

10-1-2016

# Developing Novel Machine Learning Algorithms to Improve Sedentary Assessment for Youth Health Enhancement.

Gowtham Kumar Golla

Jordan A. Carlson  
*Children's Mercy Hospital*

Jun Huan

Jacqueline Kerr

Tarrah Mitchell

*See next page for additional authors*

Follow this and additional works at: <https://scholarlyexchange.childrensmercy.org/papers>

 Part of the [Community Health and Preventive Medicine Commons](#), [Investigative Techniques Commons](#), and the [Pediatrics Commons](#)

---

## Recommended Citation

Golla, Gowtham Kumar; Carlson, Jordan A.; Huan, Jun; Kerr, Jacqueline; Mitchell, Tarrah; and Borner, Kelsey, "Developing Novel Machine Learning Algorithms to Improve Sedentary Assessment for Youth Health Enhancement." (2016). *Manuscripts, Articles, Book Chapters and Other Papers*. 1065.

<https://scholarlyexchange.childrensmercy.org/papers/1065>

This Article is brought to you for free and open access by SHARE @ Children's Mercy. It has been accepted for inclusion in Manuscripts, Articles, Book Chapters and Other Papers by an authorized administrator of SHARE @ Children's Mercy. For more information, please contact [bpfannenstiel@cmh.edu](mailto:bpfannenstiel@cmh.edu).

---

**Creator(s)**

Gowtham Kumar Golla, Jordan A. Carlson, Jun Huan, Jacqueline Kerr, Tarrah Mitchell, and Kelsey Borner



Published in final edited form as:

IEEE Int Conf Healthc Inform. 2016 October ; 2016: 375–379. doi:10.1109/ICHI.2016.67.

## Developing Novel Machine Learning Algorithms to Improve Sedentary Assessment for Youth Health Enhancement

Gowtham Kumar Golla<sup>\*</sup>, Jordan A. Carlson<sup>†</sup>, Jun Huan<sup>\*</sup>, Jacqueline Kerr<sup>‡</sup>, Tarrah Mitchell<sup>§</sup>, and Kelsey Borner<sup>§</sup>

<sup>\*</sup>Department of Electrical Engineering and Computer Science, University of Kansas, 1450 Jayhawk Blvd, Lawrence, KS 66045

<sup>†</sup>Center for Children's Healthy Lifestyles and Nutrition, Children's Mercy Hospital

<sup>‡</sup>Department of Family Medicine and Public Health, University of California, San Diego, 9500 Gilman Dr, La Jolla, CA 92093

<sup>§</sup>Department of Psychology, University of Kansas, 1450 Jayhawk Blvd, Lawrence, KS 66045

### Abstract

Sedentary behavior of youth is an important determinant of health. However, better measures are needed to improve understanding of this relationship and the mechanisms at play, as well as to evaluate health promotion interventions. Wearable accelerometers are considered as the standard for assessing physical activity in research, but do not perform well for assessing posture (i.e., sitting vs. standing), a critical component of sedentary behavior. The machine learning algorithms that we propose for assessing sedentary behavior will allow us to re-examine existing accelerometer data to better understand the association between sedentary time and health in various populations. We collected two datasets, a laboratory-controlled dataset and a free-living dataset. We trained machine learning classifiers separately on each dataset and compared performance across datasets. The classifiers predict five postures: sit, stand, sit-stand, stand-sit, and stand/walk. We compared a manually constructed Hidden Markov model (HMM) with an automated HMM from existing software. The manually constructed HMM gave more F1-Macro score on both datasets.

### I. INTRODUCTION

Sedentary behavior is defined as any waking activity with a sitting or reclining posture that produces less than 1.5 metabolic equivalents of energy expenditure [1,2]. The amount of time youth spend in sedentary pursuits has been associated with poor health indicators such as obesity, elevated blood pressure, and high cholesterol [3,4]. These associations were independent of moderate-to-vigorous physical activity, so even physically active children can be affected negatively by too much sedentary time. More recently, the patterns in which people accumulate sedentary time were found to have negative impacts on health over and above total sedentary time [5]. For example, prolonged periods of sedentary time for more than 30 minutes, and few breaks from sedentary time (i.e., standing up periodically), have been associated with several cardiometabolic risk factors. Although sedentary behavior is established as an important determinant of health, better measures are needed to improve

understanding of this relationship and the mechanisms at play, as well as to evaluate efforts to reduce sedentary time.

Accelerometers are small devices worn on the hip or wrist and are considered to be the gold standard for measuring physical activity in health research [6,7]. Although traditional cut point-based scoring methods for accelerometers are commonly used to measure sedentary time [8,9], they were not designed to capture posture (i.e., sitting vs. standing), a critical component of sedentary behavior. Traditional cut-point methods involve inferring sedentary time based on whether the accelerometer counts for a given time interval (e.g., 15 seconds) exceed a certain threshold. Research has shown that this approach is acceptable for estimating total minutes sedentary time but has limited validity for detecting sit-stand transitions [10,11].

Although other devices exist for accurately assessing posture and sit-to-stand transitions (e.g., activPAL), these devices are expensive and do not have a history of use in large health studies. Hip-worn accelerometers, on the other hand, have been used in numerous large studies with health outcome data [e.g., 12–13]. Thus, an accurate machine learning-based sedentary classifier could be employed to existing data to more accurately investigate sedentary behavior and health outcomes.

## II. BACKGROUND

Recently in public health community, using of machine learning algorithms is increasing. Having a valid machine learning-based method for assessing sedentary metrics from accelerometer data will allow better understanding of both the impacts of sedentary behavior on health as well as potential mechanisms driving these associations, which are public health priorities.

