Sleep Monitoring with Wearable Sensor Data in an eCoach Recommender System: A Conceptual Study with Machine Learning Approach

Ayan Chatterjee[1*], Andreas Prinz[1], Nibedita Pahari[2], Jishnu Das[3], and Michael Alexander Riegler[4]

[1] Department of Information Technology, Center for eHealth, University of Agder, Norway
[2] Department of Software Engineering, Knowit AS, Norway
[3] Department of Library and Information Science, University of Calcutta, India
[4] Department of Holistic Systems, Simula Research Laboratory, Norway

Abstract. The collective effects of sleep loss and sleep disorders are correlated with many adverse health outcomes, including elevated risk of high blood pressure, obesity, diabetes type II, depressive state, and cardiovascular symptoms. Research in eHealth may provide techniques to enrich personal healthcare with Information and Communication Technologies (ICTs). An eCoach system may allow people to achieve a healthy lifestyle with extended health state monitoring (e.g., sleep) and tailored recommendation generation. Using supervised machine learning (ML) techniques, this study investigated the chance of classifying sleep stages at night for adults on hourly and daily basis. The daily total sleep minutes and hourly total sleep minutes for defined sleeping period served as input for the classification models. We first used publicly available Fitbit dataset to build the initial classification models. Second, using the transfer learning approach, we re-used the top five best-performing models on a real dataset as collected from the MOX2-5 wearable medical-grade activity device. We found that support vector classifier (SVC) with "linear" kernel outdated other classifiers with a mean accuracy score of 99.92% for hourly sleep classification and a K-nearest neighbor (KNN) outpaced other classifiers with a mean accuracy score of 99.47% for daily sleep classification, for the public Fitbit datasets. Moreover, to determine the practical efficacy of the classifier models, we conceptualized to use the classifier models in an eCoach prototype system to attain tailored sleep goals (e.g., a weekly goal of 49-63 hours of sleeping).

Keywords: sleep time, sleep stage, activity sensor, machine learning, eCoach, recommendation generation.

1 Introduction

Sleep is a particular category of physical activity. Adults require 7-9 hours of proper sleeping daily, while athletes may benefit as much as 10 hours to maintain a healthy lifestyle [1,2]. Sleep deprivation can increase the causes of memory issues, trouble with thinking and concentration, accidents, mood changes, weakened immunity, risk of diabetes, elevated blood pressure, weight increase, low sexual desire, risk of heart diseases, and body imbalance [1]. Unhealthy lifestyle practices, poor sleeping hygiene, sleep disorders, work pressure, and medical conditions may result in sleep deprivation [3]. In contrast, chronic oversleeping may cause cognitive loss, daytime sleepiness, lethargy, headaches, depressive mind, and trouble in falling or staying sleeping [4]. Research [4] in 2010 showed that staying up late for 20 to 25 hours can affect individual concentration and performance, just as blood alcohol concentration (BAC) is 0.10%. In most places, people are considered legally drunk when their BAC is 0.08%. A study [4] in 2014 of 24,671 adults found evidence that more than 10 hours of sleep a night or prolonged sleep is related to depression and obesity. Long-term sleep is also associated with high blood pressure and type 2 diabetes [4]. Therefore, sleep deprivation and excessive sleep may lead to the gradual development of chronic symptoms. Chronic diseases are the most normal cause of death worldwide [5-7]. It is the leading probability of dying between ages 30 and 70 years [7]. Around 60% to 85% of the world's people live a sedentary lifestyle [5-7]. Regular physical activity has a significant impact on good sleep [4]. Still, more than 80% of the adolescent population in the world lack physical activity regularly [7].

