IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.0429000

Advanced Classifiers and Feature Reduction for Accurate Insomnia Detection Using Multimodal Dataset

AMEYA CHATUR[1], MOSTAFA HAGHI[2,3], (Member, IEEE), NAGARAJAN GANAPATHY[1], (Member, IEEE), NIMA TAHERINEJAD[3], (Member, IEEE), RALF SEEPOLD[2], (Senior Member, IEEE) and Natividad Martínez Madrid[4], (Senior Member, IEEE)

[1]Department of Biomedical Engineering, Indian Institute of Technology Hyderabad, 502284 Hyderabad, India (e-mail: ameyachatur.12@gmail.com)
[2]Ubiquitous Computing Laboratory, HTWG Konstanz – University of Applied Sciences, 78462 Konstanz, Germany (e-mail: ralf.seepold@htwg-konstanz.de)
[3]EclectX Team, Institute for Computer Engineering, Heidelberg University, 69120 Heidelberg, Germany (e-mail: mostafa.haghi, nima.taherinejad@ziti.uni-heidelberg.de)
[4]Internet of Things Laboratory, School of Informatics, Reutlingen University, 72762 Reutlingen, Germany (e-mail: natividad.martinez@reutlingen-university.de)

Corresponding author: Mostafa Haghi (e-mail: mostafa.haghi@ziti.uni-heidelberg.de).

This work was supported in part by MORPHEUS-Project, a "noninvasive system for measuring parameters relevant to sleep quality" funded by the Carl Zeiss Foundation (project number: P2019-03-003)" and "DAAD Wise" program.

ABSTRACT Sleep deprivation is a significant contributor to various diseases, leading to poor cognitive function, decreased performance, and heart disorders. Insomnia, the most prevalent sleep disorder, requires more effective diagnosis and screening for proper treatment. Actigraphic data and its combination with physiological sensors like electroencephalogram (EEG), electrocardiogram (ECG), and body temperature have proven significant in predicting insomnia using machine learning methods. Studies focusing solely on actigraphic data achieved an accuracy of 84%, combining it with other wearable devices increased accuracy to 88%, and 2-channel EEG alone yielded an accuracy of 92%, but limits scalability and practicality in real-world settings. Here we show that using the hybrid approach of incorporating both recursive feature elimination (RFE) and principal component analysis (PCA) on sleep and heart data features yields outstanding results, with the multi-layer perception (MLP) achieving an accuracy of 95.83% and an F1 score of 0.93. The top-ranked features are predominantly sleep-related and time-domain RR interval. Our findings emphasize the importance of tailoring feature sets and employing appropriate reduction techniques for optimal predictive modeling in sleep-related studies. Our results demonstrate that the ensemble classifiers generalize well on the dataset regardless of the feature count, while other algorithms are hindered by the curse of dimensionality.

INDEX TERMS Actigraphy, classification, feature reduction, heart rate variability, insomnia.

I. INTRODUCTION

SLEEP is a regular, cyclical condition of decreased responsiveness and sensitivity to external stimulation involving intricate physiological changes [1]. During sleep, the body undergoes synchronized brain activity, hormone changes, muscle relaxation, and temperature and blood pressure reductions. It is crucial for various brain processes, including neuron interactions and the creation of new neural pathways, i.e. a dynamic and complex process [2].

Sleep deprivation can lead to fatigue, impaired memory and attention, mood swings, slow judgment and reaction times, and uncoordinated movements. Chronic sleep deprivation increases the risk of high blood pressure, cardiovascular diseases, diabetes, depression, and obesity [3]. Given the extensive impact of sleep on the body, it's not surprising that many illnesses are linked to sleep deprivation. Insomnia is the most common sleep disorder, according to the American Psychiatric Association [4].

According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), insomnia is characterized by difficulty falling asleep, staying asleep, or experiencing early morning awakenings. These sleep issues must be accompanied by daytime dysfunctions, such as reduced attention or concentration difficulties. For a diagnosis of insomnia disorder, these symptoms must occur at least three times per week for a minimum of three months [5].

Hyperarousal, often associated with insomnia, is influenced by various physiological, psychological, behavioral, and cognitive factors. It is a persistent feature in individuals with insomnia, manifesting both at night and during the day, and includes physiological, cognitive, and emotional components [6]. Arousal is characterized by somatic, cognitive, and cortical activation. Individuals with insomnia are particularly sensitive to disturbances from environmental or other stimuli at sleep onset and during sleep due to enhanced sensory processing. These disturbances can directly interfere with sleep initiation and maintenance. Arousals during sleep and heightened memory formation around sleep onset can negatively impact the subjective experience of restful, undisturbed sleep [7].

Hyperarousal in insomnia is characterized by elevated heart rate, body temperature, metabolic rate, cortisol secretion, and baseline skin resistance compared to healthy sleepers. Studies have shown that the heart rate of individuals with insomnia is significantly higher than that of normal sleepers [8]. Heart rate variability (HRV) is the variation in the time interval between heartbeats and measures neurocardiac function, originating from heart-brain interactions and autonomic nervous system (ANS) processes. It assesses ANS function, reflecting the balance between parasympathetic and sympathetic activity. Low frequency (LF: 0.04–0.15 Hz) reflects parasympathetic and sympathetic activity, while high frequency (HF: 0.15–0.4 Hz) is influenced by parasympathetic activity. The LF/HF ratio indicates sympathovagal balance. Higher LF and LF/HF ratios are associated with poorer sleep quality and more severe insomnia symptoms, indicating increased sympathetic nervous system activity [9], [10].

Actigraphy, or activity-based sleep-wake monitoring, is a non-invasive method for tracking human wakefulness and sleep cycles using a wrist-worn device with an accelerometer to record movements. It is widely used in sleep medicine to assess sleep quality and diagnose sleep disorders. While not as precise as polysomnography (PSG), actigraphy provides reliable estimates of sleep patterns, including sleep latency, total sleep time, and sleep efficiency [11]. The Cole-Kripke algorithm developed to validate automatic scoring methods that distinguish sleep from wakefulness based on wrist activity, uses a weighted moving average instead of a fixed score threshold to classify each epoch as sleep or wake. In a diverse sample, including healthy controls, older adults, individuals with sleep disorders, and psychiatric patients, the Cole-Kripke Algorithm shows actigraphic estimates of sleep latency, total sleep time, and sleep efficiency that are highly correlated with PSG scores [12].

Byun et al. in [13] demonstrated a machine learning-based diagnosis of major depression using HRV. They employed a linear support vector machine (SVM) classifier for its good generalization performance with high-dimensional data and reduced risk of overfitting through appropriate regularization. The authors also used SVM-recursive feature elimination (RFE) for feature selection, which performs well when the number of samples is smaller than the number of features and has been applied to various problems [14].

Depression is linked to autonomic nervous system (ANS) dysfunction, which can be assessed using HRV. Depressed patients typically have lower HRV than healthy individuals, allowing HRV to distinguish between the two groups. Bayesian Networks, trained on the time domain, frequency domain, and non-linear HRV features, have effectively identified depression patients with 86.4% accuracy, 89.5% sensitivity, and 84.2% specificity, using the root mean square of successive differences (RMSSD) of the HRV [15].

Aoyu et al. [16] found that patients with mild cognitive impairment (MCI) exhibited lower HRV. They used five machine learning algorithms—K-nearest neighbor (KNN), Decision Tree (DT), Random Forest (RF), Naive Bayes, and eXtreme Gradient Boosting (XGB)—to classify healthy individuals and MCI patients. Standard deviation of all normal to normal RR intervals (SDNN), the proportion of pairs of successive RR intervals that differ by more than 50 ms relative to the total number of RR intervals (pNN50), and LF were identified as the most important predictors. By employing a weighted soft voting strategy that combined the outputs of the five classifiers, they achieved 88.9% accuracy, 89.9% precision, 88.2% recall, and an F1 score of 89.0%.