Ellis et al. [14,15], Kate et al. [16], Gyllensten et al. [17], and Bonomi et al. [18] used machine learning with accelerometers to predict activity types (e.g., walking, sitting, bicycling). Techniques with good performance included support vector machines [16,17], neural networks [16,17], hidden Markov models (HMM) [14], and decision trees [14,15,18]. None of these algorithms were designed or tested for their ability to capture posture or sit-stand transitions. Other limitations of these studies were that the algorithms performed at the minute-level [14,15] and the data were from lab-based activities [16,17,18].

In [17], it was observed that the machine learning classifiers applied on a laboratory-controlled dataset had low accuracy when applied on a free-living (i.e., participants perform their normal daily life) dataset, so free-living data are important for training such algorithms.

In this paper, we presented a clear empirical study of different machine learning algorithms. The present study aims to fill a critical research gap by testing machine learning algorithms for classifying posture and sit-to-stand transitions. We collected both laboratory controlled and free-living datasets, and applied different machine learning algorithms on each dataset using 1-second data. We compared algorithms across the datasets and demonstrated performance of an automated HMM vs. a manually constructed HMM.

### III. DATA COLLECTION

#### A. Participants and Procedures

**1) Laboratory-Controlled Dataset**—We recruited 9 study staff between the ages of 22 and 34 and had them engage in a protocol of sitting and standing behaviors. First, staff engaged in 5 seconds of standing followed by 5 seconds of sitting, and repeated this pattern for one minute. Next, staff engaged in 10 seconds of standing followed by 10 seconds of sitting, and repeated this pattern for one minute. Another staff member used a stop watch, synced to the device time, to direct the behavior and note any deviations in the protocol.

**2) Free-Living Dataset**—We recruited participants from two samples from a Midwestern community in the United States. The Adult Sample included 11 adult office workers (ages 25–63) at an academic medical center, and the Youth Sample included 9 youth (ages 10–17) who were patients in a behavioral weight-management program for pediatric obesity. Adult participants were enrolled in the study for one work day and were instructed to go about their day as usual, with the exception that they were asked to incorporate additional sit-stand transitions throughout the day to ensure that they were not sitting or standing for the entire data collection period. Youth participants were enrolled in the study for approximately 2 hours while attending an evening weight-management group session. The first hour of the weight-management session included group exercise activities and active games. The second hour was spent in the classroom. Weight-management staff were asked to incorporate additional sit-stand transitions during both hours of the session.

#### B. Measures

**1) Actigraph**—Participants wore a GT3X Actigraph accelerometer (ActiGraph, Pensacola, FL) on the right hip affixed to an elastic belt. Actigraphs are considered gold standard for physical activity assessment in population studies [21]. In the present study, the Actigraphs provided the data for the machine learning algorithm. Raw acceleration in g-force was recorded at 30hrz for each of the 3 axis (vertical, medio-lateral, and antero-posterior) and a vector magnitude. Accelerations counts per second, scored in ActiLife software, were derived for the laboratory-controlled dataset.

**2) activPAL**—Participants also wore an activPAL micro accelerometer (PAL Technologies, LTC) on the right thigh, affixed according to the manufacturers recommendations. The activPAL has good criterion validity for assessing posture (i.e., sitting, standing, lying) and sit-stand transitions as compared to direct observation. The event files produced from the activPAL software were used to create second-level files denoting, for each second, whether the participant was sitting, standing, standing and walking, in a sit-to-stand transition, or in a stand-to-sit transition [11,19,20]. This information served as the ground truth for developing the algorithms.

#### C. Data Processing

The activPAL and Actigraph devices were initialized on the same computer to provide time synchronization. A log was used to record when each participant put on and took off each device, and non-wear time was removed from the data. For Laboratory-Controlled dataset,

there is a total of 17.7 minutes (1065 seconds) of data. For free-living dataset, adult participants had a mean wear time of 376.9 (Standard Deviation = 77.4) minutes, and youth participants had a mean wear time of 92.1 (Standard Deviation = 16.5) minutes, for a total of 4898.3 minutes (293900 seconds) of data.

## IV. EXPERIMENTAL SETUP

We performed statistical and histogram analysis on both datasets.

### A. Laboratory-Controlled Dataset

This was a small dataset which contained 1065 records from 9 subjects. The features were acceleration counts per second from Axis1, Axis2, Axis3, and a Vector Magnitude. No additional features were extracted because we wanted to test algorithm performance when using the accelerometer count data provided by the ActiLife software. The classes to predict were: sit, stand, sit-stand and stand-sit.

Fig. 1. represents the Histogram of the class distribution of the data. Table. 1. shows the class distribution of the data. We divided the dataset into training data using 7 Subjects (104–110) with 838 samples, and testing data using 2 subjects (102,103) with 227 samples.

### B. Free-Living Dataset

This was a large dataset which contained 8,817,000 records from 11 adult and 9 youth subjects. The data included raw g force acceleration from Axis1, Axis2, and Axis3. Each record was a 33.3ms record (i.e., 30 records per second). Because of the large size of the dataset, we divided the data to smaller chunks based on the participant. The classes to predict were: sit, stand, sit-stand, stand-sit and stand\walk.

We applied the following methods to the data [22,23]. We used R programming language, R-3.3.1 Version and 64-bit, to perform these methods.