Studies related to sleep monitoring can be classified into monitoring with wearable devices or non-wearable devices [8]. Polysomnogram (PSG) has been the gold standard to assess sleep psychology [9]. Health and the clinical market have enhanced real-world sleep mode indicator [9]. Longitudinal home monitoring avoids certain limitations of laboratory PSG, such as atypical sleeping environments and single-night snapshots. Sleep is a dynamic process that changes every day, so measuring sleep over multiple nights for medical, research, and health reasons is essential [9]. Home monitoring equipment may provide a more realistic platform to capture sleep data for many nights. Longitudinal data may prove invaluable for discovering internal patterns of sleep variability or linking sleep to the timing of various other activities. The personal health goal of sleep monitoring to optimize health also needs to be achieved through longitudinal monitoring and self-tracking. In multiple environments outside the field of sleep medicine, portable monitoring can accomplish this goal. Sleep can be monitored based on brain activity signals (e.g., EEG, EMG, and EOG), automatic alerts based on movements (e.g., actigraphy, body position), bed-sleep monitoring, heart-rate-variability (HRV), body temperature, galvanic skin response (GSR), sleep images, PSG, and touch-free remote tracking (e.g., LIDAR, Wi-Fi) [9-11]. Relevant sleep monitoring devices are iBrain, Zeo, Heally Recording System, M1, Fitbit, MOX2-5, Lark, Sleep Cycle Alarm, Sleep Tracker, Up, WakeMate, Air Cushion, Early Sense Mattress, Emfit Bed Sensor, Home Health Station, Linen Sensor, Sleep Minder, Bio Harness, Health Vest, Magic vest, Radiofrequency monitor, Wrist Care [9]. Baron et al. [8] developed a "Sleep Bunny" mobile app based on wearable technology for sleep behavioral intervention; however, the app suffers from personalization, effective reminder design, and notification generation. Stucky et al. [12] used Fitbit Charge 2 wearable device to estimate sleep using Polysomnographic measures; however, it suffered from quantifying individual sleep episodes.

The idea of activity coaching may improve individual sleep health with daily and hourly sleep monitoring and tailored recommendation generation. In context, an electronic coach (eCoach) [13,14] system may generate personalized activity recommendations based on the insight from sensory observation to reach personalized activity (e.g., sleep) goals. From the literature search, the eCoach concept in eHealth is in the nascent stage, and there is very little research conducted on actual sensor data using machine learning technology. In this study, we have conceptualized a novel, personalized, and data-driven eCoaching concept that can collect activity data from participants with wearable activity sensors, process those data with different ML models to classify sleep stages, and generate personalized recommendations on individual progress to attain personalized sleep goals (e.g., daily, weekly, or monthly based on preferences). The research questions for this study are –

(RQ-1) How to classify daily and hourly sleep time into different sleep stages?
(RQ-2) How to fit the classification models in an eCoach system for recommendation generation to attain personalized sleep goal?

To demonstrate the pertinency of the study, we described how to apply the classification model to achieve personalized weekly sleep goals. The remainder of the paper is structured as follows. In section 2, we present the adopted methods. In Section 3, we discuss the experimental results, and the paper is concluded in Section 4.

2 Method

We used established statistical methods and ML models to analyze public and private sleep datasets for adults. Moreover, we assessed the performance of different ML classifiers against standard metrics to classify both hourly and daily sleep stages. The overall process includes data collection, data pre-processing, feature selection, data visualization, ML model training, testing, cross-validation, evaluation, and model reuse for personalized recommendation generation. In this study, we focused only on night sleep datasets for adults. Sleep data for the aged, children, athletes, bodybuilders, and pregnant women are beyond the scope of this study.

2.1 Data Collection

We used anonymous public Fitbit dataset for adult participants available in "Zenodo" [15] for initial ML model training and testing. The dataset has various features related to the activity; however, we selected the feature "sleep minutes" to maintain the focus of this study. We used the public dataset to discover the best performing classifiers with the defined feature in a multiclass classification problem.

