-specific EE estimation compared to conventional single-regression approaches (linear models, non-linear regression models, ANN-Acc), conventional ensemble approaches (Bagging and Random subspace), and the state-of-the-art BodyMedia device. Second, we showed that using multiple contexts can provide better accuracy than using only the activity of the user as the context. Third, our methodology is independent of the ML algorithm used for training; therefore, any algorithm that performs well on a given problem can be used. This may be beneficial if limited

processing power or memory is available, since simple algorithms such as linear regression can be used. It also means that the MCE approach can be applied to classification problems. In our previous studies we already showed that the context-based reasoning significantly improves the accuracy in activity recognition domain [41] (by using classification models) and in the fall detection domain (by using expert rules) [30].

There are also a few limitations that warrant consideration. First, since it is not easy to obtain valid, multi-sensor measurements useful for EE estimation, our method was developed using data from a limited number of people in controlled activity trials. Since ML and pattern recognition algorithms perform best when applied to population groups and/or activities that are identical or similar to those used to train the model, it remains an open question whether the models developed in the present study perform acceptably in independent samples of people performing similar or different activities remains. This issue is also relevant when comparing the MCE to the EE output of the BodyMedia sensor, whose EE estimation model was trained on a scenario different from ours (but not in the case of the ANN-Acc approach, where the model was trained and tested the same way as the MCE). Second, some may argue that the improvement in the EE-estimation accuracy is not worth the trouble of introducing such a complex methodology. However, once the context structure is defined and the models trained, the use is simple and requires relatively low computational power. The results show that the difference in the errors – if they do not cancel each other out – can amount to several hundred calories per day. This is probably most valuable for people who are particularly interested in precisely matching the caloric intake and output (because of engaging in certain sports or calorie restriction lifestyle, suffering from diabetes etc.).

In the future we plan to implement and release the MCE approach as readily usable software package or a plug-in for the WEKA ML toolkit [28]. First, this will make it accessible to researchers and practitioners from various areas and remove its complexity as a barrier to use. And second, it will make it easy to test it on various ML problems. While the approach was developed specifically for ambient-intelligence and healthcare problems, where humans are measured with sensors, it can in principle be used on any ML problem. It may no longer be possible to interpret (some of) the features as contexts, and contexts there may not be as important as in ambient intelligence and healthcare, but the MCE approach may still perform well.

ACKNOWLEDGMENTS

This work was partly supported by the Slovene Human Resources Development and Scholarship funds and partly by the CHIRON project – ARTEMIS Joint Undertaking, under grant agreement No. 2009-1-100228.

REFERENCES

[1] Trost SG, Wong WK, Pfeiffer KA, Zheng Y. Artificial neural networks to predict activity type and energy expenditure in youth. *Med Sci Sports Exerc*. 2012 Sep;44(9):1801–9.

[2] Luštrek M, Cvetković B, Kozina S. Energy expenditure estimation with wearable accelerometers. In: *Proceedings of the International Symposium on Circuits and Systems, ISCAS*; 2012 May 20-23: Seoul (Korea); 2012. p. 5–8.

[3] Vyas N, Farringdon J, Andre D, Stivoric J. Machine learning and sensor fusion for estimating continuous energy expenditure. In: *Proceedings of the 23rd Conference on Innovative Applications of Artificial Intelligence, IAAI;* 2011, Aug 9-11: San Francisco (California, USA), 2011, p. 1613–1620.

[4] Dey A, Salber D, Abowd G, Futakawa M. The conference assistant: Combining context awareness with wearable computing. In: *Proceedings of the 3rd International Symposium on Wearable Computing, ISWC*; 1999 Oct 18-19: San Francisco (California, United States of America); 1999. p. 21–28.

[5] Freedson PS, Pober D, Janz KF. Calibration of accelerometer output for children. *Med Sci Sports Exerc*. 2005; 37(11):S523–30.

[6] Hendelman D, Miller K, Bagget C, Debold E, Freedson PS. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc*. 2000 Sep;32(9):S442–9.

[7] Montoye HJ, Washburn R, Servais S, Ertyl A, Webster JG, Nagle FJ. Estimation of energy expenditure by a portable accelerometer. *Med Sci Sports Exerc*. 1983; 15(5):403–7.

[8] Swartz A, Strath SJ, Bassett DJ, O'Brien W, King GA, Ainsworth BE. Estimation of energy expenditure using CSA accelerometers at hip and wrist sites. *Med Sci Sports Exerc*. 2000 Sep;32(9):S450–6.

[9] Fruin ML, Rankin JW. Validity of a multi-sensor armband in estimating rest and exercise energy expenditure. *Med Sci Sports Exerc*. 2004; 36(6):1063–9.

[10] Bassett DR, Ainsworth BE, Swartz AM, Strath SJ, O'Brien WL, King GA. Validity of four motion sensors in measuring moderate intensity physical activity. *Med Sci Sports Exerc*. 2000; 32(9):S471–S80.