**1) Feature Extraction**—From the 3 initial g force metrics in the dataset, we extracted 21 features using a window of 30 records so that a unique value for each feature was derived for every 1-sec. The statistical features were: mean, standard deviation, coefficient of variation, median, minimum, maximum, 25th percentile, 75th percentile, correlations across each pair of axis, and mean, standard deviation, and gravity component yaw, pitch, and roll rotation. These features have been employed in previous studies that used machine learning for physical activity classification [14,15]. After applying feature extraction and aggregating the data to 1-sec, the dataset contained 293900 samples. We divided the dataset into training data using 7 adult and 5 youth subjects with 166400 samples, and testing data using 4 adult and 4 youth subjects with 127500 samples.

**2) Imbalance**—In fig. 2. we see that the dataset is biased on one class (i.e. sit). Furthermore, the class distribution is not uniform among the classes. With imbalanced data, the machine learning algorithms provide unfavorable results while predicting the classes. Thus, we used oversampling with replacement to correct the bias in the dataset.

Fig. 3. and Table. 3. represents class distribution after applying oversampling on the training dataset.

## V. EXPERIMENTS AND RESULTS

We classified posture in the laboratory-collected and free-living datasets by applying the machine learning algorithms: Support Vector Machine (SVM) [24], Random Forest (RF) [25], Conditional Random Fields (CRF) [26] and Hidden Markov Model (HMM) [27]. The machine learning algorithms were selected based on two unique characteristics of the data.

- The classes sit and stand have similar values for the features.
- The sit-stand transitions should be considered while predicting the correct posture between sit and stand.

SVM and RF predict the classes based on the values for features given at that time. CRF predicts the classes based on the values for features given at that time and the previous state of class. HMM predicts the sequence of observable classes based on the sequence of hidden states. The hidden states can be specified by start and transition probabilities. The emission probability of an observable class can be in any distribution.

We used scikit-learn, a machine learning package in python, 0.16.1 version, for running the algorithms. We trained algorithms using 10-fold cross validation for better results. We used the F1-macro metric for comparison as it calculates metrics for each label, finds their unweighted mean, and does not take class imbalance into account. The following sections explain the empirical analysis on the data.

### A. Default Model Parameters vs Tuning of Model Parameters

First, we ran algorithms with the default model parameters for each dataset. Then, we tuned model parameters for each algorithm to optimize the accuracy of predicting the postures. We compared performance of the algorithms with default model parameters and tuned the model parameters.

Fig. 4. and Fig. 5. show that by tuning the parameters, we can optimize the performance of the machine learning algorithm.

### B. Oversampling Vs No Oversampling

The free-living dataset is biased on one class (i.e. sit). We used oversampling with replacement to balance the training dataset. With this oversampling, we further optimized the performance of the machine learning algorithms. We found that we could predict the sit-stand transitions with improved accuracy when using the balanced training dataset.

Fig. 6. shows that by balancing the dataset using oversampling, we could optimize the performance of the machine learning algorithm.

### C. Manually Constructed HMM vs Automated HMM

To further increase the accuracy of prediction, we used scikit-learn hmmlearn package for training HMM on the datasets. We constructed HMM manually by calculating start, transition and Gaussian emission probabilities of the hidden states (i.e. postures). In scikit-learn the Gaussian emission of each state are the mean and covariances of the features. The manually constructed HMM performed better than the automated HMM.

Fig. 6. shows the comparison of performance between manually constructed HMM and automated HMM of scikit-learn package.

## VI. CONCLUSIONS

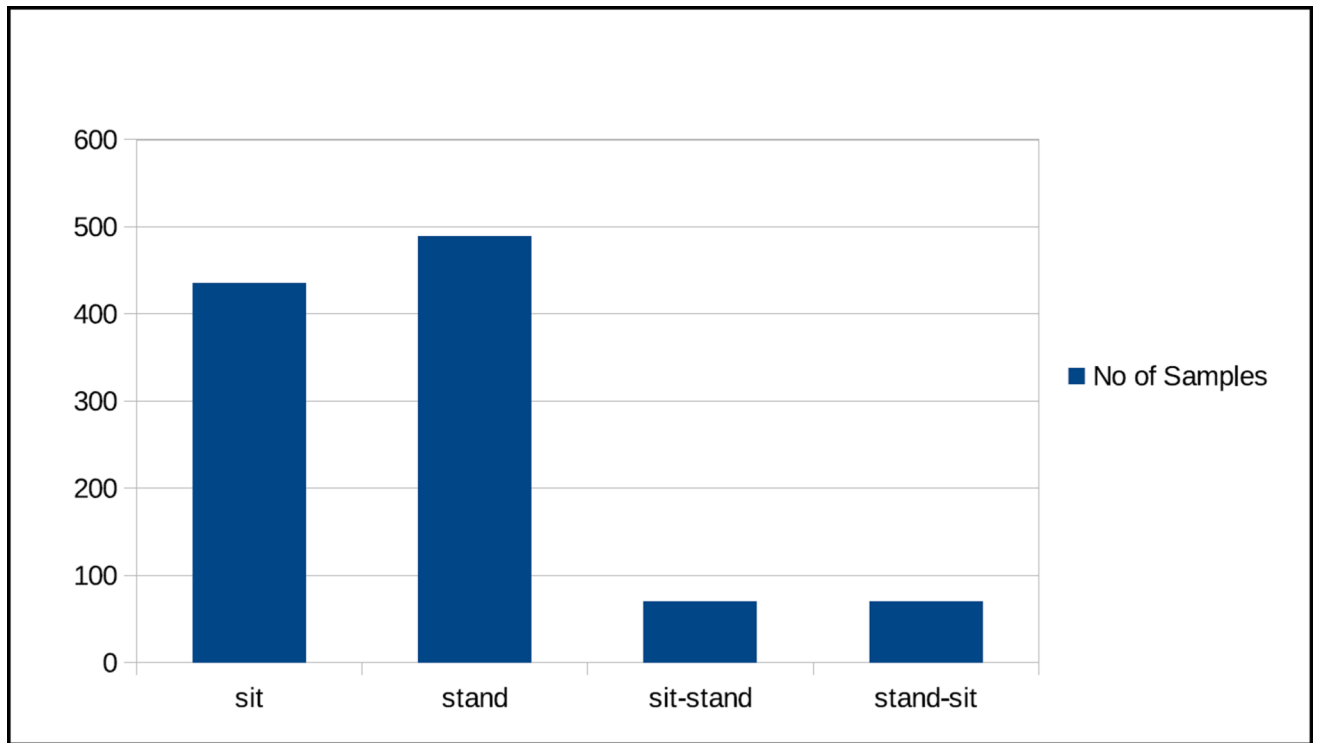
The main objective of this study was to develop a machine learning based method to accurately classify posture and sit-stand transitions in free-living behavior. To achieve this, we performed an empirical analysis of how different machine learning algorithms work on accelerometer data. We collected laboratory-controlled and free-living datasets. We applied machine learning algorithms on these datasets and showed that by carefully constructing HMM from the data, we can predict posture from accelerometer data with reasonable accuracy. Although the experiments are related to the datasets we collected and thus should be replicated, this study provides a novel approach to using machine learning algorithms for improving assessment of sedentary behavior. Applying machine learning algorithms to assess sedentary behavior in existing large health datasets will advance understanding of the impacts of sedentary behavior on health outcomes and potential approaches for minimizing the negative consequences of sedentary behavior.