Then, we applied the model to the actual dataset as collected with MOX2-5 wearable activity device [16] based on the transfer learning and incremental approach to proving the concept of personalized activity recommendation generation in an eCoach system to attain the personal sleep goal. Therefore, we collected anonymous nightly sleep data from two adults in Norway for one month using the MOX2-5 sensor following the ethical guidelines. The attributes of MOX2-5 sensor data are – timestamp, activity intensity (IMA), sedentary seconds, weight-bearing seconds, standing seconds, low physical activity (LPA) seconds, medium physical activity (MPA) seconds, vigorous physical activity (VPA) seconds, and steps per minute. IMA gives the impression if the activity is LPA or MPA or VPA. To associate the pre-trained model with the public dataset, we considered the "sedentary" feature from MOX2-5 for real-time classification. In MOX2-5 sensor, sedentary time refers to the non-activity duration, including leisure time and sleep time. Therefore, we considered ten hours of sleep data from two participants between 23:01:00 of day-(n-1) to 09:00:00 of day-n and calculated hourly and total daily sleep time. The relation between sedentary time and activity (LPA/MPA/VPA) time can be written as:

$$\sum (\text{sedentary, active, weight-bearing, standing}) = 60 \text{ seconds (sec.)}.$$

During sleep time sedentary minutes goes high (≈ 58-60 sec.) with IMA ≈ 0-20, step count ≈ 0, and activity time = 0. The IMA value can be correlated to the energy expenditure expressed in metabolic (MET) values. This makes it possible to classify:

Low Physical Activity (LPA): between 1.5 and 3 METS
Moderate Physical Activity (MPA): between 3 and 6 METS
Vigorous Physical Activity (VPA): 6.0 or more METS

For an upper leg sensor placement, the corresponding IMA thresholds are:

$4.5 < LPA \leq 11.9$ *cycles per seconds (cps)*
$11.9 < MPA \leq 26.8$ *cps*
$VPA > 26.8$ *cps.*

2.2 Data Processing and Preparation

The collected activity data are continuous. All the data are numerical in format. For the classification, we converted the data from continuous to discrete by removing the timestamp feature. We also removed participants' data which are less than one month, noisy, incomplete, or missing. We decided data for 33 participants as they performed activities more than a month, resulting in 413 records for daily sleep stage classification and 2762 records for hourly sleep stage classification. Normality test with methods, such as Shapiro–Wilk, Anderson–Darling test, and D'Agostino's K^2 [16] on each feature of the datasets revealed that data samples did not look like "Gaussian". The normality test was performed following the hypothesis testing method with P-value > $\alpha = 0.05$ (i.e., sample looks like gaussian) [16].

For the feature selection, we performed methods, such as univariate (e.g., SelectKBest), recursive feature elimination, unsupervised principal component analysis or PCA, feature importance (e.g., ExtraTreesClassifier), modeling ML pipeline with PCA and SelectKBest, and the correlation analysis. The correlation analysis with the "spearman" method revealed the strength of the linear relationship between features [17-20]. We removed features if they showed a powerful dependency score (r >= 0.6). In final, we selected the "sleep time" feature only. Afterward, we created a new feature class, "sleep stage" (on which classification would occur), based on the "sleep time" feature [4]. The "sleep stage" represents three classes – sleep deprivation (0), appropriate sleep (1), and excessive sleep (2) for daily sleep stage classification problem, and two classes – bad sleep (0) and good sleep (1) for hourly sleep stage classification problem. The rule for “sleep stage” feature class creation is defined in Table 1, based on the nature of MOX2-5 data. The feature, such as age, gender, weight, weight-bearing, standing, is not in the scope of this study.

Table 1. The Defined rules for “sleep stage” feature creation for this study.