A study by Delmastro et al. [17] used Bayesian Networks, SVM, kNN, DTs, RFs, and AdaBoost (AB) to detect stress using HRV and electrodermal activity (EDA) as markers. RF and AdaBoost outperformed the other classifiers, effectively combining multiple models through bagging and boosting methods to improve predictions.

Rossi et al. in [18] applied fifteen machine learning algorithms using 14 features to predict insomnia. They reported that SVM was the best estimator, achieving 91.6% accuracy and an F1 score of 0.92. The study concluded that vision, movement, and sleep disorders were the primary factors contributing to insomnia.

In [19], authors created an insomnia dataset based on questionnaires focused on external symptoms. They tested seven algorithms: SVM, DTs, Logistic Regression (LR), RF, KNN, Naïve Bayes, and Stochastic Gradient Descent. LR emerged as the best classifier with 98% accuracy.

Huang et al. [20] used the National Health and Nutrition Examination Survey (NHANES) to identify risk factors for sleep disorders with machine learning. They analyzed data from demographic, dietary, exercise, and mental health questionnaires and lab and physical exams. Four methods were used: XGB, RF, Adaptive Boost, and artificial neural network (ANN), all with 10-fold cross-validation. XGB outperformed the other models in terms of accuracy. The top five features identified were the Patient Health Questionnaire, depression survey, age, physician recommendation of exercise, weight, and waist circumference.

A recent study [21] aimed to develop an insomnia prediction model using electronic medical records (EMR) from 2011-2018, obtained from a statewide health information exchange. The authors evaluated five machine learning models: LR, SVM with radial basis function kernel, RF, XGB, and

a multilayer neural network. Each model's performance was assessed using the area under the receiver operating characteristic curve (AUC) on a holdout set. The XGB model, which used a combination of demographics, diagnosis, and medication data, achieved the highest average AUC of 0.80 and was selected as the best model.

To this end, the physiological domain, particularly heart rate and HRV, and sleep physical parameters acquired from actigraphy have shown significance in predicting insomnia using machine learning methods. However, most studies have focused on only one domain, which is a limitation. Combining both feature domains could potentially enhance performance and improve prediction accuracy.

This study aims to diagnose insomnia via objective data acquisition from wearable devices supporting physiological and physical sleep-related parameters. We develop an insomnia classifier using multimodal data comprising actigraphy and individual heart rate, RR interval, and HRV measurements. We utilize the publicly available Multilevel Monitoring of Activity and Sleep in Healthy People (MMASH) dataset, employing Pittsburgh Sleep Quality Index (PSQI) scores as ground truth for prediction. The task is binary classification: predicting whether a subject suffers from insomnia based on their PSQI score. A PSQI score of 6 or higher indicates poor sleep and insomnia, while a score below 6 indicates a healthy subject. We exclude anthropometric data due to its poor correlation with PSQI scores.

The rest of this paper is organized as follows: Section II provides the dataset and details the proposed methodology, including the feature extraction and classification models. Section III presents the results and performance evaluation. Finally, Section IV offers a discussion of the findings, and Section V concludes the paper.

II. MATERIALS AND METHODS

A. DATABASE DESCRIPTION

1) Participants and demographic

We used the MMASH dataset, publicly available on PhysioNet [18]. This dataset includes 24 hours of continuous beat-to-beat heart data, triaxial accelerometer data, sleep quality, physical activity, and psychological characteristics (e.g., anxiety status, stress events, and emotions) for 22 healthy participants. The participants' anthropomorphic characteristics (i.e., age, height, and weight) were recorded at the beginning of the experiment. The participants have an age range of 27.29 ± 4.21 years, a height range of 179.91 ± 8.22 cm, and a weight range of 75.05 ± 12.79 kg.

2) Data collection and monitoring devices

The participants filled in a set of initial questionnaires that provided information about their psychological status: morningness-eveningness questionnaire (MEQ), state-trait anxiety inventory (STAI-Y), PSQI, and behavioral avoidance/inhibition (BIS/BAS). During the test, participants wore a heart rate monitor to record heartbeats and beat-to-beat intervals and an actigraph to record actigraphy information such as sleep quality and physical activity. The participants wore these two devices for 24 hours.

The Polar H7 heart rate monitor (Polar Electro Inc., Bethpage, NY, USA) is a Bluetooth low-energy chest strap with an electrocardiogram (ECG) sensor that provides information about inter-beat intervals (IBI). The ActiGraph wGT3X-BT (ActiGraph LLC, Pensacola, FL, USA) is a triaxial accelerometer and one of the most commonly used devices for assessing physical activity. The sensor's dimensions are $4.6 \times 3.3 \times 1.5$ cm, with a weight of 19 g and a frequency range between 30 and 100 Hz. The accelerometer has a dynamic range of ± 8 g and a precision of 12 bits.

3) Data processing

The sleep data was processed using the Cole-Kripke algorithm [18] from the actigraph data provided in the dataset. The inter-beat interval (IBI) data was used to derive HRV, offering insights into the sympathetic-parasympathetic balance of cardiac vagal tone, an indicator of cognitive, emotional, social, and health status [19]. PSQI is a tool used to measure sleep quality and patterns, distinguishing between "poor" and "good" sleep across seven domains: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction. The PSQI score ranges from 0 to 21 [22]–[24].

B. FEATURE EXTRACTION

In this study, we used beat-to-beat heart data, accelerometer data, and sleep quality characteristics (Table 1). To assess sleep quality, we employed PSQI scores, categorizing subjects as having "good" sleep if their PSQI score was below six and "poor" sleep otherwise. The PSQI's applicability in insomnia screening is well-demonstrated across various populations, showing high test-retest reliability and good validity for patients with primary insomnia. This classification approach resulted in a class imbalance, with a higher number of healthy subjects having good sleep scores compared to those with poor sleep quality.

We derived sleep variables from actigraph data using the Cole-Kripke algorithm. These variables included: In Bed Time, Out Bed Time, Onset Time, Latency, Efficiency, Total Minutes in Bed, Total Sleep Time (TST), Wake After Sleep Onset (WASO), Number of Awakenings, Average Awakening Length, Movement Index, Fragmentation Index, and Sleep Fragmentation Index.

We used beat-to-beat heart data, consisting of RR intervals, heart rate (HR), and HRV features in the time domain, frequency domain, and nonlinear domain. For the RR interval in the time domain, we calculated RMSSD, SDNN, PNN50, the standard deviation of differences between adjacent RR intervals (SDSD), median absolute values of the successive differences between the RR intervals, and the difference between the maximum and minimum RR intervals. We also included the mean, minimum, maximum, and standard deviation (SD) of HR.

TABLE 1. Descriptions of Sleep and Heart data Features Used to Develop the Machine Learning Models

Features	Description
Sleep	
In Bed Time	Time at which the subject went to bed in minutes
Out Bed Time	Time at which the subject got out of bed in minutes
Onset Time	Time at which the subject fell asleep in minutes
Latency	Time taken by the subject to fall asleep after getting into bed in minutes
Total Minutes in Bed	Minutes spent in bed per night
Total Sleep Time (TST)	Length of sleep per night expressed in minutes
Efficiency	Percentage of total sleep time of the subject to the total time in bed
Wake After Sleep Onset (WASO)	Time spent awake after falling asleep for the first time
Number of Awakenings	Number of awakenings during the night
Average Awakening Length	Time in seconds spent awakening during the night
Movement Index	The number of minutes without movement expressed as a percentage of the movement phase (i.e., the number of periods with arm movement)
Fragmentation Index	The number of minutes with movement expressed as a percentage of the immobile phase (i.e., the number of periods without arm movement)
Sleep Fragmentation Index	The ratio of the movement and fragmentation indices
Heart Rate	
mean_nni	The mean of RR-intervals
sdnn	The standard deviation of the time interval between successive normal heart beats (i.e., the RR-intervals)
pnni_50	The proportion derived by dividing the number of interval differences of successive RR-intervals greater than 50 ms by the total number of RR-intervals
rmssd	The square root of the mean of the sum of the squares of differences between adjacent NN-intervals
median_nni	Median absolute values of the successive differences between the RR-intervals
range_nni	Difference between the maximum and minimum nn_interval
mean_hr	The mean heart rate
max_hr	Maximum heart rate
min_hr	Minimum heart rate
std_hr	Standard deviation of heart rate
sdsd	The standard deviation of differences between adjacent RR-intervals
lf	Power in HRV in the low frequency band (0.04 to 0.15 Hz)
hf	Power in HRV in the high frequency band (0.15 to 0.40 Hz)
lf_hf_ratio	LF to HF ratio
lfnu	Normalized LF power
hfnu	Normalized HF power
total_power	Total power density spectral
vlf	Power in HRV in the very low frequency band (0.003 to 0.04 Hz)
sd1	The standard deviation of projection of the Poincaré plot on the line perpendicular to the line of identity
sd2	The standard deviation of the projection of the Poincaré plot on the line of identity (y=x)
ratio_sd2_sd1	Ratio between SD2 and SD1