## References

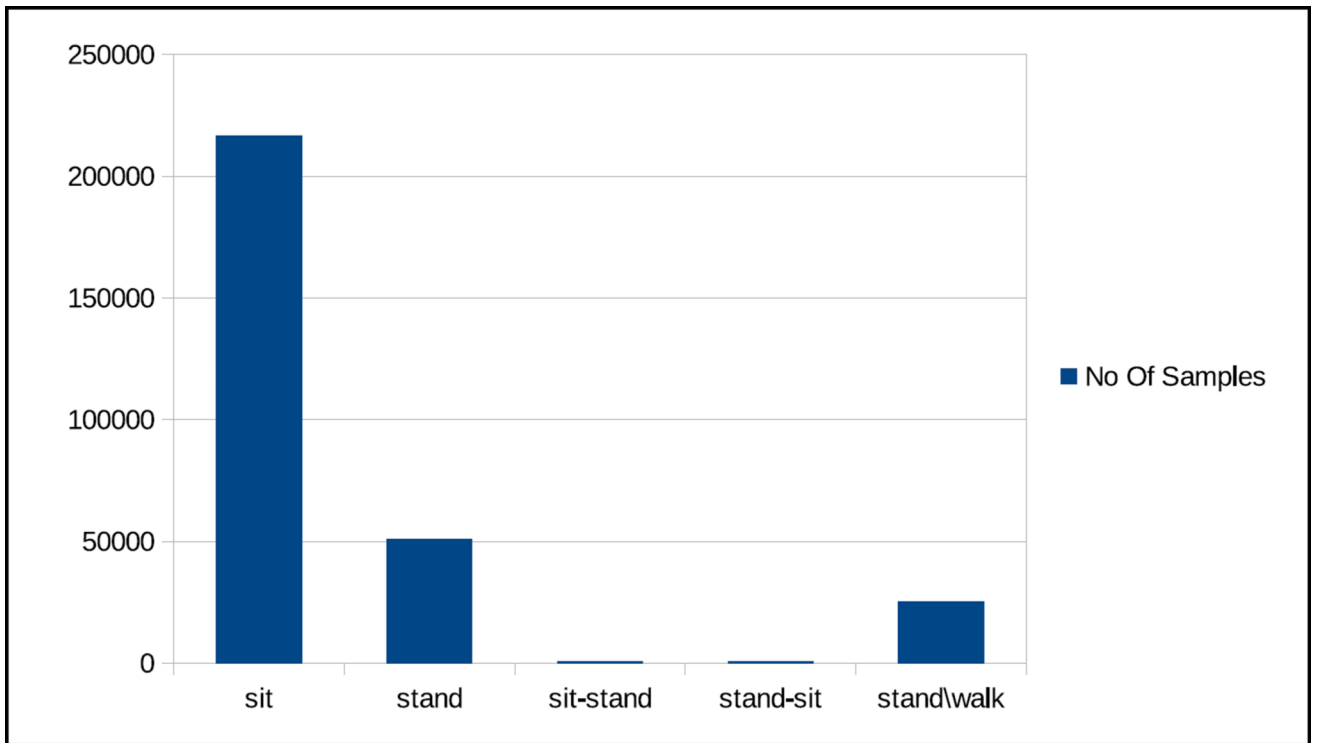
1. Pate RR, O'Neill JR, Lobelo F. The evolving definition of "sedentary". *Exercise and sport sciences reviews*. 2008; 36(4):173–178. [PubMed: 18815485]
2. Sedentary Behavior Research Network. Letter to the editor: Standardized use of the terms sedentary and sedentary behaviours. *Applied Physiology, Nutrition, and Metabolism*. 2012; 37(3):540–542. DOI: 10.1139/h2012
3. Ekelund U, Luan J, Sherar LB, Esliger DW, Griew P, Cooper A. Moderate to vigorous physical activity and sedentary time and cardiometabolic risk factors in children and adolescents. *JAMA: the journal of the American Medical Association*. 2012; 307(7):70412.
4. Tremblay MS, LeBlanc AG, Kho ME, Saunders TJ, Larouche R, Colley RC, Goldfield G, Gorber SC. Systematic review of sedentary behavior and health indicators in school-aged children and youth. *International*.
5. Healy GN, Dunstan DW, Salmon J, Cerin E, Shaw JE, Zimmet PZ, Owen N. Breaks in sedentary time: beneficial associations with metabolic risk. *Diabetes Care*. 2008; 31(4):661–666. [PubMed: 18252901]
6. Troiano RP, Berrigan D, Dodd KW, Masse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. *Medicine and Science in Sports and Exercise*. 2008; 40:181–188. [PubMed: 18091006]
7. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. *British journal of sports medicine*. 2014; 48(13):1019–1023. [PubMed: 24782483]
8. Atkin AJ, Gorely T, Clemes SA, Yates T, Edwardson C, Brage S, Biddle SJ. Methods of measurement in epidemiology: sedentary behaviour. *International journal of epidemiology*. 2012; 41(5):1460–1471. [PubMed: 23045206]