[11] Rumo M, Amft O, Tröster G, Mäder U. A stepwise validation of a wearable system for estimating energy expenditure in field-based research. *Physiol Meas*. 2011; 32(12):1983–2001.

[12] Sazonova N, Browning RC, Sazonov E. Accurate Prediction of Energy Expenditure Using a Shoe-Based Activity Monitor. *Med Sci Sports Exerc*. 2011; 43(7):1312–21.

[13] Rothney MP, Neumann M, Béziat A, Chen KY. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. *J Appl Physiol*. 2007; 103(4):1419–27.

[14] Staudenmayer J, Pober D, Crouter S, Bassett D, Freedson PS. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol*. 2009; 107:1300–7.

[15] Liu S, Gao RX, John D, Staudenmayer J, Freedson P. Multi-sensor data fusion for physical activity assessment. *IEEE Trans Biomed Eng*. 2012; 59(3):687–96.

[16] John D, Liu S, Sasaki JE, Howe CA, Staudenmayer J, Gao RX, Freedson PS. Calibrating a novel multi-sensor physical activity measurement system. *Physiol Meas*. 2011; 32(9):1473–89.

[17] Liu S, Gao RX, Freedson PS. Computational Methods for Estimating Energy Expenditure in Human Physical Activities. *Med Sci Sports Exerc*. 2012 Nov;44(11):2138–46.

[18] Crouter SE, Clowers KG, Bassett DR Jr. A novel method for using accelerometer data to predict energy expenditure. *J Appl Physiol*. 2006; 100(4):1324–31.

[19] Tapia EM. Using machine learning for real-time activity recognition and estimation of energy expenditure [dissertation]. Massachusetts Institute of Technology, MIT, 2008, p. 493.

[20] Albinali F, Intille SS, Haskell, W, Rosenberger, M. Using wearable activity type detection to improve physical activity energy expenditure estimation. In: *Proceedings of the 12th ACM international conference on Ubiquitous computing, Ubicomp*, 2010. p. 311–320.

[21] Lester J, Hartung C, Pina L, Libby R, Borriello G, Duncan G. Validated caloric expenditure estimation using a single body-worn sensor. In: *Proceedings of the 11th ACM international conference on Ubiquitous computing, Ubicomp*; 2009 Sep 30-Oct 03: Orlando (Florida, United States of America); 2009. p. 225–234.

[22] Gjoreski H, Kaluža B, Gams M, Milić R, Luštrek M. Ensembles of multiple sensors for human energy expenditure estimation. In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, UbiComp*, 2013 Sep 8-12: Zurich (Switzerland); 2013 p. 359–362.

[23] Lee JM, Kim Y, Welk GJ. Validity of Consumer-Based Physical Activity Monitors. *Med Sci Sports Exerc*. 2014 Feb 5 [Epub ahead of print].

[24] Wang Y, Witten I. Induction of model trees for predicting continuous classes. In: *Proceedings of the 9th European Conference on Machine Learning, ECML*; 1997 Apr 23-25: Prague (Czech Republic); 1997. p. 128–137.

[25] Breiman L, "Bagging Predictors," Machine Learning; 1996. 24(2): p. 123–140.

[26] Ho TK, The Random Subspace Method for Constructing Decision Forests. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1998, 20(8): p. 832-844.

[27] Zhou ZH. *Ensemble Methods: Foundations and Algorithms*. Boca Raton (FL): Chapman and Hall/CRC, 2012, p. 15.

[28] Witten I, Frank E, Hall MA. *Data Mining: Practical machine learning tools and techniques*. 3rd ed. Burlington (MA): Morgan Kaufmann, 2011, p. 403.

[29] Gjoreski H, Luštrek M, Gams M. Accelerometer Placement for Posture Recognition and Fall Detection. In: *Proceedings of the 7th International Conference on Intelligent Environments;* 2011 Jul 25-28: Nottingham (United Kingdom); 2011. p. 47–54.

[30] Gjoreski H, Gams M, Luštrek M. Context-based fall detection and activity recognition using inertial and location sensors. Journal of Ambient Intelligence and Smart Environments (JAISE), In press 2014.

[31] Mannini, A, Sabatini, AM. Machine Learning Methods for Classifying Human Physical Activities from on-body sensors. *Sensors*, 2010; 10:1154–1175.

[32] Kozina S, Gjoreski H, Gams M, Luštrek M. Three-layer activity recognition combining domain knowledge and meta-classification. *Journal of medical and biological engineering*, 2013, 33(4): 406–414.

[33] Kozina S, Gjoreski H, Gams M, Luštrek M. Efficient Activity Recognition and Fall Detection Using Accelerometers. In: Botía J, Álvarez-García JA, Fujinami K, Barsocchi P, Riedel T, editors. *Evaluating AAL Systems Through Competitive Benchmarking,* Communications in Computer and Information Science; 2013, p. 13–23.