In the frequency domain, we used the following HRV features: very-low-frequency power (VLF, below 0.04 Hz), low-frequency power (LF, 0.04-0.15 Hz) and its normalized power, high-frequency power (HF, 0.15-0.4 Hz) and its normalized power, the ratio of LF to HF, and total power (sum of power across all these frequency ranges). These features were obtained by integrating the power in the appropriate frequency ranges in the spectrum. This approach is essential because HRV data naturally consist of unevenly spaced intervals, and the sympathetic and parasympathetic nervous systems impact different portions of the spectrum.

Additionally, we used SD of the Poincaré plot, a type of recurrence plot used to quantify self-similarity in processes. This plot graphs HRV(n) on the x-axis versus HRV(n+1) on the y-axis. SD1 is the standard deviation of the Poincaré plot perpendicular to the line of identity, while SD2 is the standard deviation along the line of identity. Anthropometric data were excluded due to their poor correlation with PSQI scores.

C. DIMENSION REDUCTION

We applied principal component analysis (PCA), RFE, and a combination of both to reduce dimensionality. Additionally, we used neural networks (NN) to achieve further dimension reduction.

1) PCA

PCA aims to reduce the dimensionality of a dataset with many correlated variables while retaining as much variation as possible. The dataset, structured as a matrix with variables in columns and subjects in rows, is transformed into a new set of variables called principal components [25].

2) RFE

RFE reduces features by iteratively removing the least important ones from the dataset until the desired number remains. Initially, the model is trained on all features, and their importance is evaluated. The least important features are then discarded, and the model is retrained. This process repeats until the target number of features is achieved. Feature importance can be determined by the machine learning model or statistical methods, which evaluate the relationship between input variables and the target variable [26].

3) PCA and RFE

This process results in two types of datasets. The first applies PCA to the entire dataset for dimensionality reduction, followed by RFE to select the most important principal components. The second dataset uses RFE first to select the most important variables, which are then transformed into principal components using PCA [27].

4) Neural Network (NN)

Feature extraction using neural networks (NNs) involves training an NN on the dataset, then removing the final prediction layer. The output from the penultimate layer is used to transform the data, creating new features based on the learned representations. With fewer logits from the last hidden layer, these new features are then used to train a classifier [28].

D. CLASSIFIERS

We considered KNN, DTs, RFs, SVM, MLP, XGB, and Extreme Learning Machine (ELM) machine learning classifiers.

1) KNN

KNN is a non-parametric, supervised method used for classification and regression. It predicts the label of a new point by finding the K closest training samples, using Euclidean distance, with uniform weights. We determined $K = 3$ through cross-validation [29].

2) Decision Tree

DTs are non-parametric, supervised learning methods suitable for numerical and categorical data. They use a tree structure where nodes represent features, branches represent decision rules, and leaves represent outcomes. We used Gini impurity for the attribute selection measure [30].

3) Random Forest

RFs are ensemble learning methods that combine multiple DTs to improve prediction accuracy and control over-fitting. Each tree is built from a sample drawn with replacement, using Gini impurity as the split criterion [31].

4) SVM

SVMs are supervised methods effective in high-dimensional spaces. They aim to find the optimal hyperplane that maximizes the margin between classes. We used a linear kernel for our SVM [32].

5) MLP

MLP is a fully connected neural network with one input layer, one output layer, and multiple hidden layers. Each neuron in the hidden layer applies a weighted linear summation followed by a ReLU activation function [33].

6) XGB

XGB is an ensemble method that builds models sequentially, each correcting the errors of the previous one. This process continues until the training set is accurately predicted or a maximum number of models are added [34].

7) ELM

ELM is a single-hidden-layer feedforward neural network used for classification and regression. It uniquely trains without iterative tuning, using randomly initialized weights and biases for the hidden layer and calculating output layer weights with the Moore-Penrose inverse [35].

E. EVALUATION METRIC

Since the dataset was imbalanced, we evaluated the classifiers using Precision, Recall, and F1 scores in addition to Accuracy. These metrics ensure comprehensive performance evaluation and prevent the classifier from predicting the majority class exclusively.

Precision (P) measures the correctness of positive predictions:

$$P = \frac{TP}{TP + FP}$$

where TP is true positives and FP is false positives. It indicates how many flagged insomniacs are suffering from insomnia.

Recall (R) measures the completeness of positive predictions:

$$R = \frac{TP}{TP + FN}$$

where FN is false negatives. It shows how many insomniacs were correctly flagged out of the total insomniacs.

The F1 score (F_1) combines Precision and Recall:

$$F_1 = 2 \cdot \frac{P \cdot R}{P + R}$$

It is the harmonic mean of Precision and Recall, providing a balanced metric that is high only when both are high.

Classifiers were trained using exhaustive Grid Search for hyperparameter tuning with stratified K-fold cross-validation

(3, 5, and 8 folds). Stratified K-fold ensures each split maintains the class distribution, crucial for imbalanced datasets. Accuracy on the test data was used to select the best estimator parameters. The final estimators were trained on the split that provided the best accuracy and F1 score.

III. RESULTS

A. SLEEP FEATURES

The study showed that for sleep variables as independent variables and PSQI score as the dependent variable, the models achieved the highest accuracy of 80% among all the folds with a maximum F1 score of 0.53. We present the results based on the domains that its features used or the combination of features from multiple domains. SVM and RF achieved the highest classification accuracy of 80% under the 3-fold cross-validation configuration. This result is consistent with the findings reported in [36], where SVM also demonstrated the highest accuracy (Table 2).

B. HEART DATA FEATURES

Evaluation of the models using beat-to-beat interval features (time domain, frequency domain, and nonlinear) indicated that the highest accuracy of 86% was achieved by the MLP, with a 3-fold cross-validation and an F1 score of 0.71. This represents an improvement compared to the earlier observations using sleep features. Notably, most classifiers reported lower accuracy than the previous models. Additionally, LR exhibited poorer performance as the number of features increased (Table 3).

C. COMBINATION OF SLEEP AND HEART DATA FEATURES WITHOUT FEATURE REDUCTION

Using the combination of sleep and heart data features, the MLP once again reported the highest accuracy at 80%, while the ELM achieved the highest F1 score of 0.67. Since the total number of features exceeded the number of samples, some classifiers, such as KNN, which are susceptible to the curse of dimensionality, recorded poor performance (Table 4). To address these dimensionality issues, we employed feature reduction methods, including PCA and RFE.

D. PCA

We performed PCA on the dataset containing both sleep and heart data features to reduce the number of independent variables. However, this did not impact the performance metrics, as the highest accuracy remained around 80%, achieved by the MLP (Table 5).

E. RFE

Furthermore, we implemented feature selection methods using RFE. Consequently, the accuracy for models such as KNN increased from around 70% to 80%. As presented in Table 6, the DT achieved its highest accuracy of 85%, an improvement compared to the previous 80% accuracy, while ELM yielded the highest F1 score of 0.8. This improvement is due to the fundamental approach of RFE, which aims to select features by recursively considering smaller sets of features.