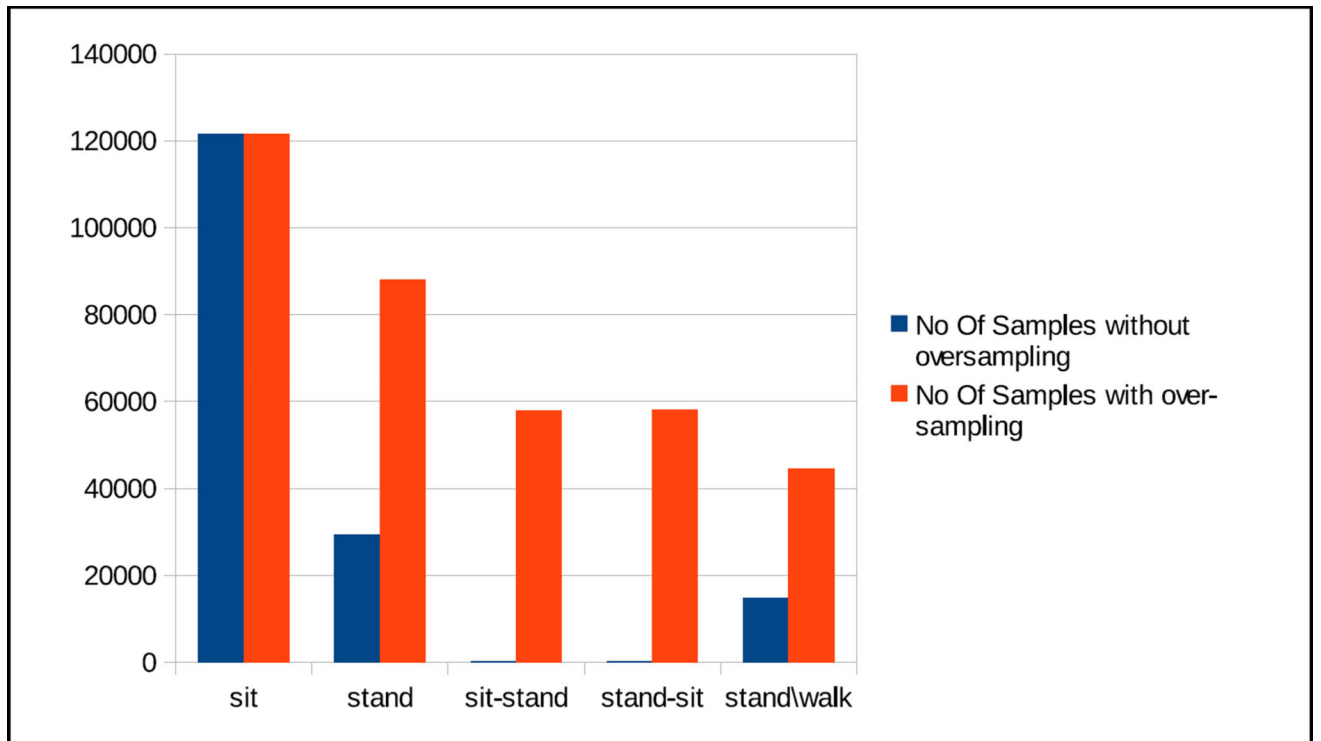
9. Matthews CE, Hagstrmer M, Pober DM, Bowles HR. Best practices for using physical activity monitors in population-based research. *Medicine and science in sports and exercise*. 2012; 44(1 Suppl 1):S68. [PubMed: 22157777]
10. Barreira TV, Zderic TW, Schuna JM, Hamilton MT, Tudor-Locke C. Free-living activity counts-derived breaks in sedentary time: Are they real transitions from sitting to standing? *Gait posture*. 2015; 42(1):70–72. [PubMed: 25953504]
11. Lyden K, Kozey-Keadle SL, Staudenmayer JW, Freedson PS. Validity of two wearable monitors to estimate breaks from sedentary time. *Medicine and science in sports and exercise*. 2012; 44(11): 2243. [PubMed: 22648343]
12. Ekelund U, Sardinha LB, Anderssen SA, Harro M, Franks PW, Brage S, Froberg K. Associations between objectively assessed physical activity and indicators of body fatness in 9-to 10-y-old European children: a population-based study from 4 distinct regions in Europe (the European Youth Heart Study). *The American journal of clinical nutrition*. 2004; 80(3):584–590. [PubMed: 15321796]
13. Katzmarzyk PT, Barreira TV, Broyles ST, Champagne CM, Chaput JP, Fogelholm M, Lambert EV. The international study of childhood obesity, lifestyle and the environment (ISCOLE): Design and methods. *BMC Public Health*. 2013; 13(1):1. [PubMed: 23280303]
14. Ellis K, Godbole S, Kerr J, Lanchriet G. Multi-sensor physical activity recognition in free-living. *Proc ACM Int Conf Ubiquitous Comput*. 2014:431–440. [PubMed: 26247061]
15. Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. *Physiological measurement*. 2014; 35(11):2191. [PubMed: 25340969]
16. Kate RJ, Swartz AM, Welch WA, Strath SJ. Comparative evaluation of features and techniques for identifying activity type and estimating energy cost from accelerometer data. *Physiological measurement*. 2016; 37(3):360. [PubMed: 26862679]
17. Gyllensten IC, Bonomi AG. Identifying types of physical activity with a single accelerometer: evaluating laboratory-trained algorithms in daily life. *Biomedical Engineering, IEEE Transactions on*. 2011; 58:9.
18. Bonomi AG, Goris A, Yin B, Westerterp KR. Detection of type, duration, and intensity of physical activity using an accelerometer. *Med Sci Sports Exerc*. 2009; 41(9):1770–1777. [PubMed: 19657292]
19. Aminian S, Hinckson EA. Examining the validity of the activPAL monitor in measuring posture and ambulatory movement in children. *IJBNPA*. 2012; 9:119. [PubMed: 23031188]
20. Davies G, Reilly J, McGowan A, Dall P, Granat M, Paton J. Validity, practical utility, and reliability of the activPAL in preschool children. *Medicine and Science in Sports and Exercise*. 2012; 44(4):761–768. [PubMed: 21983077]
21. Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exercise*. 2005; 37(11):S531–S543.
22. Preece, Stephen J., Goulermas, John Yannis, Kenney, Laurence PJ., Howard, David. A Comparison of Feature Extraction Methods for the Classification of Dynamic Activities From Accelerometer Data. *IEEE Transactions on Biomedical Engineering*. Oct; 2008 56(3):871–879. [PubMed: 19272902]
23. Batista, Gustavo EAPA., Prati, Ronaldo C., Monard, Maria Carolina. A study of the behavior of several methods for balancing machine learning training data. *Newsletter ACM SIGKDD Explorations Newsletter-Special issue on learning from imbalanced datasets Homepage archive*. Jun; 2004 6(1):20–29.
24. Smola, Alex J., Schlkopf, Bernhard. A tutorial on support vector regression. *Statistics and Computing archive*. Aug; 2004 14(3):199–222.
25. Breiman L. Random Forests. *Machine Learning*. 2001; 45(1):5–32.
26. Lafferty, J., McCallum, A., Pereira, F. 18th International Conf. on Machine Learning. Morgan Kaufmann: Conditional random fields: Probabilistic models for segmenting and labeling sequence data; p. 282-289.
27. Rabiner, Lawrence R. A tutorial on hidden Markov models and selected applications in speech recognition. 1989