[34] Gjoreski H, Kozina S, Gams M, Luštrek M. RAReFall – Real-time Activity Recognition and Fall Detection System. In: *Proceedings of the International Conference on Pervasive Computing and Communications Workshops, PERCOM;* 2014 Mar 24-28: Budapest (Hungary); 2014. p. 145–147.

[35] Venables WN, Ripley BD. Modern Applied Statistics with S. New York (NY): Springer-Verlag; 2002. p. 7.

[36] David A. Freedman (2009). Statistical Models: Theory and Practice. Cambridge University Press. p. 26.

[37] Drucker H, Burges CJC, Kaufman L, Smola AJ, Vapnik VN. Support Vector Regression Machines, Advances in Neural Information Processing Systems, NIPS; 1996, p. 155–161.

[38] David J.C. Mackay (1998). Introduction to Gaussian Processes. Dept. of Physics, Cambridge University, UK.

[39] Russell SJ, Norvig P. Artificial Intelligence: A Modern Approach. 3. Upper Saddle River (NJ): Prentice Hall; 2010. p. 729.

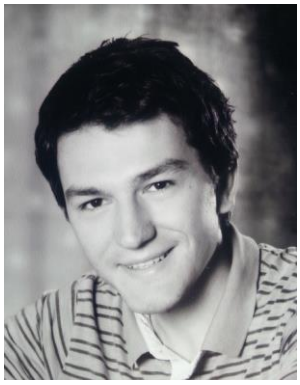
[40] Dietterich TG. Ensemble methods in Machine Learning. In Proceedings of the first International Workshop on Multiple Classifier Systems. 2000; p. 1-15.

[41] Gjoreski H, Kozina S, Luštrek M, Gams M. Using multiple contexts to distinguish standing from sitting with a single accelerometer. European Conference on Artificial Intelligence (ECAI), 2014, p. 1015-1016.

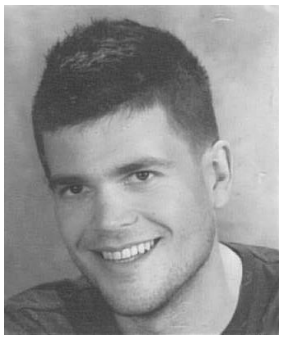
Abbreviations

MCE	=	Multiple Contexts Ensemble
EE	=	Energy Expenditure
HR	=	Heart Rate
BR	=	Breath Rate
GSR	=	Galvanic Skin Response
MET	=	Metabolic Equivalent of a Task
MLR	=	Multiple Linear Regression
SVR	=	Support Vector machine for Regression
GPR	=	Gaussian Processes for Regression
M5P	=	Model trees
ANN	=	Artificial Neural Network
ANN-Acc	=	ANN trained only on accelerometer data
RMSE	=	Root Mean Squared Error
MAE	=	Mean Absolute Error

Author bios



Dr. Hristijan Gjoreski is a researcher at the Department of Intelligent Systems at Jožef Stefan Institute. His research interests include context-based reasoning, wearable computing and ambient intelligence. He holds a Ph.D. degree in Information and Communication Technologies from the Jožef Stefan International Postgraduate School. Contact him at hristijan.gjoreski@ijs.si.



Dr. Boštjan Kaluža is the head of the Agents Group at the Department of Intelligent Systems at Jožef Stefan Institute. His research interests include agent and multi-agent systems in general, the analysis of agent behavior in ambient-intelligence and security domains, machine learning and heuristic search. He holds a Ph.D. degree in New Media and e-Science from the Jožef Stefan International Postgraduate School. Contact him at bostjan.kaluza@ijs.si.

Prof. Dr. Matjaž Gams is the head of the Department of Intelligent systems at Jožef Stefan Institute, and professor at the University of Ljubljana and Jožef Stefan Postgraduate School. His research interests include ambient intelligence, machine learning, agents, hybrid learning and reasoning. He holds a Ph.D. degree in Computer and Information Science from the University of Ljubljana. Contact him at matjaz.gams@ijs.si.

Dr. Radoje Milić, MD, is the head of the head of Exercise Physiology Lab at the Institute of Sport – Faculty of Sport, University of Ljubljana, Slovenia. His research interests include sports medicine, exercise physiology and exercise performance analysis. Contact him at radoje.milic@fsp.uni-lj.si

Dr. Mitja Luštrek is the head of the Ambient Intelligence Group at the Department of Intelligent Systems at Jožef Stefan Institute. His main research interest is ambient intelligence, particularly the analysis of human behavior using sensor data. He holds a Ph.D. degree in Computer and Information Science from the University of Ljubljana. Contact him at mitja.lustrek@ijs.si.