F. PCA AND RFE

While PCA alone did not significantly impact performance, the combination of RFE and PCA yielded improved accuracy for various classifiers, showcasing the effectiveness of these methods. The application of RFE and PCA, in combination with both sleep and heart data features, resulted in the best performance across all classifiers. LR and SVM outperformed the other models for 3-fold cross-validation. This was valid when the features consisted of independent components extracted using PCA from higher-ranked features by RFE, whereas DTs and MLP delivered higher accuracies for 3-fold cross-validation when the higher-ranked independent components were used as input features.

Choosing features using RFE followed by PCA on the top-ranked features recorded the best performance across all classifiers, with 95.83% accuracy achieved by MLP with 5-fold cross-validation. The F1 scores were also higher than those achieved with previous feature sets. When considering F1 scores as the most influential criterion for performance, the MLP with 3-fold cross-validation emerged as the best classifier, with a score of 0.93. This suggests that neural networks and perceptrons can learn representative features and generalize better on the dataset. Consequently, we developed neural networks as the next approach for feature reduction (Table 7 and 8).

G. NEURAL NETWORK

To perform feature reduction using neural networks (NNs), we leveraged the representations learned by the network to extract new features from the samples. These representations exhibited strong correlations with PSQI scores, surpassing the performance of traditional sleep variables and heart rate features (Table 9).

H. SIGNIFICANT FEATURES

Upon performing RFE, the top-ranked features were predominantly sleep-related, including Total Minutes in Bed, Number of Awakenings, Average Awakening Length, In Bed Time, Total Sleep Time (TST), and Fragmentation Index. Additionally, key HRV features were identified, such as the median, mean, and standard deviation of the RR intervals, as well as the proportion of RR intervals with an interval greater than 50 ms.

IV. DISCUSSION

Insomnia is one of the most common sleep disorders. Diagnosing insomnia is of concern as it allows for targeted treatment strategies to enhance sleep quality and overall health. Early diagnosis also contributes to identifying potential underlying conditions contributing to sleep disturbances [37]. To improve its diagnosis, we aimed to develop a classifier to predict whether a subject suffers from the disease. Given the inconveniences associated with surveys and questionnaires,

TABLE 2. Classification Metrics of Classifiers Trained on Sleep Features. SVM Outperforms the Other Classifiers in Terms of Accuracy and F1 Score

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.17	0.3	0.25	0.17	0.4	0.25	0.17	0.33	0.25	65.08	70	72.92
KNN	0	0	0	0	0	0	0	0	0	69.84	70	70.83
DT	0.33	0	0.38	0.33	0	0.38	0.33	0	0.38	79.37	70	75
RF	0.67	0.2	0.12	0.33	0.2	0.12	0.44	0.2	0.12	80.16	75	75
XGB	0.33	0.2	0.12	0.17	0.2	0.12	0.22	0.2	0.12	75.4	75	77.08
MLP	0.22	0.3	0.25	0.33	0.4	0.25	0.27	0.33	0.25	75.4	75	79.17
SVM	0.44	0.5	0.44	0.67	0.6	0.5	0.53	0.53	0.46	80.16	80	79.17
ELM	0	0.99	0.99	0	0.99	0.99	0	0.99	0.99	50	99	99

TABLE 3. Classification Metrics of Classifiers Trained on Heart Data Features: MLP Improves Overall Performance in Both F1 Score and Accuracy Compared to Sleep Features

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.51	0.3	0.12	0.44	0.3	0.12	0.4	0.27	0.12	53.57	58	58.33
KNN	0.17	0.2	0.06	0.17	0.2	0.12	0.17	0.2	0.08	68.45	65	68.75
DT	0.5	0.2	0.25	0.28	0.1	0.25	0.36	0.13	0.25	73.21	73	68.75
RF	0.67	0.2	0.25	0.28	0.1	0.25	0.39	0.13	0.25	77.38	73	72.92
XGB	0	0	0	0	0	0	0	0	0	68.45	69	68.75
MLP	0.99	0.4	0.38	0.56	0.4	0.38	0.71	0.4	0.38	86.31	79	79.17
SVM	0	0	0	0	0	0	0	0	0	68.45	69	68.75
ELM	0.5	0.99	0	0.5	0.99	0	0.5	0.99	0	71.43	99	66.67

TABLE 4. Classification Metrics of Classifiers Trained on Sleep and Heart Data Features: Combining Both Data Types Without Feature Reduction Does Not Yield Improvement

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.25	0.27	0.12	0.33	0.6	0.12	0.28	0.37	0.12	60.32	60	64.58
KNN	0.17	0	0	0.17	0	0	0.17	0	0	69.84	70	70.83
DT	0.17	0.5	0.25	0.17	0.5	0.25	0.17	0.47	0.25	69.84	80	72.92
RF	0.33	0.2	0.25	0.17	0.2	0.25	0.22	0.2	0.25	74.6	75	79.17
XGB	0.33	0.2	0.12	0.17	0.2	0.12	0.22	0.2	0.12	75.4	75	77.08
MLP	0.56	0.4	0.12	0.5	0.4	0.12	0.49	0.4	0.12	80.16	80	77.08
SVM	0	0.13	0.25	0	0.2	0.25	0	0.16	0.25	69.84	75	72.92
ELM	0.99	0	0.99	0.5	0	0.99	0.67	0	0.99	83.33	50	99

TABLE 5. Classification Metrics from Classifiers When Trained on Independent Components Extracted from Sleep and Heart Data Features Using PCA

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.28	0.27	0.35	0.67	0.8	0.62	0.39	0.39	0.44	53.97	40	50
KNN	0.17	0	0	0.17	0	0	0.17	0	0	69.84	70	70.83
DT	0.5	0	0.12	0.5	0	0.12	0.44	0	0.12	74.6	70	75
RF	0.33	0.2	0	0.17	0.2	0	0.22	0.2	0	74.6	75	70.83
XGB	0.5	0.55	0.38	0.83	0.7	0.62	0.61	0.58	0.46	69.84	65	70.83
MLP	0.5	0.67	0.31	0.5	0.7	0.38	0.5	0.63	0.33	76.19	80	81.25
SVM	0	0.13	0.25	0	0.2	0.25	0	0.16	0.25	69.84	75	72.92
ELM	0.4	0.99	0	0.99	0.99	0	0.57	0.99	0	57.14	99	99

TABLE 6. Classification Metrics from Classifiers When Trained on Top-Ranked Sleep and Heart Data Features Using RFE

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0	0	0	0	0	0	0	0	0	69.84	70	70.83
KNN	0.67	0	0	0.33	0	0	0.44	0	0	80.16	70	70.83
DT	0.89	0.5	0.25	0.67	0.6	0.25	0.71	0.53	0.25	84.92	85	75
RF	0.5	0.4	0	0.5	0.4	0	0.5	0.4	0	79.37	80	70.83
XGB	0.83	0.4	0.12	0.5	0.4	0.12	0.61	0.4	0.12	80.16	75	77.08
MLP	0.33	0.4	0.12	0.17	0.3	0.12	0.22	0.33	0.12	75.4	80	77.08
SVM	0	0.1	0	0	0.1	0	0	0.1	0	69.84	70	70.83
ELM	0.67	0.99	0.99	0.99	0.5	0.99	0.8	0.67	0.99	85.71	75	99

TABLE 7. Classification Metrics from Classifiers When Trained on Top-Ranked Independent Components Extracted from Sleep and Heart Data Features Using PCA

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.72	0.77	0.5	0.83	0.99	0.62	0.77	0.83	0.54	69.84	85	87.5
KNN	0.56	0.2	0.38	0.5	0.1	0.38	0.49	0.13	0.38	80.16	70	83.33
DT	0.5	0.4	0.38	0.5	0.4	0.38	0.44	0.4	0.38	84.92	85	83.33
RF	0.67	0.4	0.38	0.5	0.3	0.38	0.56	0.33	0.38	79.37	80	83.33
XGB	0.78	0.7	0.31	0.83	0.8	0.38	0.76	0.73	0.33	80.16	85	79.17
MLP	0.89	0.8	0.69	0.83	0.991	0.75	0.82	0.87	0.71	75.4	90	95.83
SVM	0.83	0.7	0.56	0.67	0.8	0.62	0.67	0.73	0.58	69.84	90	91.67
ELM	0.99	0	0.99	0.5	0	0.99	0.67	0	0.99	85.71	75	99

TABLE 8. Classification Metrics from Classifiers When Trained on Independent Components Extracted from Top-Ranked Sleep and Heart Data Features Using PCA

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.72	0.4	0.5	0.83	0.5	0.62	0.77	0.43	0.54	85.71	75	87.5
KNN	0.56	0.2	0.38	0.5	0.1	0.38	0.49	0.13	0.38	79.37	70	83.33
DT	0.5	0.5	0.38	0.5	0.6	0.38	0.44	0.53	0.38	74.6	85	83.33
RF	0.33	0.4	0.38	0.33	0.3	0.38	0.33	0.33	0.38	79.37	80	83.33
XGB	0.78	0.27	0.31	0.83	0.4	0.38	0.76	0.3	0.33	84.92	75	79.17
MLP	0.89	0.9	0.69	0.99	0.9	0.75	0.93	0.87	0.71	95.24	90	95.83
SVM	0.5	0.7	0.56	0.33	0.8	0.62	0.39	0.73	0.58	75.4	90	91.67
ELM	0.99	0.99	0	0.5	0.99	0	0.67	0.99	0	83.33	99	99

TABLE 9. Classification Metrics from Classifiers When Trained on Features Learned by Neural Networks

Metrics	Precision			Recall			F1			Accuracy		
Models\Folds	3	5	8	3	5	8	3	5	8	3	5	8
LR	0.99	0.99	0.62	0.83	0.9	0.62	0.89	0.93	0.62	95.24	95	95.83
KNN	0.99	0.99	0.62	0.83	0.9	0.62	0.89	0.93	0.62	95.24	95	95.83
DT	0.89	0.99	0.62	0.83	0.9	0.62	0.82	0.93	0.62	89.68	95	89.58
RF	0.89	0.9	0.62	0.67	0.9	0.62	0.71	0.87	0.62	84.92	90	83.33
XGB	0.67	0.99	0.62	0.5	0.9	0.62	0.56	0.93	0.62	85.71	95	83.33
MLP	0.99	0.99	0.62	0.83	0.9	0.62	0.89	0.93	0.62	95.24	95	95.83
SVM	0.89	0.99	0.62	0.99	0.9	0.62	0.93	0.93	0.62	94.44	95	95.83
ELM	0.99	0.99	0	0.5	0.5	0	0.67	0.67	0	85.71	75	99

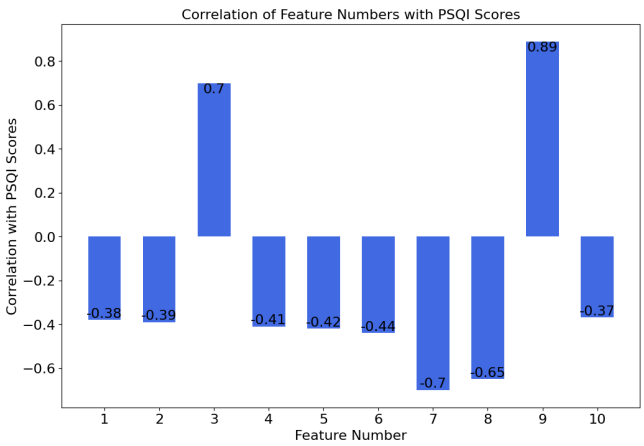

FIGURE 1. **Correlation scores of feature representations from the NN with the PSQI scores. TST and HR standard deviation are the most relevant features in insomnia detection.**

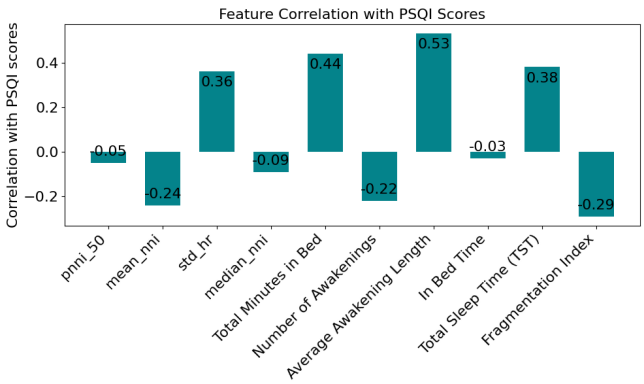

FIGURE 2. **Correlation scores of the top-ranked features given by RFE with the PSQI scores. Sleep features contribute significantly more to insomnia detection than heart rate data features.**

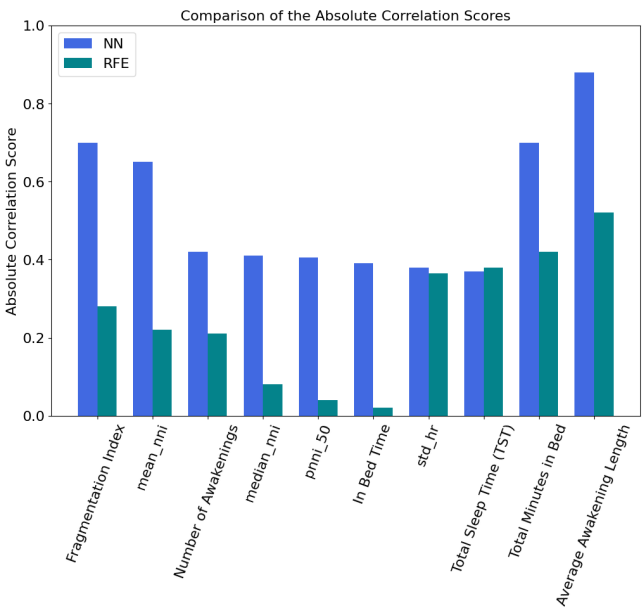

FIGURE 3. **Comparison of the absolute correlation scores of the top-ranked features given by RFE and the NN with the PSQI scores.**

such as poor self-evaluation and high participant dropout rates, we used the MMASH dataset from PhysioNet, which contains time series data from actigraphy and heart data sensors.

Furthermore, we derived sleep variables from the time series data using the Cole-Kripke algorithm. We used PSQI scores as the dependent variable, given their high correlation with sleep quality and insomnia. Bitkina et al. [36] developed various algorithms using these sleep variables and PSQI scores to model sleep quality, achieving a satisfactory accuracy of approximately 80–86%.

To further enhance the model, we included HRV variables, which correlate well with PSQI scores. These variables were derived from frequency domain, time domain, and nonlinear HRV data analysis. Additionally, we employed ensemble classifiers such as DTs, RFs, and XGB and algorithms similar to NNs, such as MLPs and ELMs.

The size of the input features caused many classifiers to perform poorly due to the curse of dimensionality. To address this issue, we employed several feature reduction techniques based on linear algebra and statistical principles, including PCA, to derive the most critical components from the dataset. Additionally, we employed statistical methods to recursively analyze the correlation with the target variables, ranking the features and eliminating those that did not contribute to the performance.

We also applied multi-layered NNs for feature learning, aiming to extract specific features from the dataset. This allowed us to compare traditional manual feature selection approaches with a fully automated approach independent of domain knowledge. These techniques offer advantages over conventional methods, such as improved efficiency, adaptability, and higher accuracy. However, they often require significant computational resources, have low interpretability, and are prone to overfitting [38]. To mitigate overfitting, we employed standard practices such as normalization, regularization, and learning rate decay methods [39].

The newly learned features demonstrated a good correlation with PSQI scores and performed well, comparable to other feature reduction techniques. From the correlation plots, these new features exhibited higher correlation magnitudes with PSQI scores compared to the top-ranked features obtained through recursive feature elimination (Figures 1, 2, and 3). Using these strategies, we achieved accuracy exceeding 95% and F1 scores of 0.93.

The ensemble classifiers generalized the dataset well and showed the highest performance regardless of the number of features. Consequently, combining feature selection techniques and ensemble methods has improved the insomnia classifiers' performance. Using objective data collected from wearable devices, along with high-performance classifiers, shows promising results in the timely diagnosis and detection of insomnia [40], [41].

Classifiers such as LR, KNN, and SVM, which initially delivered poor performance compared to tree-based and ensemble models like DTs, RF, and XGBoost, showed signifi-

cant improvement (see Table 9). We observed that the highest accuracy of 95.83% was achieved by simpler classifiers such as LR, KNN, SVM, and MLP. These classifiers also attained an F1 score of 0.93, matching the performance obtained using PCA and RFE.

A. COMPARISON WITH THE EXISTING WORK

Several studies on insomnia detection using actigraphic data highlight the effectiveness and complexity of different machine-learning models and data modalities. Studies using actigraphic data alone, such as those employing SVM and RF models, showed an accuracy of up to 84%, demonstrating the practical utility of actigraphy in monitoring sleep patterns. However, integrating more complex electroencephalography (EEG) data with deep neural networks (DNNs) and feedforward neural networks (FNNs) can enhance accuracy up to 92%, but at the cost of increased complexity, the need for more obtrusive equipment,reduced subject comfort, and increased data processing requirements, limiting their practical application in large-scale or everyday settings (Table 10). Our work using actigraphic and ECG data supported with RFE and PCA outperformed previous models, achieving an accuracy of 95.83% and F1 score of 0.93 while preserving applicability, unobtrusiveness, and user comfort.

Overall, our findings emphasize tailoring feature sets and employing appropriate reduction techniques for optimal predictive modeling in sleep-related studies. The success of various classifiers and reduction methods provides valuable insights for researchers and practitioners seeking to enhance the accuracy of PSQI score predictions, ultimately contributing to the advancement of sleep quality assessment in clinical and research settings.

To this end, advanced machine learning classifiers have demonstrated high accuracy and reliability in diagnosing insomnia by analyzing physiological and sleep-related data collected from wearable devices. These classifiers can process complex, multi-dimensional data, providing robust and scalable real-time sleep disorder detection and monitoring solutions. This approach offers a promising method for continuous and non-invasive insomnia diagnosis [40], [48].

B. RESTRICTIONS AND OPPORTUNITIES

However, several concerns must be addressed to ensure the reliability and user-friendliness of these solutions. Data privacy is a significant issue, as sensitive health information is transmitted and stored. Additionally, wearable devices' battery life and sensor accuracy must be improved to provide consistent and long-term monitoring without frequent interruptions or errors. Ensuring data security and employing energy-efficient algorithms are critical for the widespread adoption of wearable sleep monitoring technology [49].

Despite these concerns, the integration of wearable device data with Cognitive Behavioral Therapy for Insomnia (CBT-I) presents a highly promising avenue for personalized treatment. By providing objective, real-time insights into sleep patterns and behaviors, this integration can support more accurate assessments, monitor therapy progress, and potentially improve patient outcomes by tailoring interventions based on precise, continuous data [50].

C. FUTURE RESEARCH AND DIRECTION

Future research should address the challenges associated with wearable technology for sleep monitoring. These include improving the accuracy and longevity of sensors, developing robust data security measures, and creating energy-efficient machine learning algorithms [51]. Additionally, further work is needed to explore how to integrate these technologies with CBT-I, ensuring that the combination of objective data and therapeutic interventions maximizes patient benefit.

Further work on insomnia classification should include using a more extensive and balanced dataset encompassing individuals from diverse age groups, genders, and ethnicities, as the MMASH dataset primarily includes data from young male adults. Implementing these algorithms in wearables and smartwatches could give users a comprehensive assessment of their sleep habits. These classification techniques, combined with digital CBT-I, can assist health professionals in more accurately diagnosing their patients' health and evaluating the efficacy of current treatment plans.

V. CONCLUSION

We investigated diverse feature sets and classification models to predict insomnia based on PSQI scores, leveraging sleep variables and heart data features. Our findings highlighted MLP as a robust classifier, consistently delivering high accuracy and F1 scores across different feature sets and reduction techniques. Combining RFE and PCA proved a powerful feature reduction strategy, significantly enhancing classifier performance. The hybrid approach of incorporating both RFE and PCA on sleep and heart rate features yielded outstanding results, with the MLP achieving an accuracy of 95.83% and an F1 score of 0.93 compared to the existing reported accuracy of 92% and an F1 score of 0.92.

Our exploration of feature reduction using NNs demonstrated the potential of learned representations in capturing meaningful information related to PSQI scores, surpassing traditional feature performance. Additionally, classifiers such as LR, KNN, SVM, and MLP showed significant improvements and matched the performance of tree-based models through careful feature selection and dimensionality reduction.

ACKNOWLEDGMENT

The authors acknowledge no conflict of interest.

REFERENCES

[1] Mostafa Haghi, Natividad Martínez Madrid, and Ralf Seepold. In-home, smart sleep monitoring system for cardiorespiratory estimation and sleep apnea detection: proof of concept. *IEEE sensors journal*, 2024.

[2] Jean-Philippe Chaput, Andrew W McHill, Rebecca C Cox, Josiane L Broussard, Caroline Dutil, Bruno GG da Costa, Hugues Sampasa-Kanyinga, and Kenneth P Wright Jr. The role of insufficient sleep and circadian misalignment in obesity. *Nature Reviews Endocrinology*, 19(2):82–97, 2023.

TABLE 10. Comparison of Insomnia Detection Studies with the Proposed Work. Note: KK, Temp., body P., and env. light Stand for Kohen Kappa, Temperature, body position, and environmental light

Study	Data	Model	Accuracy(%)	Sensitivity(%)	F1 Score
This work	Actigraphic + ECG	MLP	95.83	90	0.93
[42]	Actigraphic	SVM	81	77	0.78
[43]	Actigraphic	RF	80	76	NA
[44]	Actigraphic	RF	84	NA	NA
[45]	Temp., motor activity, body p., and env. light	DT	88	88	0.92
[46]	EEG 2-channel	DNN	92	NA	0.84 (KK)
[47]	EEG + ECG	FNN	81.3	NA	NA

[3] Vibha Madan Jha. The prevalence of sleep loss and sleep disorders in young and old adults. *Aging brain*, 3:100057, 2023.

[4] Colin A Espie. Insomnia: conceptual issues in the development, persistence, and treatment of sleep disorder in adults. *Annual review of psychology*, 53(1):215–243, 2002.

[5] Dieter Riemann, Lukas B Krone, Katharina Wulff, and Christoph Nissen. Sleep, insomnia, and depression. *Neuropsychopharmacology*, 45(1):74–89, 2020.

[6] Natalie A Grima, Bei Bei, and Darren Mansfield. Insomnia theory and assessment. *Australian journal of general practice*, 48(4):193–197, 2019.

[7] Dieter Riemann, Fee Benz, Raphael J Dressle, Colin A Espie, Anna F Johann, Tessa F Blanken, Jeanne Leerssen, Rick Wassing, Alasdair L Henry, Simon D Kyle, et al. Insomnia disorder: State of the science and challenges for the future. *Journal of sleep research*, 31(4):e13604, 2022.

[8] Benedetto Farina, Serena Dittoni, Salvatore Colicchio, Elisa Testani, Anna Losurdo, Valentina Gnoni, Chiara Di Blasi, Riccardo Brunetti, Anna Contardi, Salvatore Mazza, et al. Heart rate and heart rate variability modification in chronic insomnia patients. *Behavioral sleep medicine*, 12(4):290–306, 2014.

[9] Rollin McCraty and Fred Shaffer. Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. *Global advances in health and medicine*, 4(1):46–61, 2015.

[10] Denise C Jarrin, Hans Ivers, Manon Lamy, Ivy Y Chen, Allison G Harvey, and Charles M Morin. Cardiovascular autonomic dysfunction in insomnia patients with objective short sleep duration. *Journal of sleep research*, 27(3):e12663, 2018.

[11] Roger J Cole, Daniel F Kripke, William Gruen, Daniel J Mullaney, and J Christian Gillin. Automatic sleep/wake identification from wrist activity. *Sleep*, 15(5):461–469, 1992.

[12] Mirja Quante, Emily R Kaplan, Michael Cailler, Michael Rueschman, Rui Wang, Jia Weng, Elsie M Taveras, and Susan Redline. Actigraphy-based sleep estimation in adolescents and adults: a comparison with polysomnography using two scoring algorithms. *Nature and science of sleep*, pages 13–20, 2018.

[13] Sangwon Byun, Ah Young Kim, Eun Hye Jang, Seunghwan Kim, Kwan Woo Choi, Han Young Yu, and Hong Jin Jeon. Detection of major depressive disorder from linear and nonlinear heart rate variability features during mental task protocol. *Computers in biology and medicine*, 112:103381, 2019.

[14] Ke Yan and David Zhang. Feature selection and analysis on correlated gas sensor data with recursive feature elimination. *Sensors and Actuators B: Chemical*, 212:353–363, 2015.

[15] Danni Kuang, Rongqian Yang, Xiuwen Chen, Guohui Lao, Fengchun Wu, Xiong Huang, Ruixue Lv, Lei Zhang, Chuanxu Song, and Shanxing Ou. Depression recognition according to heart rate variability using bayesian networks. *Journal of psychiatric research*, 95:282–287, 2017.

[16] Aoyu Li, Jingwen Li, Dongxu Zhang, Wei Wu, Juanjuan Zhao, and Yan Qiang. Synergy through integration of digital cognitive tests and wearable devices for mild cognitive impairment screening. *Frontiers in Human Neuroscience*, 17:1183457, 2023.

[17] Franca Delmastro, Flavio Di Martino, and Cristina Dolciotti. Cognitive training and stress detection in mci frail older people through wearable sensors and machine learning. *Ieee Access*, 8:65573–65590, 2020.

[18] Alessio Rossi, Eleonora Da Pozzo, Dario Menicagli, Chiara Tremolanti, Corrado Priami, Alina Sirbu, David Clifton, Claudia Martini, and David Morelli. Multilevel monitoring of activity and sleep in healthy people. *PhysioNet*, 2020.

[19] Sylvain Laborde, Emma Mosley, and Julian F Thayer. Heart rate variability and cardiac vagal tone in psychophysiological research–recommendations for experiment planning, data analysis, and data reporting. *Frontiers in psychology*, 8:238557, 2017.

[20] Alexander A Huang and Samuel Y Huang. Use of machine learning to identify risk factors for insomnia. *PloS one*, 18(4):e0282622, 2023.

[21] Emma Holler, Farid Chekani, Jizhou Ai, Weilin Meng, Rezaul Karim Khandker, Zina Ben Miled, Arthur Owora, Paul Dexter, Noll Campbell, Craig Solid, et al. Development and temporal validation of an electronic medical record-based insomnia prediction model using data from a statewide health information exchange. *Journal of Clinical Medicine*, 12(9):3286, 2023.

[22] Giuseppe Curcio, Daniela Tempesta, Simone Scarlata, Cristina Marzano, Fabio Moroni, Paolo Maria Rossini, Michele Ferrara, and Luigi De Gennaro. Validity of the italian version of the pittsburgh sleep quality index (psqi). *Neurological Sciences*, 34:511–519, 2013.

[23] Mohammed Salahuddin, Tarekegn Tesfaye Maru, Abera Kumalo, Seithikurippu R Pandi-Perumal, Ahmed S Bahammam, and Md Dilshad Manzar. Validation of the pittsburgh sleep quality index in community dwelling ethiopian adults. *Health and quality of life outcomes*, 15:1–7, 2017.

[24] Jutta Backhaus, Klaus Junghanns, Andreas Broocks, Dieter Riemann, and Fritz Hohagen. Test–retest reliability and validity of the pittsburgh sleep quality index in primary insomnia. *Journal of psychosomatic research*, 53(3):737–740, 2002.

[25] Shirin Shafazand, Douglas M Wallace, Kristopher L Arheart, Silvia Vargas, Corneliu C Luca, Henry Moore, Heather Katzen, Bonnie Levin, and Carlos Singer. Insomnia, sleep quality, and quality of life in mild to moderate parkinson's disease. *Annals of the American Thoracic Society*, 14(3):412–419, 2017.

[26] Mostafa Shahin, Lamana Mulaffer, Thomas Penzel, and Beena Ahmed. A two stage approach for the automatic detection of insomnia. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 466–469. IEEE, 2018.

[27] Ali Şenol, Tarık Talan, and Cemal Aktürk. A new hybrid feature reduction method by using mcmstclustering algorithm with various feature projection methods: a case study on sleep disorder diagnosis. *Signal, Image and Video Processing*, 18(5):4589–4603, 2024.

[28] RK Tripathy and U Rajendra Acharya. Use of features from rr-time series and eeg signals for automated classification of sleep stages in deep neural network framework. *Biocybernetics and Biomedical Engineering*, 38(4):890–902, 2018.

[29] Thomas Cover and Peter Hart. Nearest neighbor pattern classification. *IEEE transactions on information theory*, 13(1):21–27, 1967.

[30] J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1:81–106, 1986.

[31] Andy Liaw, Matthew Wiener, et al. Classification and regression by randomforest. *R news*, 2(3):18–22, 2002.

[32] Chih-Chung Chang and Chih-Jen Lin. Libsvm: a library for support vector machines. *ACM transactions on intelligent systems and technology (TIST)*, 2(3):1–27, 2011.

[33] David E Rumelhart, Geoffrey E Hinton, and Ronald J Williams. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

[34] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pages 785–794, 2016.

[35] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. Extreme learning machine: theory and applications. *Neurocomputing*, 70(1-3):489–501, 2006.

[36] Olga Vl Bitkina, Jaehyun Park, and Jungyoon Kim. Modeling sleep quality depending on objective actigraphic indicators based on machine learning

methods. *International Journal of Environmental Research and Public Health*, 19(16):9890, 2022.

[37] Nishant Sharma, Manish Sharma, Hardik Telangore, and U Rajendra Acharya. Automated accurate insomnia detection system using wavelet scattering method using ecg signals. *Applied Intelligence*, pages 1–18, 2024.

[38] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

[39] John D Kelleher. *Deep learning*. MIT press, 2019.

[40] Sungkyu Park, Sang Won Lee, Sungwon Han, Meeyoung Cha, et al. Clustering insomnia patterns by data from wearable devices: algorithm development and validation study. *JMIR mHealth and uHealth*, 7(12):e14473, 2019.

[41] Melissa Aji, Nick Glozier, Delwyn J Bartlett, Ronald R Grunstein, Rafael A Calvo, Nathaniel S Marshall, David P White, and Christopher Gordon. The effectiveness of digital insomnia treatment with adjunctive wearable technology: A pilot randomized controlled trial. *Behavioral sleep medicine*, 20(5):570–583, 2022.

[42] Shitanshu Kusmakar, Chandan Karmakar, Ye Zhu, Sergiy Shelyag, Sean PA Drummond, Jason G Ellis, and Maia Angelova. A machine learning model for multi-night actigraphic detection of chronic insomnia: development and validation of a pre-screening tool. *Royal Society open science*, 8(6):202264, 2021.

[43] S Rani, Sergiy Shelyag, Chandan Karmakar, Ye Zhu, R Fossion, JG Ellis, SPA Drummond, and M Angelova. Differentiating acute from chronic insomnia with machine learning from actigraphy time series data. *Frontiers in Network Physiology*, 2:1036832, 2022.