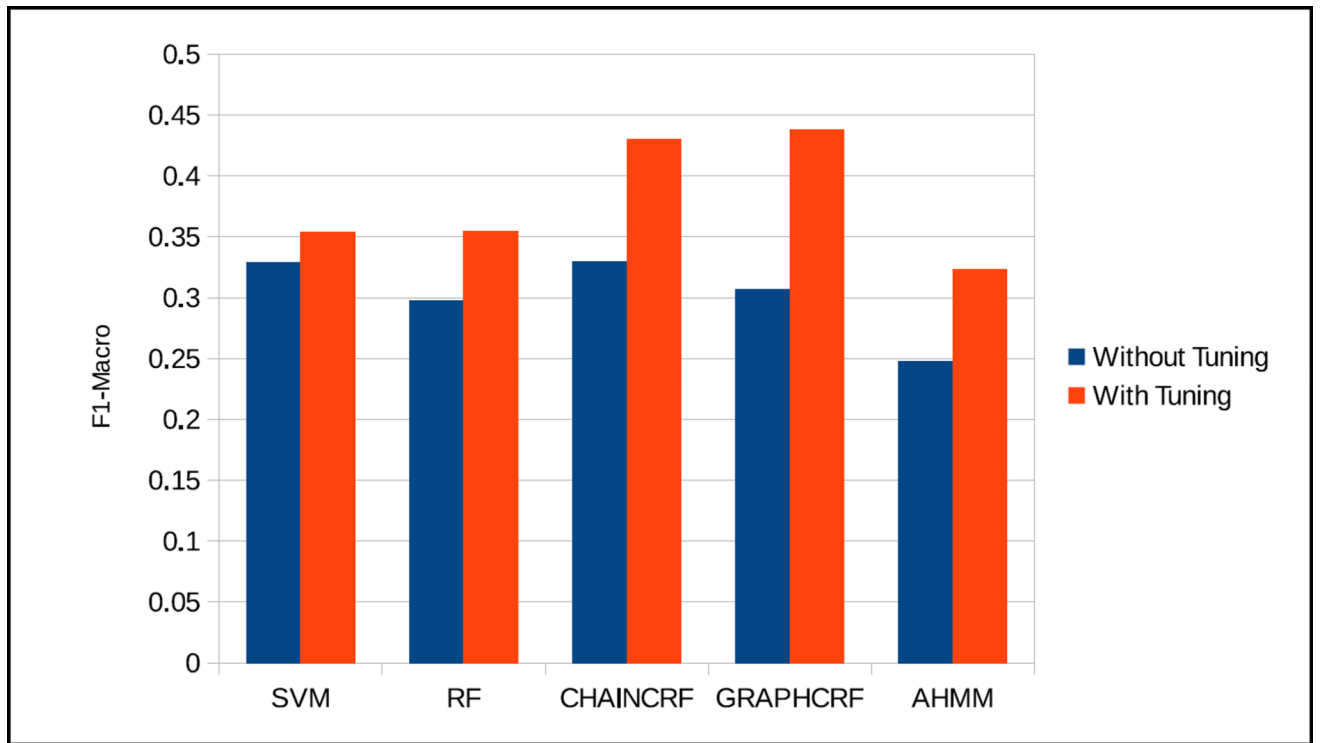
**Fig. 1.**  
Histogram for class distribution of laboratory-controlled dataset



