Original Paper

Neural Network–Based Algorithm for Adjusting Activity Targets to Sustain Exercise Engagement Among People Using Activity Trackers: Retrospective Observation and Algorithm Development Study

Ramin Mohammadi[1,2], PhD; Mursal Atif[2], BA; Amanda Jayne Centi[2], PhD; Stephen Agboola[2,3,4], MD, PhD; Kamal Jethwani[2,3,4], MD, PhD; Joseph Kvedar[2,3,4], MD; Sagar Kamarthi[1,2], PhD

[1]Northeastern University, Boston, MA, United States

[2]Partners connected for health, Boston, MA, United States

[3]Massachusetts General Hospital, Department of Dermatology, Boston, MA, United States

[4]Harvard University, Harvard Medical School, Boston, MA, United States

Corresponding Author:
Sagar Kamarthi, PhD
Northeastern University
360 Huntington Ave
Boston, MA, 02115
United States
Phone: 1 6173733070
Email: sagar@coe.neu.edu

Abstract

Background: It is well established that lack of physical activity is detrimental to the overall health of an individual. Modern-day activity trackers enable individuals to monitor their daily activities to meet and maintain targets. This is expected to promote activity encouraging behavior, but the benefits of activity trackers attenuate over time due to waning adherence. One of the key approaches to improving adherence to goals is to motivate individuals to improve on their historic performance metrics.

Objective: The aim of this work was to build a machine learning model to predict an achievable weekly activity target by considering (1) patterns in the user's activity tracker data in the previous week and (2) behavior and environment characteristics. By setting realistic goals, ones that are neither too easy nor too difficult to achieve, activity tracker users can be encouraged to continue to meet these goals, and at the same time, to find utility in their activity tracker.

Methods: We built a neural network model that prescribes a weekly activity target for an individual that can be realistically achieved. The inputs to the model were user-specific personal, social, and environmental factors, daily step count from the previous 7 days, and an entropy measure that characterized the pattern of daily step count. Data for training and evaluating the machine learning model were collected over a duration of 9 weeks.

Results: Of 30 individuals who were enrolled, data from 20 participants were used. The model predicted target daily count with a mean absolute error of 1545 (95% CI 1383-1706) steps for an 8-week period.

Conclusions: Artificial intelligence applied to physical activity data combined with behavioral data can be used to set personalized goals in accordance with the individual's level of activity and thereby improve adherence to a fitness tracker; this could be used to increase engagement with activity trackers. A follow-up prospective study is ongoing to determine the performance of the engagement algorithm.

XSL•FO

RenderX

Introduction

Background

Studies have reported the efficacy of physical activity in reducing the risk of disease; however, physical inactivity is on rise in the United States [1]. Considering that physical inactivity was the fourth leading cause of mortality in 2016 [2], there is much emphasis on developing effective methods to maintain healthy levels of physical activity. One promising solution is wearable fitness trackers that enable individuals to monitor their activity levels and patterns to ensure a healthy level of physical activity [3].

It is reported that about 20% of the general health-tracking population use smart devices such as medical gadgets, mobile phone apps, or online tools to track their health data [4]. The use of technology to objectively monitor physical activity is associated with higher levels of activity [5]. However, the potential benefits derived from the use of physical activity trackers are challenged by the limited and transient adoption of devices that necessarily require sustained use to exert their intended effect. Continued engagement with fitness trackers is an issue that warrants further investigation [5] A previous study [6] found two factors associated with the adoption and sustained use of physical activity trackers: (1) the number of digital devices owned by the participants, and (2) the use of activity fitness trackers and other smart devices by the participants' family members. The existence of these two factors bode well for the increased use of activity trackers; one study [1] demonstrated that motivational factors are associated with physical activity levels. Time constraints, fatigue, and aversion to exercise are some of the barriers to engaging in physical activity [1]. It has been reported that adjustments to activity targets are likely to enhance the users' commitment to physical activity and engagement with fitness trackers [7].

With the rise of machine learning and availability of activity tracker data, it is possible to create a model that can learn from users' behavior and adjust activity targets. Machine learning methods have broadly been applied to many health care areas such as cancer staging, risk assessment, and drug recommendation systems [8]. Researchers have studied the accuracy of activity trackers for energy expenditure assessment [9]. Having an automated personal trainer enabled by data mining techniques can be useful for an amateur athlete who cannot afford a personal trainer [10]. Similarly, effective feedback methods can be used for helping both athletes and coaches [11].

Although, there have been some attempts to study the benefit of activity trackers, there is room for studies on how to make use of activity trackers on a sustained basis. This study, to the best of our knowledge, is the first of its kind to develop a machine learning method to adjust the activity target for activity trackers. Machine learning techniques such as, but not limited to, lasso regression, ridge regression, Bayesian ridge regression, neural networks, random forest regression, and support vector regression have been used for prediction in the medical field [12].

Feature selection is an important step for improving model performance [13]. Prior to applying machine learning techniques, it is essential to study the data to find features that might negatively or positively affect the model [8]. In this work, we applied two feature selection techniques with a support vector machine [14]: principal component analysis [15] and recursive feature elimination.

We compared predictive models developed over (1) all features, (2) features generated by principal component analysis, (3) features selected by recursive feature elimination, and (4) features found from the authors' previous study [6] that characterizes participants environments. Figure 1 shows the study flow diagram highlighting the key steps undertaken in this study.

Figure 1. Study flow diagram.
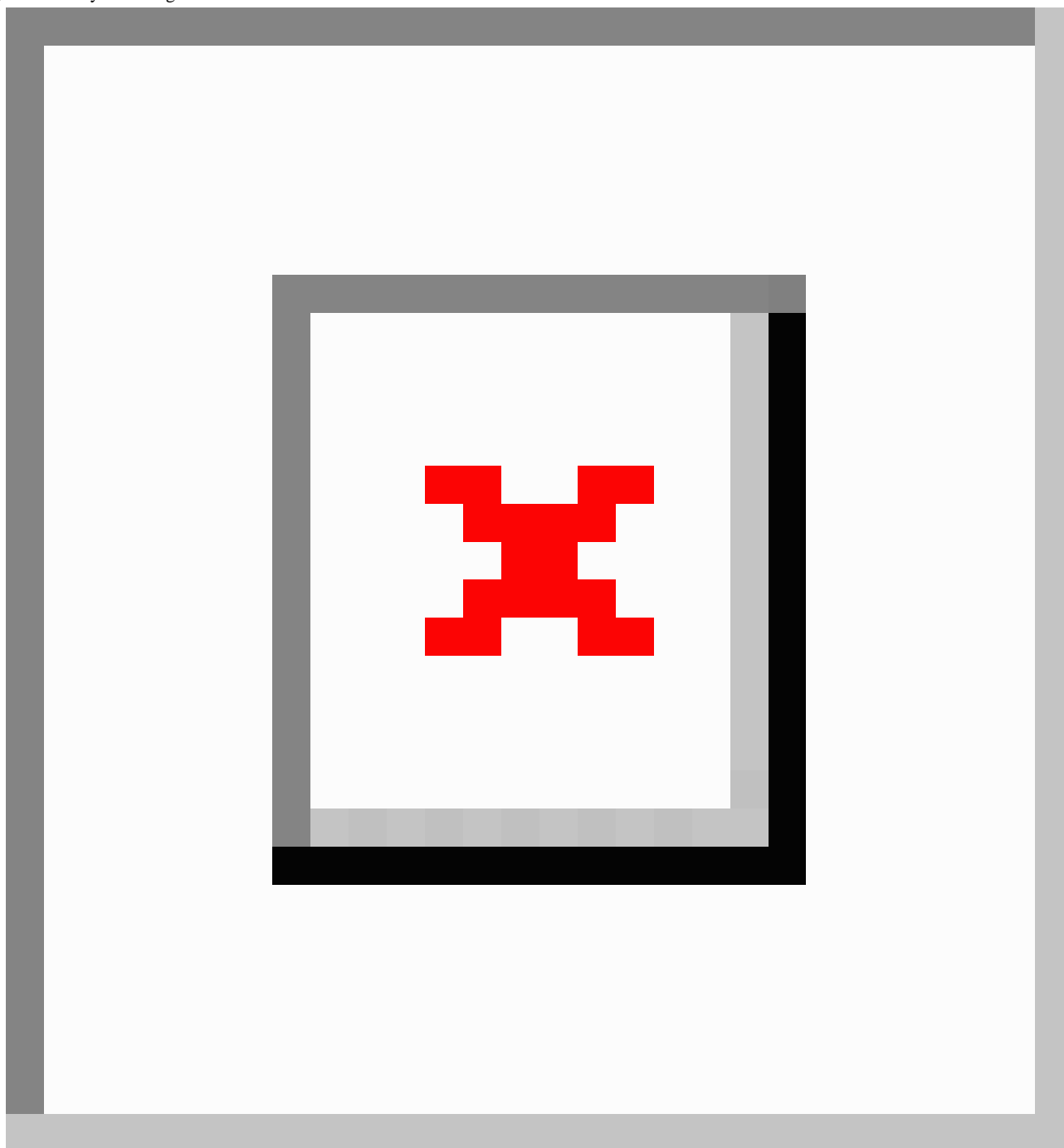
Objective

The purpose of this study was to develop a predictive model to estimate achievable weekly step goals. By setting realistic targets that are neither too easy nor too difficult to achieve, activity tracker users can be encouraged to continue to meet their weekly activity goals, and at the same time, to find utility of the tracker. We chose individuals who were overweight as our first use case, because the benefits of sustaining or even increasing physical activity in this population are well known, while there have been a lack of sustained interventions addressing this issue [16-18].

Methods

Data Collection

The study was designed as a 9-week, nonrandomized pilot in which the data were analyzed retrospectively. For this purpose, adults (N=30) with a BMI of 25 kg/m^2 or greater were enrolled from a local Massachusetts General Hospital–affiliated clinic. After screening the participants and seeking their consent, the research team directed the study participants to visit the Wellocracy website [19] to read information regarding the study and review different types of activity trackers (and their features) available for their use during the 9-week study. The study staff assisted the participants with the device setup process as needed.

For the study, 27 participants chose to use the FitBit Charge (Fitbit Inc) model, 2 chose the FitBit One (Fitbit Inc), and 1 chose the FitBit Zip (Fitbit Inc); 10 participants data were not used for analysis (7 participants either did not use the activity trackers at all or used less than 3 days a week, and an additional 3 participants demonstrated irregular use of their activity trackers on week by week basis). The data from the remaining 20 participants were used for model building.

The following surveys were collected from each participant both at enrollment and at the closeout stages: Behavioral Regulation in Exercise Questionnaire (BREQ-2) [20]; Barriers to Being Active (BBA) [21]; Patient Health Questionnaire (PHQ-8) [22]; Prochaska Stage of Change [23]; and Patient-Reported Outcomes Measurement Information System (PROMIS) Global-10 [24]. These surveys included questions about technology use and the ownership of electronic devices (BREQ-2), questions related to perceived barriers to exercise and activity (BREQ-2 and BBA), depression screening questions (PHQ-8), stages of change (Prochaska), and general health questions (PROMIS Global-10).

The BREQ-2 is designed to gauge the extent to which reasons for exercise are internalized and self-determined based on the following categories: motivation, external, introjected, identified, and intrinsic. In contrast, the BBA assesses whether participants gauge certain categories as reasons for inactivity and includes energy, willpower, time, and resources. A score of 5 or greater for a category indicates that it is a substantial barrier to a person's ability to exercise.

Participants were instructed to continuously wear the activity tracker for the entire period of the study. The first week was treated as a run-in period. Participants were contacted minimally during the remaining 8-week period to facilitate observation of participants' activity tracker habits without interference. At the end of the study, participants completed a closeout survey either online or in paper format and underwent a phone interview to gather information on their experiences during the study. All interviews were conducted and transcribed for analysis by a trained neuropsychologist.

Experimental Setting

We divided the data into disjoint training and testing sets. We chose 16 participants' data as training data set and the remaining 4 participants' data were used as the testing data set. We explored and fine-tuned hyperparameters using the training data set. Each participant was informed and had a fixed weekly activity goal for weeks 2 through 8; it was fixed for each participant at 110% of their week 1 average daily step count. Since we didn't have a means to adjust the activity goal for each week for each participant, this value was used as an estimate of the personalized average daily step count goal for each week. We ignored the week if a participant used the tracker device for less than 3 days during the week.

Data Preprocessing

We collected all data from questionnaires and the participants' daily step count data from their activity trackers. The questionnaires generated 96 variables. The variables were screened to determine candidate predictors for building a machine learning model. In the first pass of the variable screening process, 11 variables whose variance was zero were eliminated. In the second pass, we examined pairwise correlations among all remaining variables to eliminate redundant ones. Since we found that all pairwise correlations were less than 0.6, we considered all variables to be nonredundant.

Machine Learning Techniques

Figure 2 presents the models used to predict participants' activity target. Selection of these models was based on their suitability and capabilities. Bayesian ridge regression estimates a probabilistic model of a ridge regression [25]. Lasso regression is robust to overfitting due to its regularization penalty [26]. Random forest models are versatile for numeric and categorical predictors and for classification and regression tasks. Random forest models can also be interpreted easily and are less susceptible to underfitting [27]. Neural network models consist of hundreds or thousands of neurons that perform mathematical operations to recognize patterns [28]. Support vector regression models are regression models whose optimization is unaffected by the dimensionality of the data [14].

We employed these models in four cases: (1) using all features, (2) using new features built by principal component analysis, (3) using important features found by recursive feature elimination, and (4) using a subset of features found from the authors' previous study [6] that characterizes participants environments.

Figure 2. A schematic depicting models developed using the training data.
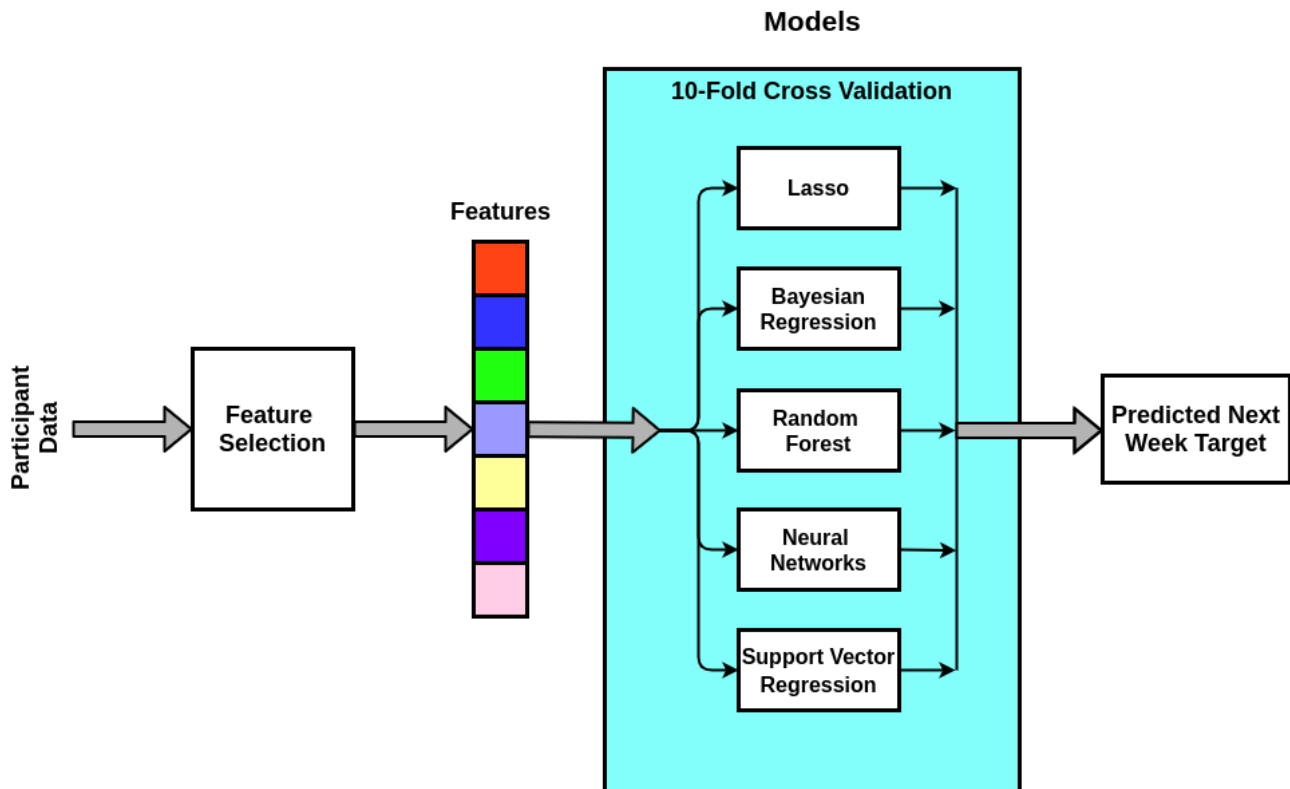
Feature Selection and Extraction

In this study, we used participants' mental health, behavioral information, and weekly activity performance as inputs. For every 7-day moving window, the subsequent 7 days were considered as the forthcoming week. The questionnaires provided a highly redundant and low variance data set. These features coupled with 7 daily steps counts of the week and a normalized Shannon entropy (E_s) value of the weekly step count were considered candidate predictors for target step count for each individual participant for the forthcoming week [29]. The normalized Shannon entropy was computed as $E_s = \frac{-\sum_{i=1}^{N} p_i \ln(p_i)}{\ln N}$ where i denotes the day, p_i denotes the portion of the total weekly steps completed on day i, and N denotes the number of days per week (ie, 7). The normalized Shannon entropy varies between 0 and 1. A value close to 0 indicates that the daily step counts of the participant throughout the week were irregular; in contrast, a value close 1 indicates that the step counts were consistent.

In total there were 93 candidate features for building the predictive models: 85 features from the questionnaires, 7 features from the daily steps of the week, and one weekly feature (ie, Shannon entropy). We used all 93 features for models developed in case 1.

For case 2, we developed a principal component analysis model. Principal component analysis is a dimension reduction technique that is widely used for extracting uncorrelated features components from correlated variables [14]. The principal component analysis used the 85 features from the questionnaires. We combined the principal components with the 7 daily and 1 weekly step count features.

In case 3, to identify the important features, we performed recursive feature elimination with support vector regression using the 85 questionnaire variables. We augmented these features with the 7 daily step count features and the normalized Shannon entropy value of the weekly step count.

Lastly, for case 4, we selected 2 features that were found to be important from the questionnaire in our previous work [6] studying the link between the participants' environmental factors and the adherence to the use of activity trackers. These features were (1) the number of digital devices owned by the participants, and (2) the use of activity fitness trackers and other smart devices by the participants' family members. We appended these two features with the daily step count features, and Shannon entropy of the weekly step count.

Model Evaluation

We trained the models in this study with 10-fold cross-validation using the training data set. We compared the performance of the models using mean absolute error (MAE) and adjusted R^2. We tested the final models (models A, B, C and D) of each case (cases 1, 2, 3, and 4) on an unseen test data set.

Statistical Analysis

We used Python (version 3.5.0) and R (version 3.4.1) for model development and statistical tests. We performed a one-sided paired t test ($P<.05$ were deemed significant) for statistical comparison between the models. The null hypothesis was that the mean error of two models were equal, with the alternative hypothesis that the candidate model had a lower mean error than the mean error given by the comparison model.

Results

Participant Characteristics

We present the characteristics of the study participants in Table 1. Among them, 10 participants were eventually removed from the study since they stopped using the activity trackers.

Over the span of 8 weeks, not all participants met their week-by-week step objectives. By and large, under half of participants met their progression objective every week (see Table 2). We present the distribution step count for all participants for each of the 9 weeks in Figure 3. We also presented the distribution of steps for each day of the week during the 9-week study period (see Figure 4).

Table 1. Patient demographic data.

Variable	Enrolled (N=30)	Participants (n=20)
Age (years), mean (SD)	48.9 (9.5)	47.7 (10.2)
Gender, n (%)		
Male	9 (30)	6 (30)
Female	21 (70)	14 (70)
BMI at enrollment		
Mean (SD)	32.5 (4.6)	32.8 (4.7)
Range	25.0-41.2	25.0-41.2
Race, n (%)		
White	21 (70)	14 (70)
American Indian or Alaskan Native	1 (3)	1 (5)
Black or African American	3 (10)	2 (10)
Hispanic	3 (10)	3 (15)
Unknown	2 (6)	0 (0)
Marital status, n (%)		
Married	8 (27)	6 (30)
Divorced or separated	8 (27)	5 (25)
Single (never married)	8 (27)	6 (30)
Living with partner	3 (10)	2 (10)
Widowed	1 (3)	1 (5)
No response	2 (7)	0 (0)
Education, n (%)		
12 years or completed high school or GED	5 (17)	3 (15)
Some college	5 (17)	2 (10)
College graduate	9 (30)	8 (40)
Post–high school	2 (7)	2 (10)
Postgraduate	2 (7)	1 (5)
Less than high school	3 (10)	2 (10)
Unknown	4 (13)	1 (5)
Employment status, n (%)		
Employed/self-employed	15 (50)	12 (60)
Disabled	5 (17)	3 (15)
Unemployed	5 (17)	2 (10)
Student	1 (3)	1 (5)
Retired	1 (3)	1 (5)
Unknown	3 (10)	1 (5)

Table 2. Participants meeting their average daily step goal for the week (110% of the average daily step count in week 1) over the course of the study.

Week	Participants (n=20) who met goal, n (%)
2	4 (20)
3	10 (50)
4	9 (45)
5	4 (20)
6	8 (40)
7	6 (30)
8	4 (20)
9	5 (25)

Figure 3. Weekly distribution of steps among all 20 participants.
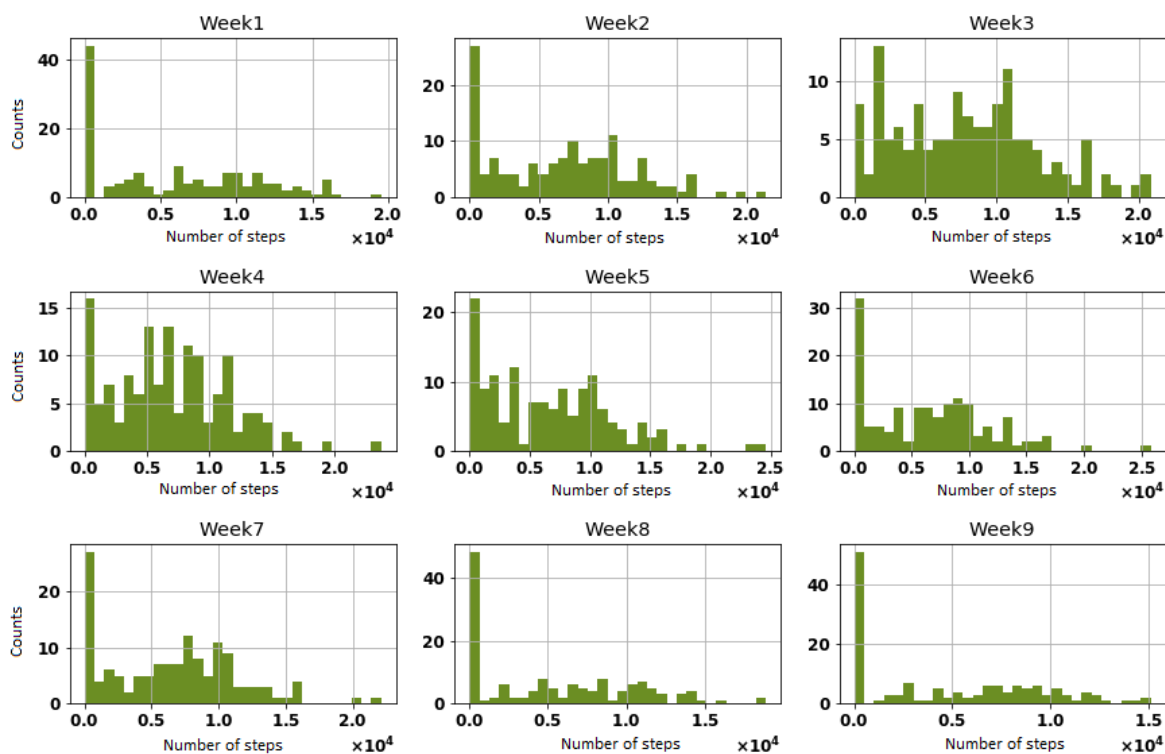
Figure 4. Step count distribution on each day of the week.
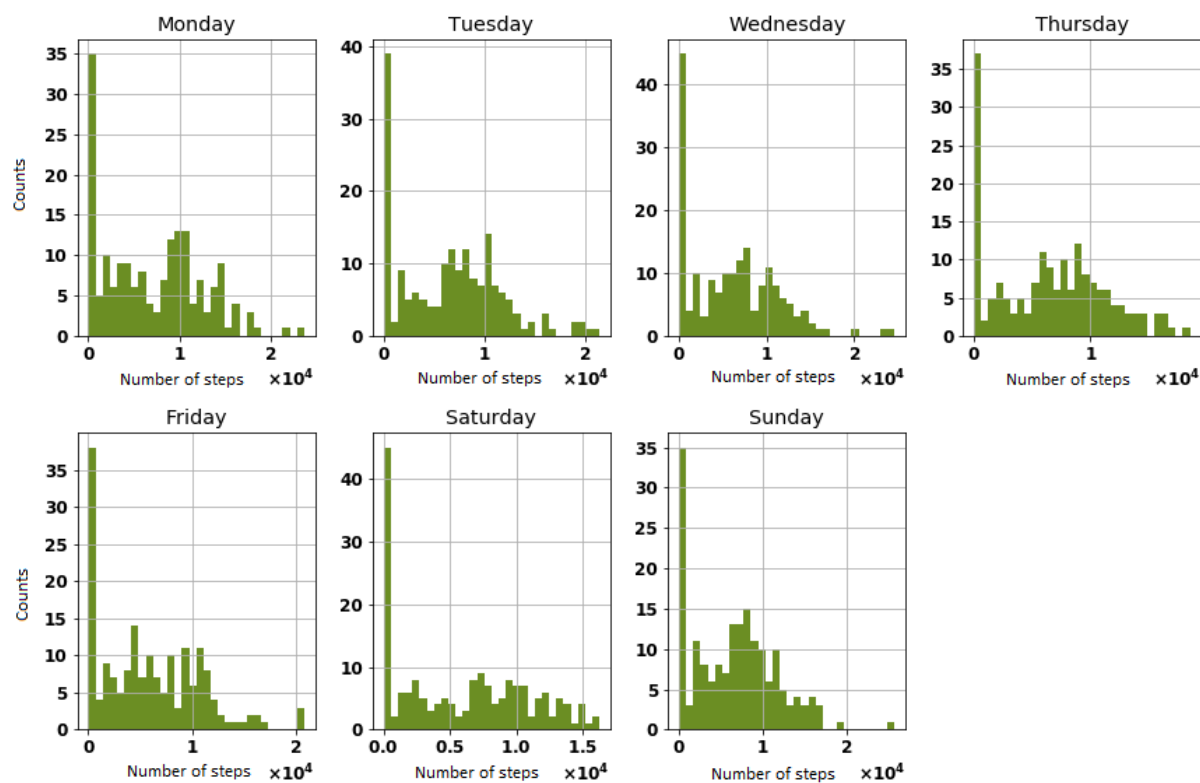
Parameter Tuning and Feature Selection

We performed a grid-search with 10-fold cross validation over the training data set. Lasso regression with α=0.01, Bayesian ridge regression with α=0.0001, and support vector regression with polynomial kernel of degree 2 with a regularization parameter of 0.00001 gave the best performance. Optimal parameters for random forests and neural networks depend on feature selection techniques. We found 30 variables from the questionnaires to be important using recursive feature elimination with support vector regression (see Figure 5). In the case of principal component analysis, the top 14 principal components explained 100% of the variance of the variables in questionnaires as shown in Figure 6.

Figure 5. Number of features found to be important by support vector regression.

XSL•FO

RenderX
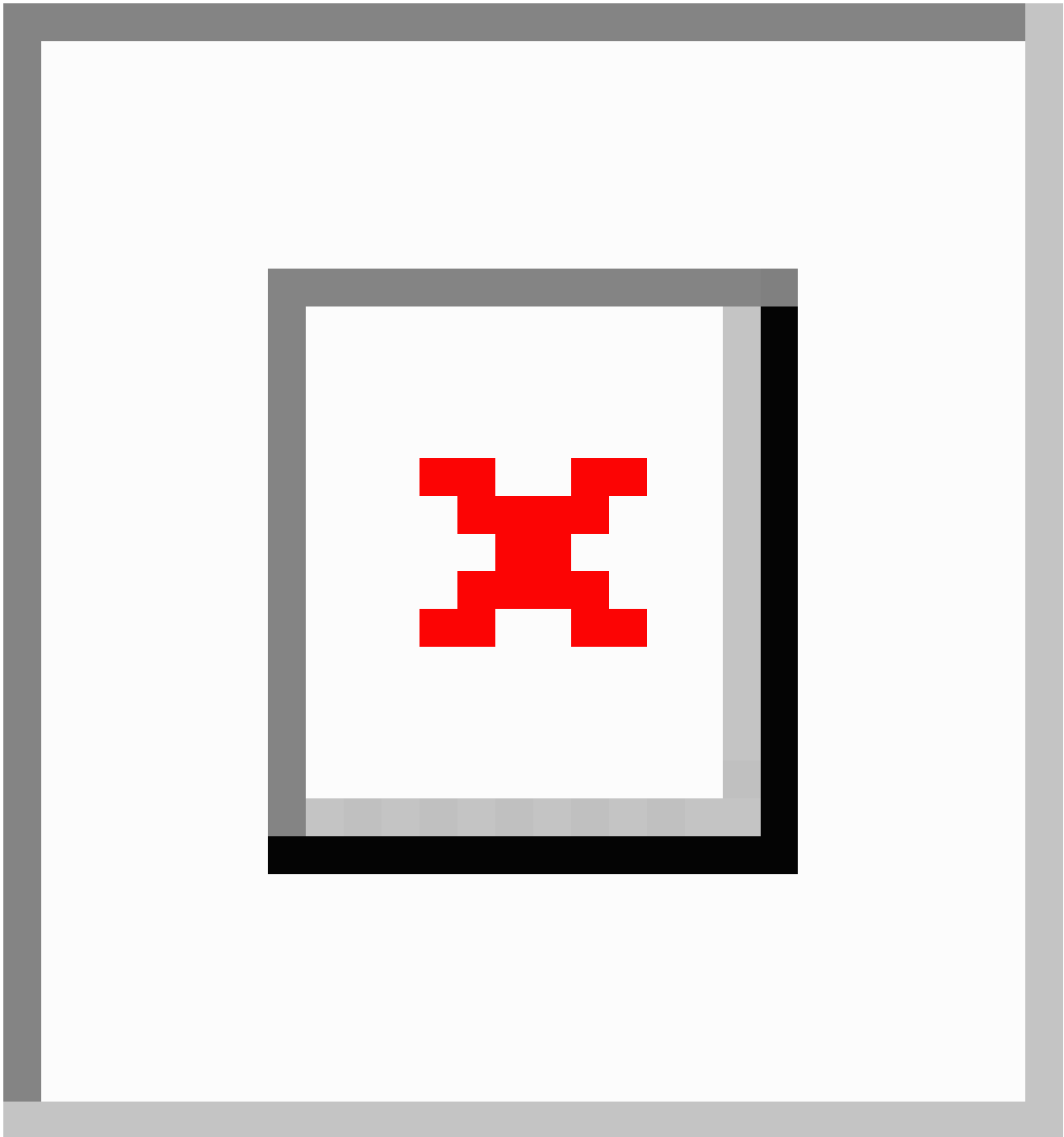
Figure 6. Number of components found by principal component analysis.
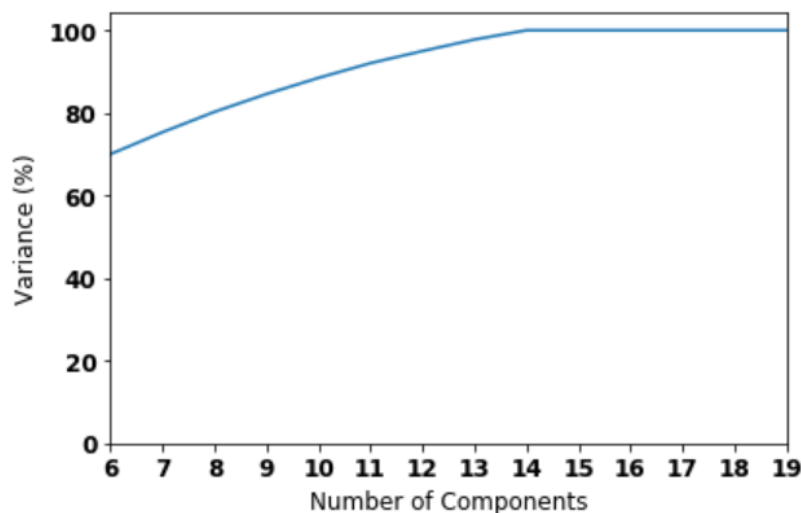
Model Performance

We report the performance for all models of 10-fold cross-validation with the training data set in Table 3. We used mean absolute error (MAE) and adjusted R^2 for the model comparison. Among all models developed, the Bayesian ridge regression model that used 93 features (case 1) gave the best performance for the training set (see Figure 7; MAE 1672, 95% CI 1640-1704; adjusted R^2=0.85). The Bayesian ridge regression model is referred to as model A in the rest of this paper.

Similarly, in case 2, we appended 14 principal components extracted from principal component analysis to the 7 daily and 1 weekly step count features resulting in 22 features. A random forest model (MAE 1700, 95% CI 1664-1737; adjusted R^2=0.91) gave the best performance among all models (see Figure 8). This random forest model is referred to as model B for the rest of this paper.

Table 3. Results of 10-fold cross-validation with training set.

Case and models	Mean MAE	95% CI	P value
1 All features			
Bayesian ridge regression[a]	1672	1640-1704	Model ref[b]
Lasso	2016	1985-2047	.002
Random forest	2425	2386-2464	.002
Neural network	1856	1813-1899	.002
Support vector regression	2425	2386-2464	.002
2 Feature selection using principal component analysis			
Bayesian ridge regression	2139	2107-2171	<.001
Lasso	2036	2005-2067	<.001
Random forest[c]	1700	1664-1737	Model ref
Neural network	1956	1926-1985	.03
Support vector regression	2938	2862-3013	<.001
3 Feature selection using recursive feature elimination			
Bayesian ridge regression	2131	2100-2163	.002
Lasso	2026	1995-2057	.002
Random forest[d]	1774	1739-1809	Model ref
Neural network	1906	1855-1958	.002
Support vector regression	2548	2473-2624	.002
4 Feature selected from previous study			
Bayesian ridge regression	2564	2537-2592	<.001
Lasso	2457	2429-2485	<.001
Random forest	1810	1768-1852	.04
Neural network[e]	1622	1589-1655	Model ref
Support vector regression	2940	2864-3016	<.001

[a]Model A (ie, this model gave the best performance in case 1).

[b]Reference model for comparisons.

[c]Model B (ie, this model gave the best performance in case 2).

[d]Model C (ie, this model gave the best performance in case 3).

[e]Model D (ie, this model gave the best performance in case 4).

XSL•FO

RenderX

Figure 7. Cross-validation performance of models using all questionnaire features, 7 daily step count features, and weekly entropy feature. BRIDGE: Bayesian ridge regression; NN: neural network; RF: random forest; SVR: support vector regression.
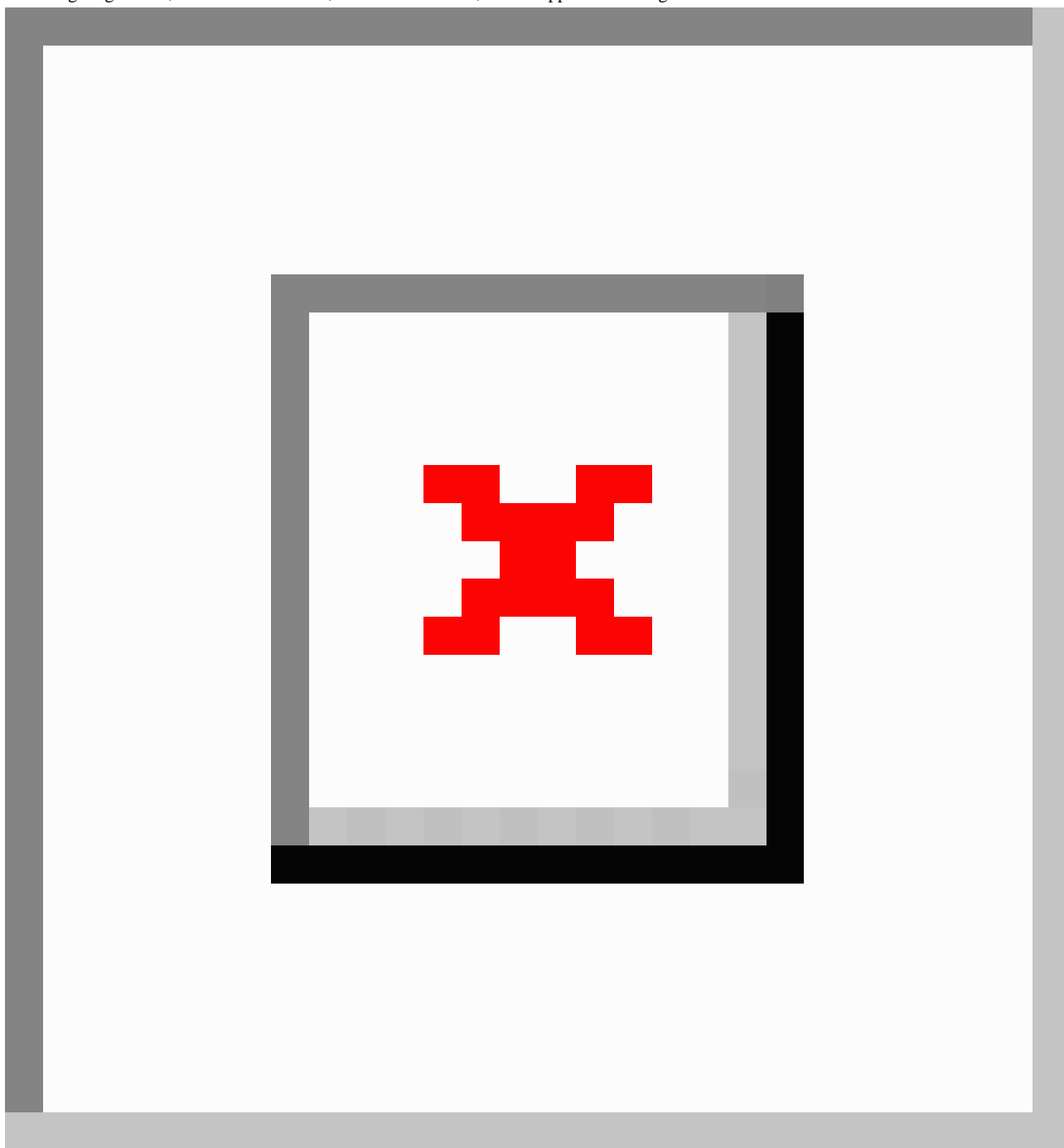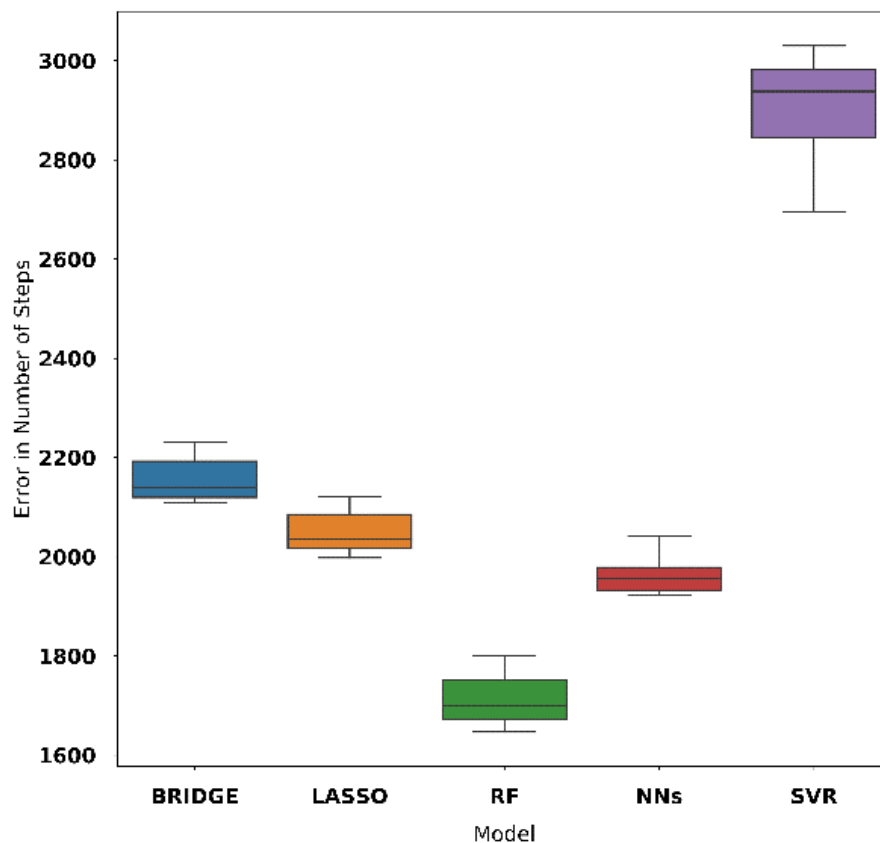
Figure 8. Cross-validation performance of models using features generated by principal component analysis, daily step count features, and weekly entropy feature. BRIDGE: Bayesian ridge regression; NN: neural network; RF: random forest; SVR: support vector regression.


In case 3, we appended variables found by recursive feature elimination to the 7 daily step counts and 1 weekly entropy feature which resulted in 39 features. A random forest model (MAE 1774, 95% CI 1739-1809; adjusted R^2=0.81) offered the best performance among all the models developed with these features (see Figure 9). We refer to this random forest model as model C.

Finally, for case 4, we coupled two features found in the previous study [6] with the 7 daily step counts and 1 weekly entropy feature which resulted in 10 features. A neural network gave the best performance across all models (MAE 1622 steps, 95% CI 1589-1655; adjusted R^2=0.89) (see Figure 10). We refer to this neural network model as model D.

Model D gave the best predictive performance among all the models. It had the lowest MAE across all models explored. We compared the predictive power of the model D (neural networks) with those of models A, B, and C using the testing data set. We found that model D gave a better predictive performance as shown in Figure 11. We performed comparisons using *t* tests between the errors generated by the model D and errors generated by models A, B, and C as shown in Table 4. We observed that model D's lower errors in comparison to those of Bayesian ridge regression (model A: *P*=.01), random forest (model B: *P*<.001), and random forest (model C: *P*=.01) models were significant.

Figure 9. Cross-validation performance of models using all features given by recursive feature elimination, 7 daily step count features, and weekly entropy feature. BRIDGE: Bayesian ridge regression; NN: neural network; RF: random forest; SVR: support vector regression.
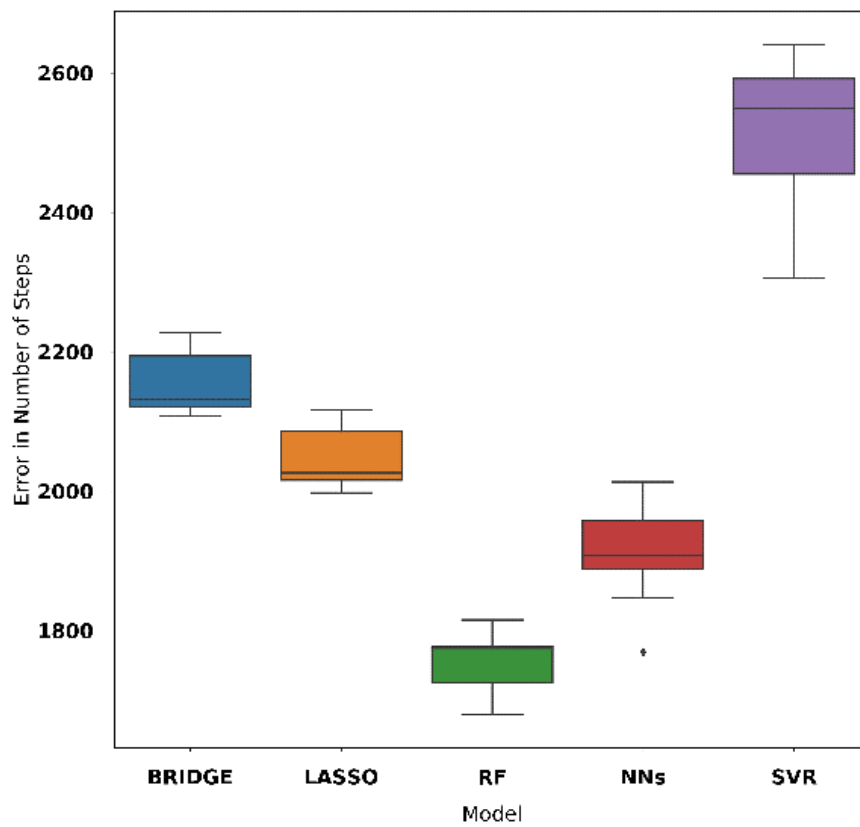
Figure 10. Cross-validation performance of models using features generated from previous knowledge, 7 daily step count features, and weekly entropy feature. BRIDGE: Bayesian ridge regression; NN: neural network; RF: random forest; SVR: support vector regression.
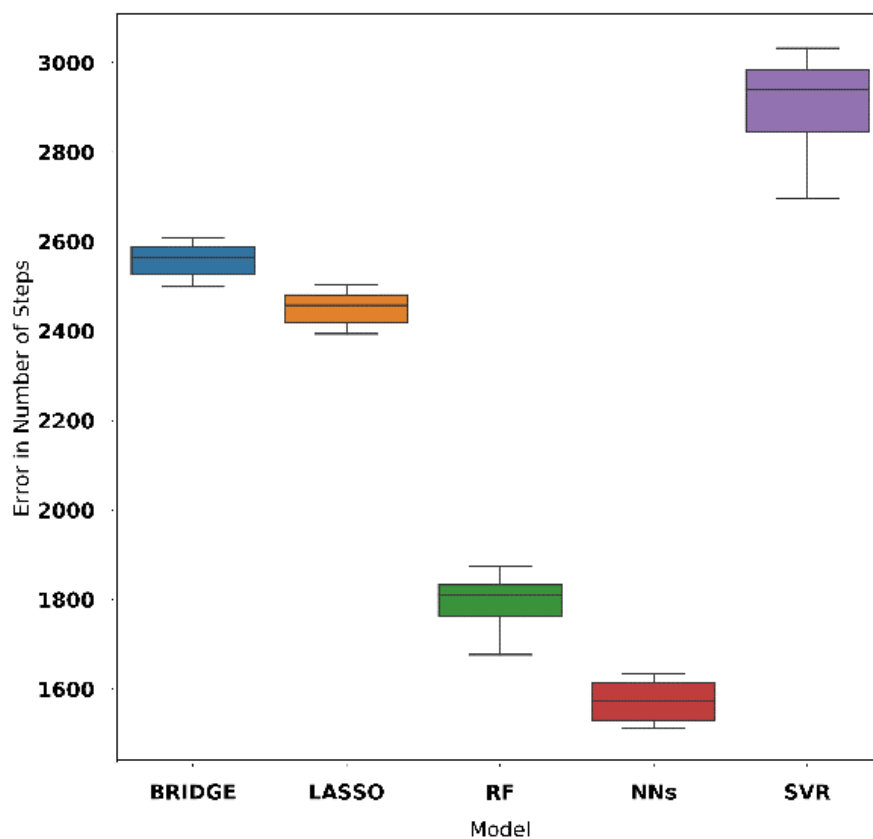
Figure 11. Boxplot of errors in terms of steps over the test set for Model A (Bayesian ridge regression), Model B (random forest), Model C (random forest), and Model D (neural network). MAE: mean absolute error.
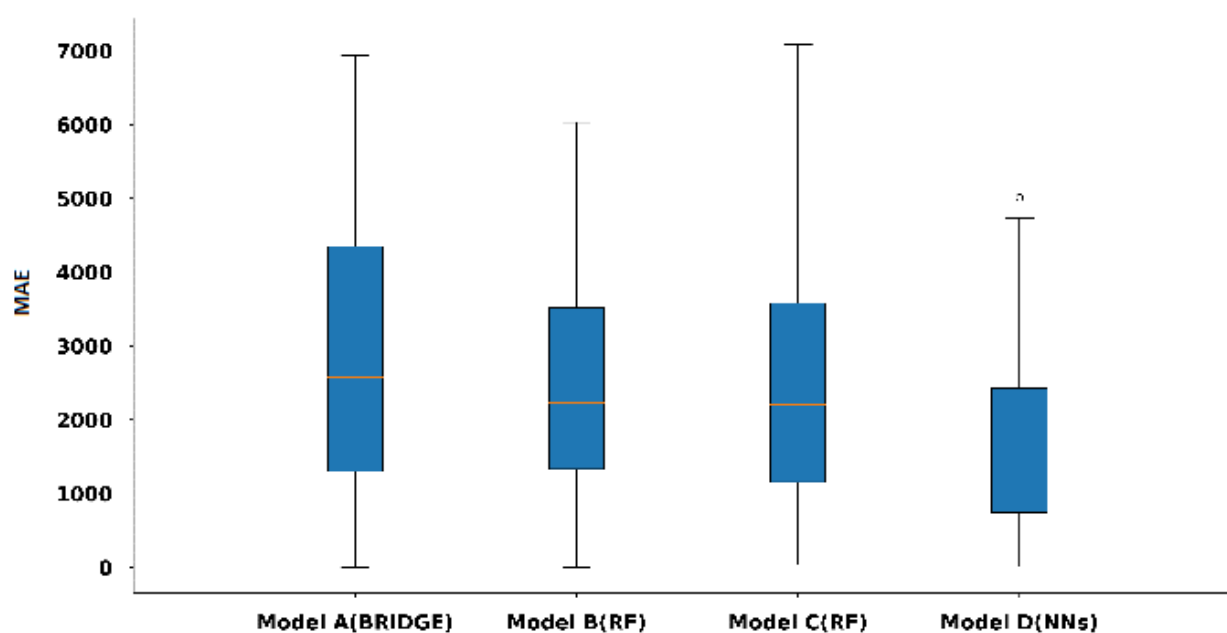
Table 4. Breakdown of model results over the test data set.

Model (type)	Mean MAE[a]	95% CI	P value
Model D (neural network)	1545	1383-1706	Model ref[b]
Model C (random forest)	2210	1990-2420	.01
Model B (random forest)	2230	2015-2445	<.001
Model A (Bayesian ridge regression)	2578	2310-2845	.01

[a]MAE: mean absolute error.

[b]Reference model for comparisons.

We evaluated the naïve rule: using participants' average daily step count of the week as the prediction for the subsequent week's activity goal [30]. This is a reasonably competitive approach because the weekly target exhibited strong autocorrelation. The naïve rule achieved an MAE of 1664 steps while model D achieved an MAE of 1545 steps (95% CI 1383-1706) for the 4 test participants over a period of 8 weeks.

Sensitivity Analysis

Its reported that the Fitbit activity devices have margin of ±5% error in their step count readings [31,32]. We tested the performance of our best model (model D) on three noisy data sets generated from the original test data set by adding ±1%, ±3%, and ±5% noise. The first noisy data set was generated by adding random noise between –1% and +1% to the Fitbit readings. Similarly, the second and third noisy data sets were generated by adding ±3% and ±5% noise to the original data set. To conduct the sensitivity analysis, we generated 100 replications of each noisy data sets by adding different random noises with limit specified for each data set. These data sets were used to evaluate the performance of model D.

The model D achieved an MAE of 1606 steps (95% CI 1490-1755) for the noisy test data set generated using margin of ±1% error which is approximately 4% higher than the MAE achieved for the original noise-free test data set. Similarly, model D's MAE was 1670 steps (95% CI 1571-1840) for the margin of ±3% error which was approximately 8% higher than the MAE achieved for the original noise-free test data set. Finally, for the data set generated using margin of ±5% error, model D's MAE of 1710 steps (95% CI 1621-1908) was approximately 10% higher than the MAE achieved for the original noise-free test data set.

This led to the empirical observation that 2%, 6%, and 10% noise in Fitbit readings leads to approximately 4%, 8%, and 10% increases in prediction error, respectively. Therefore, we assume that the model error sublinearly increases with margin of error in activity tracker readings. However, one might need to validate this observation with further evaluation data.

Feature Importance

We experimented with integrated gradients [33] in order to analyze the features of model D. This method provides a score that reflects the contribution of each variable to the response

variable by calculating the integral of the gradient of the response variable with respect to that variable. We report the top features for model D in Figure 12.

Figure 12. Importance of the features as measured by the integrated gradient method.
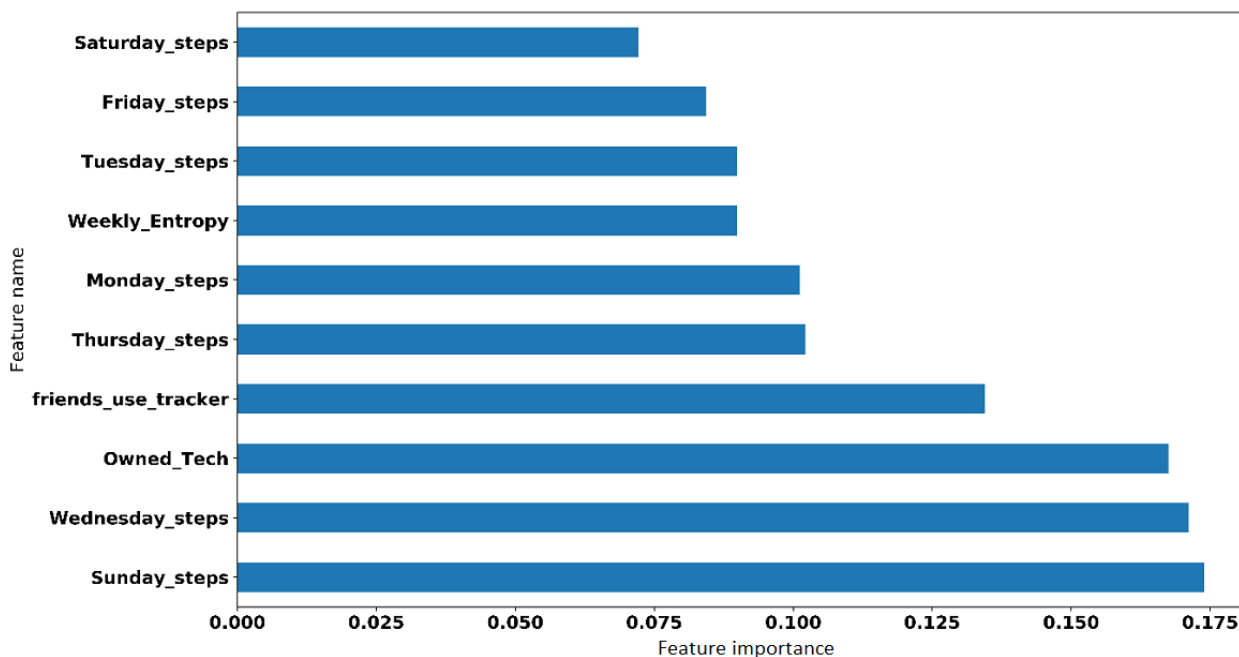
Discussion

Precise prescient calculations that consolidate information are one of the main focuses in predictive analyses [34]. To the best of our knowledge, this work is one of the first studies to explore machine learning models with the aim of adjusting step count goals. Inputs to these models use an individual's personal, social, and environmental factors as well as weekly activity data. A recent study concluded that activity tracker users feel unmotivated despite having knowledge about the benefits of physical activity [35]. Users can become unmotivated if they cannot meet their activity goal [36]. Step count goals that are too easy to achieve may lead to abandonment of the activity tracker, and those that are too high will discourage the individual [37].

In previous work [6], we studied the factors influencing the use of activity trackers and identified two factors that likely promote the continued use of activity trackers: (1) the number of digital devices that the participant owns, and (2) whether or not members of the participants' family use activity fitness trackers and other smart devices. Extending the previous work [6], in this study, we explored different predictive models to estimate achievable weekly step goals to encourage the use of activity trackers. This study can be used to set goals for individuals and can be accompanied by proper motivation messages to improve the sustained use of the activity trackers [38-40].

This study has some limitations. The number of participants was low, and all of the participants were selected from a cohort with BMI of 25 kg/m^2 or greater from the same geographical area. Participants who choose to participate in this study were more likely to use from 3 models of Fitbit activity trackers than those who chose not to participate in the study. Moreover, only participants who completed the 9-week study were considered for further analysis. Finally, in this study, we used 110% of week 1 averaged daily step count as the best estimate of each participant's personalized goal for each week. The goal may have been on heavy side for some participants and on the easy side for others. This is, of course, a limitation of the study in the absence of a mechanism for a correct estimate of the goal for each participant.

As an extension to this work, a new study with the goal of validating the developed model with a large number of participants has been undertaken. The new study recruited 120 individuals from the general public to use the neural network-based predictive model developed herein over a period of 6 months. This model, hosted on a server, provides each user with achievable daily steps goals and updates the model parameters on a weekly basis. The data from the study that is underway will be used to fine tune the predictive model and to gain insight into activity tracker users to motivate and manage their physical activity.

Authors' Contributions

RM and SK worked on data exploration, dimension reduction through factor analysis, implementing and designing the neural network–based algorithm, generating results, and drafting the paper. SK is one of the guarantor authors of the paper. MA and AJC worked on the design and data collection. SA and KJ worked on the design, and JK is one of the guarantor authors.

Conflicts of Interest

None declared.

References

1. Bassuk SS, Manson JE. Epidemiological evidence for the role of physical activity in reducing risk of type 2 diabetes and cardiovascular disease. Journal of Applied Physiology 2005 Sep;99(3):1193-1204. [doi: 10.1152/japplphysiol.00160.2005]

2. Mercer K, Li M, Giangregorio L, Burns C, Grindrod K. Behavior Change Techniques Present in Wearable Activity Trackers: A Critical Analysis. JMIR mHealth uHealth 2016 Apr 27;4(2):e40. [doi: 10.2196/mhealth.4461]

3. Eheman C, Henley SJ, Ballard-Barbash R, Jacobs EJ, Schymura MJ, Noone A, et al. Annual Report to the Nation on the status of cancer, 1975-2008, featuring cancers associated with excess weight and lack of sufficient physical activity. Cancer 2012 Mar 28;118(9):2338-2366. [doi: 10.1002/cncr.27514]

4. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: A meta-regression. Health Psychology 2009;28(6):690-701. [doi: 10.1037/a0016136]

5. Wang JB, Cataldo JK, Ayala GX, Natarajan L, Cadmus-Bertram LA, White MM, et al. Mobile and Wearable Device Features that Matter in Promoting Physical Activity. JournalMTM 2016 Jul;5(2):2-11. [doi: 10.7309/jmtm.5.2.2]

6. Centi AJ, Atif M, Golas SB, Mohammadi R, Kamarthi S, Agboola S, et al. Factors Influencing Exercise Engagement When Using Activity Trackers: Nonrandomized Pilot Study. JMIR Mhealth Uhealth 2019 Oct 24;7(10):e11603 [FREE Full text] [doi: 10.2196/11603]

7. Rúben G, Evangelos K, Marc H. How do we engage with activity trackers?: a longitudinal study of habito. 2015 Sep 07 Presented at: ACM International Joint Conference on Pervasive Ubiquitous Computing; 2015; New York p. 1305-1316. [doi: 10.1145/2750858.2804290]

8. Koh HC, Tan G. Data mining applications in healthcare. J Healthc Inf Manag 2005;19(2):64-72. [Medline: 15869215]

9. Phelan S, Phipps MG, Abrams B, Darroch F, Grantham K, Schaffner A, et al. Does behavioral intervention in pregnancy reduce postpartum weight retention? Twelve-month outcomes of the Fit for Delivery randomized trial. Am J Clin Nutr 2014 Feb;99(2):302-311 [FREE Full text] [doi: 10.3945/ajcn.113.070151] [Medline: 24284438]

10. Fister, Iztok, Duan F, Simon F. Data Mining in Sporting Activities Created by Sports Trackers. 2013 Aug 26 Presented at: International Symposium on Computational and Business Intelligence. IEEE; 2013; India. [doi: 10.1109/iscbi.2013.25]

11. Liebermann DG, Katz L, Hughes MD, Bartlett RM, McClements J, Franks IM. Advances in the application of information technology to sport performance. Journal of Sports Sciences 2002 Jan;20(10):755-769. [doi: 10.1080/026404102320675611]

12. Tomar D, Agarwal S. A survey on Data Mining approaches for Healthcare. IJBSBT 2013 Oct 31;5(5):241-266. [doi: 10.14257/ijbsbt.2013.5.5.25]

13. Chandrashekar G, Sahin F. A survey on feature selection methods. Computers & Electrical Engineering 2014 Jan;40(1):16-28. [doi: 10.1016/j.compeleceng.2013.11.024]

14. Drucker, Harris. Support vector regression machines. Advances in neural information processing systems. Jolliffe, Principal component analysis, Springer 1997:155-161.

15. Jolliffe IT. Principal Component Analysis, 2nd ed. New York: Springer; 2002.

16. Söderlund A, Fischer A, Johansson T. Physical activity, diet and behaviour modification in the treatment of overweight and obese adults: a systematic review. Perspect Public Health 2009 May 07;129(3):132-142. [doi: 10.1177/1757913908094805]

17. Jakicic JM, Otto AD, Lang W, Semler L, Winters C, Polzien K, et al. The Effect of Physical Activity on 18-Month Weight Change in Overweight Adults. Obesity 2010 Jun 10;19(1):100-109. [doi: 10.1038/oby.2010.122]

18. Bickmore T, Schulman D, Yin L. Maintaining Engagement in Long-term Interventions with Relational Agents. Appl Artif Intell 2010 Jul 01;24(6):648-666 [FREE Full text] [doi: 10.1080/08839514.2010.492259] [Medline: 21318052]

19. http://www.wellocracy.com. URL: http://www.wellocracy.com [accessed 2020-07-28]

20. David M, Vannessa T. A modification to the Behavioural Regulation in Exercise Questionnaire to include an assessment of amotivation. Journal 2004:191-196. [doi: 10.1123/jsep.26.2.191]

21. Overcoming Barriers to Physical Activity. website. URL: https://www.cdc.gov/physicalactivity/basics/adding-pa/barriers.html [accessed 2020-07-28]

22. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. Journal of Affective Disorders 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026]

23. Holmen H, Wahl A, Torbjørnsen A, Jenum AK, Småstuen MC, Ribu L. Stages of change for physical activity and dietary habits in persons with type 2 diabetes included in a mobile health intervention: the Norwegian study in RENEWING HEALTH. BMJ Open Diab Res Care 2016 May 12;4(1):e000193. [doi: 10.1136/bmjdrc-2016-000193]

XSL•FO
RenderX

24. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009 Jun 19;18(7):873-880. [doi: 10.1007/s11136-009-9496-9]

25. Christopher M B. Pattern recognition and machine learning. In: springer. New York, NY: Springer; 2006.

26. Santosa F, Symes WW. Linear Inversion of Band-Limited Reflection Seismograms. SIAM J. Sci. and Stat. Comput 1986 Oct;7(4):1307-1330. [doi: 10.1137/0907087]

27. Genuer R, Poggi J, Tuleau-Malot C. Variable selection using random forests. Pattern Recognition Letters 2010 Oct;31(14):2225-2236. [doi: 10.1016/j.patrec.2010.03.014]

28. Simon H. Neural networks: a comprehensive foundation. In: book. NJ, United States: Prentice Hall PTR; 1994.

29. Cover, Thomas M. ,Joy A. Thomas.Elements of information theory. John Wiley & Sons 2012.

30. Mohammadi, Jain S, Agboola S, Palacholla R, Kamarthi S, Wallace BC. Learning to Identify Patients at Risk of Uncontrolled Hypertension Using Electronic Health Records Data. AMIA Jt Summits Transl Sci Proc 2019;2019:533-542 [FREE Full text] [Medline: 31259008]

31. Alinia P, Cain C, Fallahzadeh R, Shahrokni A, Cook D, Ghasemzadeh H. How Accurate Is Your Activity Tracker? A Comparative Study of Step Counts in Low-Intensity Physical Activities. JMIR Mhealth Uhealth 2017 Aug 11;5(8):e106. [doi: 10.2196/mhealth.6321] [Medline: 28801304]

32. An H, Jones GC, Kang S, Welk GJ, Lee J. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 2017 Apr;17(3):360-368. [doi: 10.1080/17461391.2016.1255261] [Medline: 27912681]

33. Sundararajan, Mukund, Ankur T, Qiqi Y. Axiomatic attribution for deep networks. 2017 Aug 11 Presented at: Proceedings of the 34th International Conference on Machine Learning-Volume 70. JMLR. Org; 2017; Sydney, Australia.

34. Rykov Y, Thach T, Dunleavy G, Roberts AC, Christopoulos G, Soh CK, et al. Activity Tracker–Based Metrics as Digital Markers of Cardiometabolic Health: Cross-Sectional Study. JMIR Mhealth Uhealth 2020 Jan 27;8(1):e16409. [doi: 10.2196/16409]

35. Glenn F, Shaun G, Andrew K, Leslie S. Why People Stick With or Abandon Wearable Devices. Catalyst Journal 2017 Sep 14.

36. Kerner C, Goodyear VA. The Motivational Impact of Wearable Healthy Lifestyle Technologies: A Self-determination Perspective on Fitbits With Adolescents. American Journal of Health Education 2017 Jul 28;48(5):287-297. [doi: 10.1080/19325037.2017.1343161]

37. Elizabeth M, David H, Gueorgi K. Mobile health apps: adoption, adherence, and abandonment. In: UbiComp/ISWC'15 Adjunct: Adjunct Proceedings. 2015 Sep 14 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings; 2015; Osaka Japan p. 261-264. [doi: 10.1145/2800835.2800943]

38. Couper MP, Alexander GL, Zhang N, Little RJ, Maddy N, Nowak MA, et al. Engagement and Retention: Measuring Breadth and Depth of Participant Use of an Online Intervention. J Med Internet Res 2010 Nov 18;12(4):e52. [doi: 10.2196/jmir.1430]

39. Bossen D, Buskermolen M, Veenhof C, de Bakker D, Dekker J. Adherence to a Web-Based Physical Activity Intervention for Patients With Knee and/or Hip Osteoarthritis: A Mixed Method Study. J Med Internet Res 2013 Oct 16;15(10):e223. [doi: 10.2196/jmir.2742]

40. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective Associations Between Intervention Components and Website Engagement in a Publicly Available Physical Activity Website: The Case of 10,000 Steps Australia. J Med Internet Res 2012 Jan 11;14(1):e4. [doi: 10.2196/jmir.1792]

Abbreviations

BBA: Barriers to Being Active
BREQ: Behavioral Regulation in Exercise Questionnaire
MAE: mean absolute error
PHQ: Patient Health Questionnaire
PROMIS: Patient-Reported Outcomes Measurement Information System

XSL·FO
RenderX