[44] Maia Angelova, Chandan Karmakar, Ye Zhu, Sean PA Drummond, and Jason Ellis. Automated method for detecting acute insomnia using multi-night actigraphy data. *IEEE Access*, 8:74413–74422, 2020.

[45] Beatriz Rodriguez-Morilla, Eduard Estivill, Carla Estivill-Domènech, Javier Albares, Francisco Segarra, Angel Correa, Manuel Campos, Maria Angeles Rol, and Juan Antonio Madrid. Application of machine learning methods to ambulatory circadian monitoring (acm) for discriminating sleep and circadian disorders. *Frontiers in neuroscience*, 13:1318, 2019.

[46] Mostafa Shahin, Beena Ahmed, Sana Tmar-Ben Hamida, Fathima Lamana Mulaffer, Martin Glos, and Thomas Penzel. Deep learning and insomnia: assisting clinicians with their diagnosis. *IEEE journal of biomedical and health informatics*, 21(6):1546–1553, 2017.

[47] H Abdullah, Thomas Penzel, and Dean Cvetkovic. Detection of insomnia from eeg and ecg. In *The 15th International Conference on Biomedical Engineering: ICBME 2013, 4th to 7th December 2013, Singapore*, pages 687–690. Springer, 2014.

[48] Jiaxing Liu, Yang Zhao, Boya Lai, Hailiang Wang, Kwok Leung Tsui, et al. Wearable device heart rate and activity data in an unsupervised approach to personalized sleep monitoring: algorithm validation. *JMIR mHealth and uHealth*, 8(8):e18370, 2020.

[49] Xiao Li, Jessilyn Dunn, Denis Salins, Gao Zhou, Wenyu Zhou, Sophia Miryam Schüssler-Fiorenza Rose, Dalia Perelman, Elizabeth Colbert, Ryan Runge, Shannon Rego, et al. Digital health: tracking physiomes and activity using wearable biosensors reveals useful health-related information. *PLoS biology*, 15(1):e2001402, 2017.

[50] Annemarie I Luik, Tanja van der Zweerde, Annemieke van Straten, and Jaap Lancee. Digital delivery of cognitive behavioral therapy for insomnia. *Current psychiatry reports*, 21:1–8, 2019.

[51] Manal Mohamed, Nourelhuda Mohamed, and Jae Gwan Kim. Advancements in wearable eeg technology for improved home-based sleep monitoring and assessment: A review. *Biosensors*, 13(12):1019, 2023.

AMEYA CHATUR completed his bachelor studies at the Indian Institute of Technology Hyderabad with a major in Biomedical Engineering and a minor in Computer Science. Ameya will be pursuing a master's degree in Biomedical Engineering from Johns Hopkins with a focus on Neuroengineering and Brain-Computer Interfaces. His major interests are sensors and healthcare data analysis through Machine Learning and Statistical techniques as well as Computational Neuroscience and Neuromorphic computing.

MOSTAFA HAGHI received the bachelor's degree in electronics engineering from Urmia University, Urmia, Iran, in 2009, the M.Tech degree (First Class Hons. with Distinction) in embedded systems from Jawaharlal Nehru Technological University, Hyderabad, India, in 2014, and the Ph.D. degree from the Center for Life Science Automation, University of Rostock, Rostock, Germany, in 2019. He is currently a Postdoctoral Researcher with the Computer Architecture Team of the Institute for Computer Engineering at Heidelberg University, Germany. Mostafa is currently the co-chair of the "Accident and Emergency Informatics Working Group" of the International Medical Informatics Association and a member of the Young Professionals Committee of IEEE Engineering in Medicine and Biology Society (EMBS). His major interests are wearable healthcare, hardware/software-embedded systems, bio-electronics, and sleep medicine.

NAGARAJAN GANAPATHY (Member, IEEE) is an Assistant Professor in the Department of Biomedical Engineering at Indian Institute of Technology Hyderabad, India. He completed his postgraduate studies in Computer Science and Engineering and earned his Ph.D. in Biomedical Engineering in 2019. His doctoral research was jointly supervised at the Indian Institute of Technology Madras, India, and TU Braunschweig, Germany. Then, he got postdoctoral training at Peter L. Reichertz for Medical Informatics of TU Braunschweig and Hannover Medical School (PLRI), Germany. Presently, he heads the Biomedical Medical Informatics lab at the Indian Institute of Technology, Hyderabad. Currently, He is a visiting scientist at the PLRI. His main research interests are affective computing systems, Artificial Intelligence, biosignals and facial expression analytics, wearables and sensors, multimodal emotion recognition, mental wellbeing, and behavioral Informatics. He has been involved in various scientific projects. He has published more than 51 papers in referred journals and conferences. He coordinated the FAIR4Health and ZDIN (Future Lab for Healthcare) Horizon 2020 projects. Currently, He is the PI for Private Spaces (IIT Hyderabad), SMARTBEING (Indo-Japan JICA Program), and Pilot wellbeing projects. He has received a DAAD fellowship and the DAAD Research Ambassador award.

NIMA TAHERINEJAD (Member, IEEE) received the Ph.D. degree in ECE from The University of British Columbia (UBC), Vancouver, Canada, in 2015. He is currently a Full Professor at Heidelberg University, Germany, and affiliated with TU Wien, Vienna, Austria. He has published three books, four patents, and more than 90 articles. His research interests include in-memory computing, cyber-physical and embedded systems, SoC, memristive circuits and systems, and health care. He received several awards and scholarships from universities, conferences, and competitions he has attended.

RALF SEEPOLD (Senior Member, IEEE) is Professor with the Chair for Ubiquitous Computing at HTWG Konstanz (Germany), Visiting Professor at the School of Computer Engineering at the University of Seville (Spain) and the Southern Taiwan University of Science and Technology (Taiwan). Ralf was a Professor of the Institute of Digital Medicine at Sechenov First Moscow State Medical University (Russia). Ralf obtained his MSc in Computer Science from the University of Paderborn (Germany) and his Ph.D. from the University of Tübingen (Germany). He was Department Manager at the Computer Science Research Center (FZI) at the Karlsruher Institut für Technologie (KIT) in Karlsruhe (Germany), and he was appointed Professor for Telecommunication Engineering and Telematics at the University Carlos III of Madrid (Spain). Ralf is Director of the Ubiquitous Computing Laboratory and the Interdisciplinary Sleep Lab at HTWG, and he leads the research initiative "ReSiliente Systeme (ReSiS) Strengthening resilience in complex systems" at HTWG. His research fields include noninvasive physiological measurement, medical IoT, sleep medicine, ubiquitous AI, and the aging society. Ralf is an IEEE Senior Member, a member of the IEEE Engineering in Medicine and Biology Society, a Member of the Working Group Telemedicine in the German Sleep Society (DGSM), and the European Sleep Research Society (ESRS). He served in IEEE as General Chair, Program Chair, Local Organizing Committee member and Reviewer for several conferences and Journals.

NATIVIDAD MARTÍNEZ MADRID (Senior Member, IEEE) is Professor at the School of Informatics, Reutlingen University (Germany) and appointed Visiting Professor at the Escuela Técnica Superior de Ingeniería Informática (ETSII) in the University of Seville (Spain) and at Southern Taiwan University of Science and Technology in Tainan (Taiwan). Natividad is head of the IoT Lab and Director of the AAL-Living Lab in Reutlingen. She was a Professor at the Sechenov Medical State University in Moscow (Russia) for Information and Internet Technology. Nati obtained her Master and Ph.D. in Telecommunication from the Universidad Politécnica de Madrid (Spain), and her habilitation degree at the Universidad Carlos III de Madrid (Spain), where she worked as Associate Professor. Nati is an IEEE Senior Member, a member of the IEEE Engineering in Medicine and Biology Society, a Founding member and Leader of the Workgroup Innovation of the Smart-Home-Living Society B.W., and a member of several professional organizations, like the German Sleep Society (DGSM), and the European Sleep Research Society (ESRS). Her research focuses on artificial intelligence in healthcare, especially in sleep medicine, care systems, IoT, and wearable medical devices.