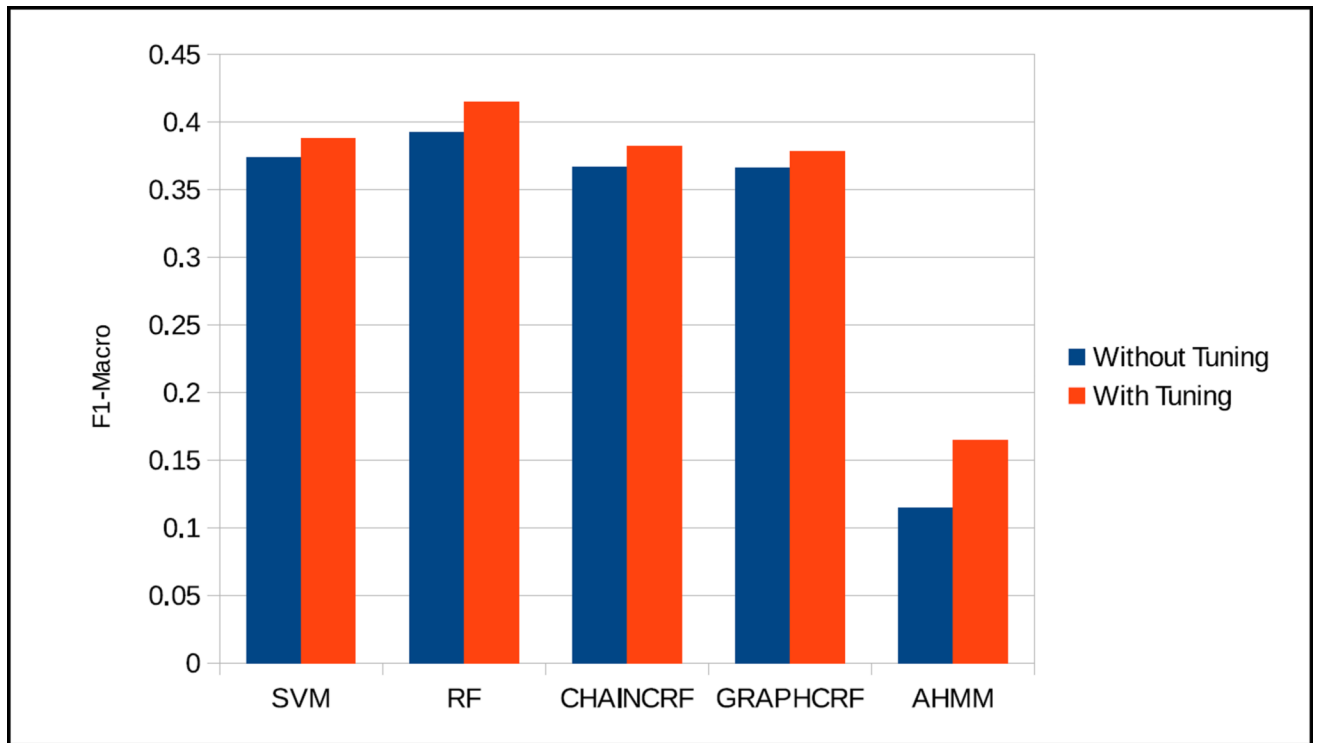
**Fig. 2.**  
Histogram for class distribution of free-living dataset



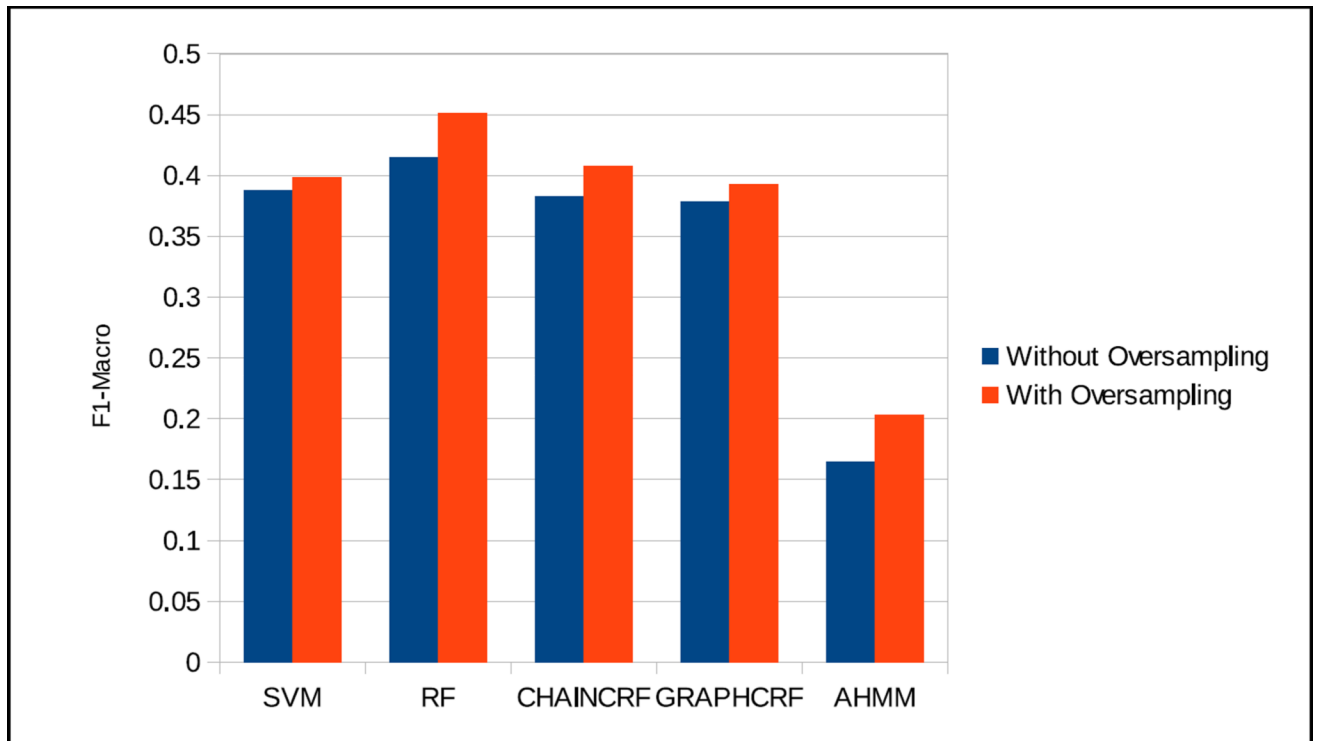
**Fig. 3.**  
Histogram for class distribution of Training dataset