Classification type	Active class	Rule
Hourly sleep	bad sleep	58 < sedentary minutes AND steps > 2
	good sleep	58 <= sedentary minutes <=60 AND 0 <= steps <= 2
Daily sleep	sleep deprivation	Sleep time < 7 hrs. / day
	appropriate sleep	7 <= Sleep time <= 8 hrs. / day
	excessive sleep	Sleep time > 7 hrs. / day

We used Python 3.8.5 supported language libraries, such as pandas (v. 1.1.3), NumPy (v. 1.21.2), SciPy (v. 1.5.2), Matplotlib (v. 3.3.2), Seaborn (v. 0.11.0), Plotly (v. 5.2.1), scikit-learn or sklearn (v. 0.23.2), and Graph Viz (v. 2.49.1) to process data and build the machine learning models. We set up the intended Python environment in Windows 10 Enterprise system using Anaconda Distribution and used the Spyder 5.x IDE for the development, debugging, and data visualization.

2.3 Model Training and Testing

In this study, all the selected machine learning models for classification are described in Table 2 with corresponding optimization methods. To better use data, initially, we shuffled the dataset, then split the dataset into training and testing with a random state integer value. To boost the performance of the machine learning model, we used k-fold cross-validation where k>=1. Moreover, we adopted Grid Search parameter optimization technique for ML model tuning as appropriate selection of learning rate (alpha (α) and gamma (γ)) in gradient descent, and proper selection of components, such as PCA components, criterion, and max_depth is important for tree-based models. Ensembles [19] can give a boost to ML results in combination with several supervised models based on the approaches, such as parallel ensemble (bagging), sequential ensemble (boosting), and voting. Gradient descent follows a convex optimization technique.

Table 2. Machine Learning Classifier models with optimization methods.

Models	Optimization Method
SVM (kernel = linear or rbf)	Gradient descent
Logistic Regression	Gradient descent
Naïve Bayes (NB)	Gradient descent
Decision Tree (DT)	Information Gain, Gini

K-Nearest Neighbor (KNN)	‘auto’, ‘ball_tree’, ‘kd_tree’, ‘brute’
Random Forest (RF)	Ensemble - Bagging
Linear Discrimination Analysis (LDA)	Gradient descent
Bagging classifier	Ensemble - Bagging
AdaBoost Classifier (ADA)	Ensemble - Boosting
Extra Trees Classifier (ET)	Ensemble - Bagging
Gradient Boosting Classifier (GB)	Ensemble - Boosting
Voting Classifier	Ensemble - Voting

We executed each ML classification model for five times and calculated their mean performance score for comparison. The general pseudocode is stated below:

```
Input: An instance of ML classifier model, mlcSleep
Input: A value-set to train from, value
Input: Necessary parameters for data splitting, param
Input: A value for cross validation, kfold
Input: A value-set for optimization technique, optValue
Input: Number of times model execution, count
Output: Predictions, classified_class, best_params, mean(best_score)
Begin
  value ← shuffle_rows (value)
  X, y ← split (value, param)
  arr ← list ()
  While n < count do
    model ← calculate (mlcSleep, optValue, kfold, 'accuracy')
    model.fit (X, y)
    arr.append(model.classified_class),
    arr.append(model.best_params)
    arr.append(model.best_score)
    n ← n + 1
  end
  return top_five(arr)
end
```

2.4 Model Evaluation Metrics

In this study, performance of a ML-based classification models has been evaluated with discrimination measures. Discrimination metrics are – precision, recall, specificity, accuracy score, F1 score, classification report, and confusion matrix. A confusion matrix is a 2-dimensional table (“actual” vs “predicted”), and both dimensions have “True Positives (TP)”, “False Positives (FP)”, “True Negatives (TN)”, and “False Negatives (FN)” [17-20]. The equations for calculating metrices are [17-20]:

$Accuracy = (TP+TN) / (TP+FP+FN+TN)$,

$Precision\ (P) = TP / (TP+FN)$,

Recall (R) or Sensitivity (S) = TP / (TP+FN),
Specificity = (1 - Sensitivity) = TN / TN+FP,
*F1 score = (2*P*R) / (P+R).*

Also, we used cross-validation score to determine overfitting and underfitting and learning curve to visualize the convergence status of training score with the cross-validation score. We tested if the standardization technique on the entire dataset before learning can improve the performance of the models by reducing data leakage.

2.5 Transfer Learning for Recommendation Generation

The eCoach prototype system aims to collect individual activity data from wearable activity sensors at a daily level (day-n) and classify the sleep data into the identified three classes using machine learning models. In the procedure, participants can set personal preferences (e.g., daily goal, weekly goal, monthly goal, recommendation time, and the mode of recommendation) in the eCoach mobile app for tailored recommendation generation and its delivery. In this study, we have focused on daily sleep goal of 7-9 hrs. to achieve a weekly goal of 49-63 hrs. of sleeping. We considered a good hourly sleep as ≈ 58-60 minutes of sedentary time with IMA ≈ 0-20, step count ≈ 0, and activity time = 0. This study will help participants to identify their hourly sleep as well as daily sleep variation to achieve personalized sleep goals.

Different classification models are available; however, we can't determine "a priori" which classifier will perform the best. It requires enormous data for training, validation, and testing. We collected real-activity data for two adults using the MOX2-5 activity sensor over thirty days. However, that volume of data is not sufficient to determine the accuracy of the best classifier. Therefore, we adopted the concept of transfer learning and incremental training approach. Initially, we trained all the potential classifiers (see Table 2) with public Fitbit data using Kfold = 10 and radom_state = 7. Afterward, we selected the top five best performing classifiers and saved them as pickle files. Then, we used those pre-trained models for individual hourly and daily sleep stage classification. In this study, we classified the MOX2-5 sensor data with collected over 30-days from two participants. The sleep stage classification method entails two steps – a. training of pre-trained models with individual sleep data and model storing for individual participants, and b. sleep classification for day-n with individual models and re-train the models with the individual classification result of that day for the following day (day-n+1) classification. Models trained with personalized sleep data are disjoint with the trained models for other participants. We selected the classification results from the individual classifiers with the highest mean accuracy. The process can be applied to other participant datasets.

3 Experimental Results and Discussion

This section describes - first, the analyses on public Fitbit datasets with ML classifier models, second, the selection of top-five models with their best parameters to train MOX2-5 activity data for personalized sleep classification, and third, the representation of personalized sleep goal achievement in an eCoach prototype mobile app. We prepared public Fitbit data for hourly and daily sleep classifications with 21 different variants of classifiers and corresponding average accuracy scores for five passes described in Table 3. The SVM (kernel = "linear") outpaced other classifiers in the hourly sleep classification, and ExtraTreesClassifier outperformed different classifiers in daily sleep classification. The learning curves for both the highest-ranked classifiers in each category are depicted in Figure 1 and Figure 2. The result shows neither overfit nor underfit. The top five models in the respective category are bold in Table 3 and used for transfer learning as described in Section 2.

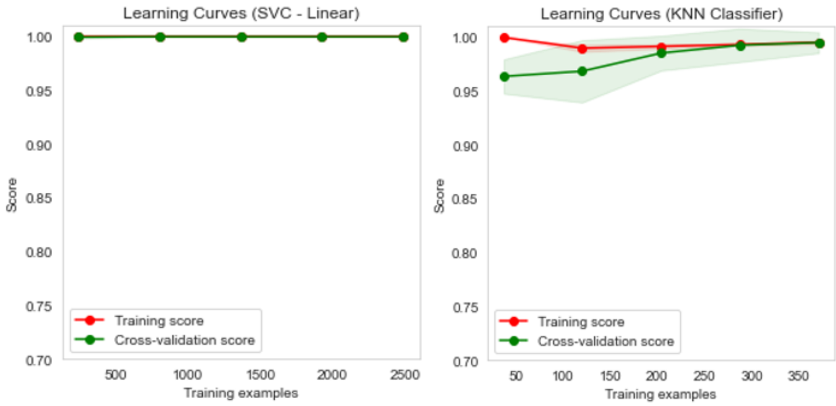

Fig. 1. Learning curve for SVC in hourly sleep classifications, and **Fig. 2.** Learning curve for KNN in daily sleep classifications.

Table 3. Performance of the Machine Learning Classifier models for different classification approaches.

ML classifier models with high level specification	Mean accuracy of hourly sleep classification	Mean accuracy of daily classification
SVC (kernel='linear')	**99.92**	**99.32**
SVC (kernel='rbf')	99.8	98.64
LogisticRegression()	99.8	98.5
GaussianNB ()	95.6	95.1
BernoulliNB ()	85.3	53.0
ComplementNB ()	14.7	44.06
DecisionTreeClassifier (criterion="gini")	99.8	99.18
DecisionTreeClassifier (criterion="entropy")	99.8	99.18
RandomForestClassifier (n_estimators = 25)	99.8	99.18
RandomForestClassifier (n_estimators = 50)	99.8	99.18
RandomForestClassifier (n_estimators = 100)	**99.9**	99.18
KNeighborsClassifier (n_neighbors = 2)	99.8	**99.47**
KNeighborsClassifier (n_neighbors = 4)	99.8	99.47
LinearDiscriminantAnalysis ()	95.0	90.27
BaggingClassifier (base_estimator = Decision-	99.8	99.18

TreeClassifier ())		
AdaBoostClassifier (n_estimators = num_trees, random_state = seed)	**99.9**	**99.18**
ExtraTreesClassifier (n_estimators=25, max_features=max_features)	**99.9**	99.28
ExtraTreesClassifier (n_estimators=50, max_features=max_features)	99.9	**99.32**
ExtraTreesClassifier (n_estimators=100, max_features=max_features)	99.9	99.22
GradientBoostingClassifier	**99.9**	99.17
VotingClassifier (estimators)	99.8	**99.42**

The best optimization parameters (as obtained with grid search method) for those top five models under each category are described in Table 4 and Table 5. Furthermore, during training data preparation we investigated if the pipeline execution concept can improve the performance of the ML classifiers or not! Thus, we created data preparation pipeline models for the best performing classifier. We tried to standardize the whole datasets in each data preparation pipeline and then classify the sleep data. However, the data preparation pipeline improved the performance of the models after addressing typical non-Gaussian nature of datasets. The results of SVC with linear kernel and KNN in pipeline execution are in Table 6.

Table 4. Optimized parameters for top-5 models in Hourly classification.

ML classifier models	Parameter list	Best parameter
ExtraTreesClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=8, criterion = 'gini'
SVC (kernel='linear')	alphas (α) = [0.001, 0.01, 0.1, 1, 10] gammas (γ) = [0.001, 0.01, 0.1, 1]	α = 0.01, γ =0.001
AdaBoostClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=6, criterion = 'gini', α = 0.001
RandomForestClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=2, criterion = 'gini'
GradientBoostingClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=4, criterion = 'gini', α = 0.01

Table 5. Optimized parameters for top-5 models in Daily classification.

ML classifiers	Parameter list	Best parameter
SVC (kernel='linear')	alphas = [0.001, 0.01, 0.1, 1, 10] gammas = [0.001, 0.01, 0.1, 1]	α = 0.01, γ =0.001
ExtraTreesClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=8, criterion = ‘entropy’

KNN	metrics: [‘minkowski’, ‘euclidean’, ‘manhattan’] weights: [‘uniform’, ‘distance’] n_neighbors: [2, 3, 4, 5, 6]	metrics = ‘minkowski’, weights = ‘uniform’, n_ neighbors = 2
AdaBoostClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=4, criterion = 'entropy', α = 0.001
RandomForestClassifier	criterion = ['gini', 'entropy'] max_depth = [2,4,6,8,10,12]	max_depth=2, criterion = 'gini'

Table 6. Result of pipelined ML model execution

Classifier Type	Model	Accuracy score
Hourly data	SVC (kernel = ‘linear’)	99.93
Daily data	KNN	99.51

To prove the practical usefulness of the classifier models, we conceptualized to use them in an eCoach prototype system to achieve personalized activity goals of active sleeping (7-9 hrs. of sleeping/day) for an entire week. We collected activity data from two adult participants over thirty days with the MOX2-5 activity sensor and measured sleep time from features, such as IMA, sedentary time, LPA, and steps. Initially, we assumed that their goal was to maintain active sleeping for an entire week (i.e., the last seven days of the 30 days). Therefore, we used the top five classifiers mentioned in Table 3 to train, validate, and test MAX2-5 datasets for 2 participants and develop personalized ML models using an incremental approach. We divided 30 days into two parts – a. 23 days data for training, and b. remaining seven days data for testing. Based on the training performance up to date-(n-1), we classified individual sleep data for the day-(n) with the best classifier. Next, we trained each participant's five classifiers based on their sleep stage classification result on the day-(n). It helped for sleep stage classification on the day-(n+1). We repeated the same incremental process until the 7-days goal periods got over. The whole result has been captured for the nth day (e.g., n=30) in Table 7 and the last seven days in Table 8.

Table 7. Hourly Sleep stage classification over a Day (MOX2-5 dataset).

Participant	Set of hours (hrs.)	Classifier Model with mean accuracy on day-29	Set of sleep stages
P-1 (Age:34, and Normal weight)	{hr-1, hr-2, hr-3, hr-4, hr-5, hr-6, hr-7, hr-8, hr-9, hr-10}	SVC (kernel = ‘linear’) Mean accuracy = 99.966%	{0, 0, 1, 1, 1, 1, 1, 1, 0, 1}
Text Recommendation		*“You had a good sleep of 7 hours, which is optimal and initial 2 hours the sleep was not perfect”*	
P-2 (Age:40, Obese)	{hr-1, hr-2, hr-3, hr-4, hr-5, hr-6, hr-7, hr-8, hr-9, hr-10}	GradientBoostingClassifier Mean accuracy = 99.96%	{1, 1, 1, 1, 1, 1, 1, 1, 1, 0}
Text Recommendation		*“You had a great sleep of 9 hours, which is suffi-*	

cient and every hour the sleep was proper"

Table 8. Daily Sleep stage classification over a WEEK (MOX2-5 dataset).

Partici-pant	Set of Days	Classifier Model	Accuracy day-(n-1)	Sleep stage on day-n
P-1 (Age:34, Normal weight)	Day-1	KNN	99.51	2
	Day-2	KNN	99.52	1
	Day-3	KNN	99.52	2
	Day-4	KNN	99.52	2
	Day-5	Random Forest	99.52	2
	Day-6	KNN	99.53	2
	Day-7	KNN	99.53	1
Text Recommendation		*"You have slept more than adequate sleeping hours for a week. Try to reduce your sedentary bouts for next week"*		
P-2 (Age:40, Obese)	Day-1	Extra Trees	99.52	1
	Day-2	KNN	99.51	1
	Day-3	KNN	99.52	1
	Day-4	KNN	99.52	2
	Day-5	KNN	99.52	2
	Day-6	KNN	99.53	2
	Day-7	KNN	99.53	2
Text Recommendation		*"You have slept more than adequate sleeping hours for a week. Try to reduce your sedentary bouts for next week"*		

A motivation with an eCoach may improve self-behavior by keeping up an active pace of sleeping over the day or weeks or months. The daily sleep stage will give a reflection on daily sleep status and the hourly sleep stage classification will explain in which hour the sleep was good or bad. In real coaching, to attain a weekly sleep goal, the eCoach module will generate personalized recommendations based on the sleep outcome on each day and followed by a predictive analysis to achieve the weekly or monthly goal. In our future study, we will address it with more participants (N>15).

4 Conclusion

This study has shown a direction to use ML technology to design and develop an intelligent eCoach system to generate automatic, meaningful, evidence-based, and tailored sleep recommendations to attain personal sleep goals. Improvement of physical activity in sequence with wearable activity sensors and digital activity trackers, eCoach features can be encouraging. The concept, such as transfer learning, exists; its re-use with incremental training and testing approach in a sleep eCoaching concept is novel. Moreover, this is the first study conducted on MOX2-5 datasets on sleep monitoring and the conceptualization of tailored recommendation generation. This study has presented a detailed analysis of different ML classifiers on sleep data at a granular level. In the future study, we will focus on classifying leisure time and sleep time from sedentary time based on temporal feature analysis.

References

1. Carden, K.A., 2020. Sleep is essential: a new strategic plan for the American Academy of sleep medicine. Journal of Clinical Sleep Medicine, 16(1), pp.1-2.
2. Vitale, K.C., Owens, R., Hopkins, S.R. and Malhotra, A., 2019. Sleep hygiene for optimizing recovery in athletes: review and recommendations. International journal of sports medicine, 40(08), pp.535-543.
3. Naitoh, P., Kelly, T.L. and Englund, C., 1990. Health effects of sleep deprivation.
4. How Much Sleep Do I Need? Webpage: https://www.cdc.gov/sleep/about_sleep/how_much_sleep.html.
5. The GBD 2015 Obesity Collaborators. Health effects of overweight and obesity in 195 countries over 25 years. New England Journal of Medicine. 12 Jun 2017. DOI: 10.1056/NEJMoa1614362.
6. GBD 2017 Diet Collaborators. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. 2 April 2019. doi: 10.1016/S0140-6736(19)30041-8.
7. Physical activity. Webpage: https://www.who.int/news-room/fact-sheets/detail/physical-activity.
8. R. Robbins, et al, "Four-Year Trends in Sleep Duration and Quality: A Longitudinal Study Using Data from a Commercially Available Sleep Tracker," Journal of Medical Internet Research, 2020, vol. 22(2), e14735.
9. J.M. Kelly at al. "Recent developments in home sleep-monitoring devices," International Scholarly Research Notices, 2012.
10. X. Liu et al. "Wi-sleep: Contactless sleep monitoring via wifi signals," IEEE Real-Time Systems Symposium, 2014, pp 346-366.
11. H. Hong et al. "Microwave sensing and sleep: Noncontact sleep-monitoring technology with microwave biomedical radar," IEEE Microwave Magazine, 2019, vol 20(8), pp 18-29.
12. K.G. Baron, et al, "Technology-Assisted Behavioral Intervention to Extend Sleep Duration: Development and Design of the Sleep Bunny Mobile App," Journal of Medical Internet Research, 2018, vol. 5(1), e3.
13. A. Chatterjee, et al, "Human Coaching Methodologies for Automatic Electronic Coaching (eCoaching) as Behavioral Interventions with Information and Communication Technology: Systematic Review," Journal of Medical Internet Research, 2021, vol. 23(3), e23533.
14. H. Rutjes, et al, "Understanding Effective Coaching on Healthy Lifestyle by Combining Theory-and Data-driven Approaches," PPT@ PERSUASIVE, 2016, pp. 26-29.
15. Crowd-sourced Fitbit datasets 03.12.2016-05.12.2016. Webpage: https://doi.org/10.5281/zenodo.53894.
16. MOX2 Bluetooth LE activity monitor. Webpage: https://www.accelerometry.eu/products/wearable-sensors/mox2/.
17. A.Chatterjee, et al, "Identification of risk factors associated with obesity and overweight—A machine learning overview," Sensors, 2020, Vol. 20(9), pp. 2734.
18. Sklearn Page. Available online: https://scikit-learn.org/stable/supervised_learning.html.
19. A. Chatterjee, et al, "Comparing Performance of Ensemble-Based Machine Learning Algorithms to Identify Potential Obesity Risk Factors from Public Health Datasets," Emerging Technologies in Data Mining and Information Security, Springer, 2021, pp. 253-269.
20. S. Brandt, et al, "Statistical and Computational Methods in Data Analysis," North-Holland Publishing Company, 1970.