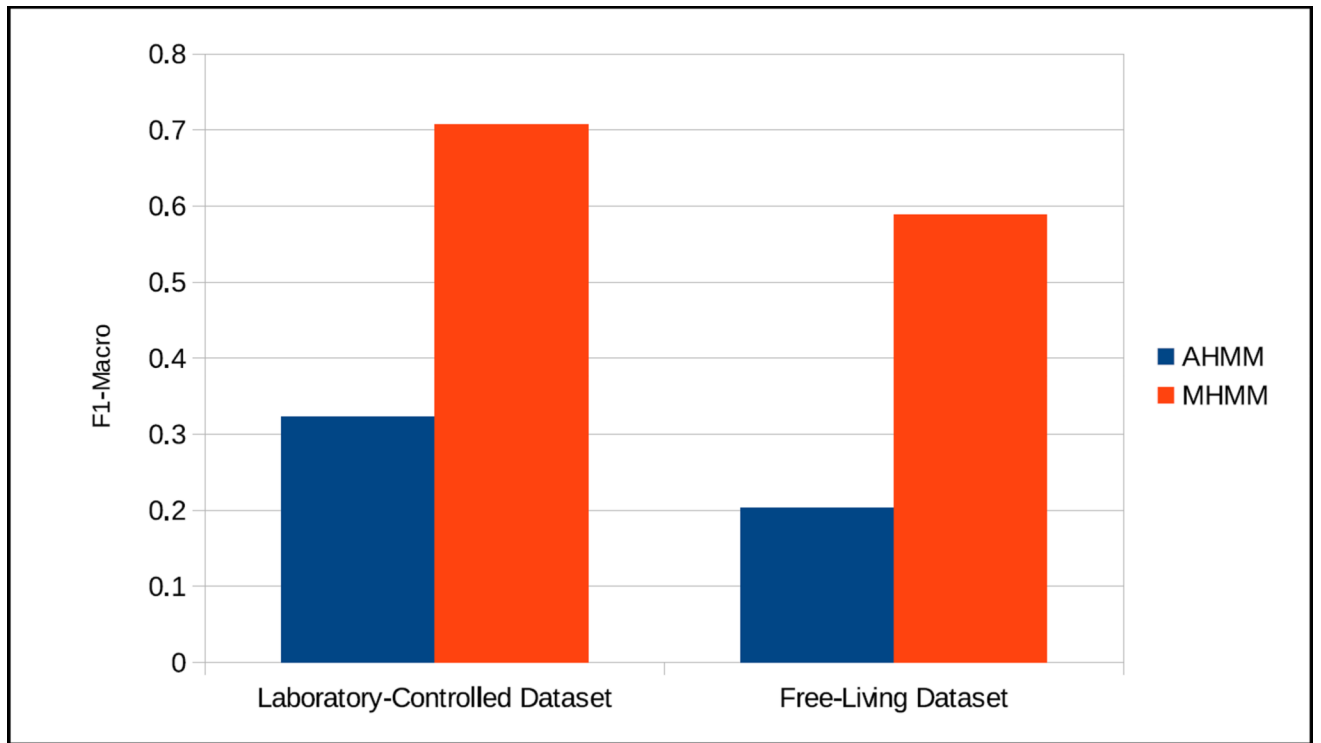
**Fig. 4.**  
Comparison with tuned parameters of laboratory-controlled dataset



**Fig. 5.**  
Comparison with tuned parameters of free-living dataset



**Fig. 6.**  
Comparison with oversampling of free-living dataset



**Fig. 7.**  
Comparison between manually constructed HMM and automated HMM



**TABLE I**

Class distribution of laboratory-controlled dataset

Class	No Of Samples
sit	435
stand	489
sit-stand	71
stand-sit	70

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE II**

Class distribution of free-living dataset

Class	No Of Samples
sit	216734
stand	50799
sit-stand	538
stand-sit	538
stand\walk	25291

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**TABLE III**

Class distribution of training dataset

Class	No Of Samples without oversampling	No Of Samples with oversampling
sit	121635	121635
stand	29333	87999
sit-stand	287	57974
stand-sit	288	58176
stand\walk	14857	44571

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript