Nature and Science of Sleep

Elevate Series

Open Access Full Text Article

METHOD

Research on Sleep Staging Based on Support Vector Machine and Extreme Gradient Boosting Algorithm

Yiwen Wang[1], Shuming Ye[2], Zhi Xu[3], Yonghua Chu[1], Jiarong Zhang[4], Wenke Yu[5]

[1]Clinical Medical Engineering Department, The Second Affiliated Hospital, Zhejiang University School of Medicine, HangZhou, ZheJiang, People's Republic of China; [2]Department of Biomedical Engineering, Zhejiang University, HangZhou, ZheJiang, People's Republic of China; [3]China Astronaut Research and Training Center, BeiJing, People's Republic of China; [4]Baidu Inc, BeiJing, People's Republic of China; [5]Radiology Department, ZheJiang Province Qing Chun Hospital, HangZhou, ZheJiang, People's Republic of China

Correspondence: Yiwen Wang; Shuming Ye, Email karenkaren2010@zju.edu.cn; ysmln@vip.sina.com

Nature and Science of Sleep downloaded from https://www.dovepress.com/
For personal use only.

Purpose: To develop a sleep-staging algorithm based on support vector machine (SVM) and extreme gradient boosting model (XB Boost) and evaluate its performance.

Methods: In this study, data features were extracted based on physiological significance, feature dimension reduction was performed through appropriate methods, and XG Boost classifier and SVM were used for classification. One hundred and twenty training sets and 80 test sets were randomly composed of the first 200 groups of data from the SHH1 database. The polysomnography (PSG) data of 20 real individuals in the clinic were selected as the experimental data. The C3 electroencephalogram (EEG), left and right electro-oculogram (EOG), electromyogram (EMG), and other signals were analyzed. Finally, the stages were adjusted based on human sleep laws. The standard staging of the database and the doctor's diagnosis staging was used as the standard.

Results: The SHHS1 database test results were as follows: the average accuracy was 83.24%, the precision and recall of Stage Wake and Stage 2 NREM sleep (N2) were over 80%, and the precision, F1-Score and recall of Stage 3 NREM sleep (N3) and Rapid Eye Movement (REM) were more than 70%. The clinical data test results were as follows: the average accuracy rate was 76.37%; for Wake and N3, the precision reached 85%; for Wake, N2, and REM, the recall rate reached over 70%; for Wake, the F-1 Score reached over 90%.

Conclusion: This study shows that the sleep staging results of the algorithm for the database and clinical data were similar. The staging results meet the requirements at the medical level.

Keywords: sleep staging, physiological significance, feature dimension reduction, databases and clinical trials, confusion matrix

Introduction

In 2007, the American Academy of Sleep Medicine (AASM) developed new sleep interpretation guidelines - The AASM Guidelines. The AASM guidelines define sleep stages as Wake, non-rapid eye movement (NREM), and rapid eye movement (REM). NREM is divided into NREMI stage (N1), NREMII stage (N2) and NREMIII stage (N3). When falling asleep, the sleep stage of normal people first enters the NREM stage, gradually shifts from N1 stage to N3 stage, and then enters the REM stage at N2 stage or N3 stage. The sleep stage passes from NREM to REM, representing a complete sleep cycle[1–5].

Sleep staging results can be linked to certain diseases. Rapid eye movement sleep behavior disorder (RBD), as an early biological marker for predicting neurodegenerative diseases, such as Parkinson's disease, is of great significance for the prevention and treatment of the latent stage of neurodegenerative diseases. RBD diagnosis requires accurate REM staging.[2] Depressed patients had low sleep efficiency and poor sleep quality. Accurate and detailed sleep staging results are required for both diagnosis and pharmacological intervention, particularly N3.[3,4] The proportion of REM sleep time is reduced in adolescents with bipolar depressive disorder (BDD). The duration of N2 was shorter in adult BDD patients than in normal adults.[5] Although the diagnosis of sleep apnea syndrome (SAS) requires respiratory indications, and blood oxygen indications, the results of sleep staging also play an important role in the determination and research of SAS of the central and obstructive types.[6]

Clinically, polysomnography (PSG) is generally used for sleep-stage classification in patients.[7] It takes a doctor two hours to determine a patient's sleep stage using PSG (particularly electroencephalogram (EEG)). The workload is large and it depends on the experience of the doctor. Patients are required to wear EEG electrodes overnight. The measurement method is inconvenient and subject to greater subjective influence.[8,9] The lack of medical resources and low diagnostic efficiency have

Received: 13 July 2024
Accepted: 15 October 2024
Published: 26 November 2024

© 2024 Wang et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (http://creativecommons.org/licenses/by-nc/3.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

resulted in some shortcomings in professional diagnosis of related diseases and insufficient popularization of sleep quality preliminary screening. Therefore, an efficient sleep staging diagnostic method is needed to assist clinical decision making.

Many researchers attempt to develop automatic sleep staging methods. The main signal of the automatic sleep staging algorithm is EEG. Since the 1980s, artificial intelligence has been rapidly developing and has become a research hot spot. There have been numerous studies on the combination of machine learning and sleep staging in the past decade. In 2010, Güneş et al used a novel algorithm known as k-means clustering-based feature weighting (KMCFW) to increase the accuracy from 55% to 82%.[10] A study by Silveira in 2017 had an accuracy higher than 90%. The method consisted of decomposing EEGs using a discrete wavelet transform and computing the kurtosis, skewness, and variance of its coefficients at selected levels. A random forest predictor was trained to classify each epoch into one of the R or K' stages.[11] Seifpour designed an accurate and robust computer-assisted sleep stage scoring system using a single-channel EEG signal by proposing a novel time-domain feature known as statistical behavior of local extrema (SBLE), which extracted and defined various patterns in 2018. The average accuracy rate was 97.9% for the six-stage to two-stage sleep classification of the Sleep-EDF dataset.[12] Fifty-five time and frequency-domain features were extracted from the EEG signals and were then constituted as the inputs to the LSTM networks. The average accuracy rate was 83.6% on the Sleep-EDF dataset.[13] Fang and Tao of South China Normal University achieved an accuracy of 88% based on the energy characteristics of EEG signals and the regular changes in fuzzy entropy in different sleep stages, combined with support vector machine classification.[14] Wenbing et al extracted the relative energy mean of the rhythm wave using the wavelet transform algorithm and the multi-scale entropy of 9 to 13 and input it into the feedback propagation neural network classifier. Through the statistical analysis of the experimental results, the average staging accuracy of the proposed method was obtained as 85.81%.[15]

In this study, 120 training sets and 80 test sets were randomly composed of the first 200 groups of data from the National Sleep Research Resource Database of the United States. Two types of classifiers were used in the database. Then, the superior performer was selected, optimized, and used on the clinical data. The accuracy, precision, recall, and F1 score were used to evaluate the algorithm.

Materials

Shhs1

The SHHS1 database[16] contains considerable sleep data of healthy subjects. The Sleep Heart Health Study (SHHS) conducted by the National Heart, Lung, and Blood Institute shares the physiological signal data collected during the clinical sleep trials in the National Sleep Research Repository. Among them, the PSG experimental samples in the SHHS1 dataset were derived from 6441 individuals collected from 1995 to 1998.[17] The subjects were all aged 40 years or older, had no sleep-disordered breathing syndrome, and had not undergone tracheostomy and home oxygen therapy. Each sample contains C3-A2 and C4-A1 EEG, left and right eye, EMG, and ECG signals, and each sample contains experts' sleep staging results in a 30-second frame. There are six movement categories. In this study, the first 200 data groups in the SHHS1 dataset were indiscriminately selected as experimental data, and the staging results annotated by experts were used as the standard staging results. The EEG adopts the third lead C3-A2 signal, the EMG adopts the fifth lead EMG signal, and the EOG adopts the sixth lead EOGL signal.

Clinical Data

This study complies with the provisions of the Helsinki Declaration and has obtained the written informed consent of the patients. The study was approved by the Human Research Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine. The batch number is (2021) LunShenYanDi (0053). The PSG instrument was used to collect 51 real individuals, aged 18–65 years, without sleep respiratory syndrome and without tracheostomy or home oxygen therapy. Each measurement should take no less than 8 hours. The subjects had no sleep-related diseases and had a good schedule. In the experiment, the environment should be quiet, no noise, no light stimulation and the temperature should be moderate. C3 EEG, left and right eye EEG, and EMG were collected, and the sleep results

were recorded at one frame every 10 seconds. The results were calculated at one frame in 30 seconds by integrating algorithms, and the staging results were evaluated by doctors. Finally, 20 data that meet the requirements were sorted out.

Methods

The algorithm of this project included signal preconditioning, feature extraction and screening, classifier design, stage adjustment, and other steps.

Signal Processing

Preconditioning

ECG, EEG, EMG, and EOG signals are weak physiological signals with low amplitude and are easily disturbed. The noise was first removed. The EMG and EOG signals have distinct time-domain characteristics, so the baseline and poles were first removed; then, a zero-stage filter was used for filtering. The effective frequency band of eye electricity was mainly concentrated in the 0–10 hz range, and the passband frequency of the EOG filter was selected as 0.5–10 hz. The effective frequency band of EMG was 2–500 hz, and the passband frequency of the EMG filter was selected as 5–50 hz. The passband frequency of the EEG filter was selected as 0.3–50 Hz.[18]

Feature Extraction and Screening

The EMG amplitude was the highest in the wake stage and the lowest in the REM stage. The amplitude of EOG was larger in the wake stage and REM stage. Therefore, the relationship between EOG, EMG and sleep phases is mainly reflected in the time domain. Different sleep stages have different brain electrical rhythms, including α rhythm, slow wave activity during sleep, apex wave, sleep spindle wave and K complex wave. The difference of EEG rhythm is mainly reflected in the frequency domain, so the relationship between EEG and sleep stages is mainly reflected in the frequency domain.

For the non-stationary signal with weak amplitude, there are many time domain characteristics, such as integral mean, variance and coefficient of variation. Among them, the mean can express the signal strength, the variance can express the signal activity, and the coefficient of variation can balance the mean and variance and unify the dimension. In the wake stage and REM stage, the integral mean value and variance amplitude of EOG are significantly larger than those in other sleep stages, which is consistent with the physiological characteristics of the more active EOG in the wake stage and REM stage mentioned above. Therefore, the time-domain characteristics proposed above are of great significance for distinguishing the wake stage, REM stage and other sleep stages. The variance of EMG is the largest in the wake stage, and with the deepening of sleep, the characteristic value gradually decreases and approaches zero in the rapid eye movement stage, which accords with the weakest characteristic of EMG in the rapid eye movement stage analyzed above. Wake and REM stages can be distinguished according to the amplitude of the EMG features. According to the AASM standard, artificial staging mainly depends on the distribution of different EEG rhythms to judge the sleep stage, and EEG rhythms are signals located in different frequency bands. Therefore, frequency domain features of EEG signals can be extracted to analyze the dominant EEG rhythms in different stages, so as to determine the sleep stage according to the distribution of EEG rhythms. In this paper, spectral relative energy, spectral energy ratio of different EEG rhythms and relative power of different frequency bands are used to describe the frequency domain of EEG signals.

The quantity of the feature parameters and the correlation between the feature parameters and classification results directly affect the classification results. It is required to select several features that contribute the most to the classification. First, the features are standardized. Then, the *T*-test, variance filtering, mutual information, embedding, recursive feature elimination, and other methods were used to select the features. For feature dimension reduction, filtering method is used to remove some irrelevant features, embedding method is used to rank the importance of each feature, packaging method is used to select the least and best feature collection under the premise of no decline in accuracy, and PCA method is used for dimension reduction. Finally, 20 features are selected. The stratified sampling K-fold cross-validation method was used to verify the performance of the classifier. Thus, simplified features with a large amount of information and weak cross-correlation were obtained.

Support Vector Machine (SVM) Algorithm and Extreme Gradient Algorithm

Svm

The core purpose of the SVM classifier is to separate two or more samples by constructing a split surface to determine the separating hyperplane, which correctly divides the training dataset and has the largest geometric separation. SVM is suitable for solving small-sample, high-dimensional, and nonlinear problems, and can improve generalization performance.[19,20]

First, the eigenvalues were uniformized to improve the model accuracy and convergence speed. Then, a kernel function and different parameters were selected to set the model. Different parameters affect the correctness and accuracy of classification and tolerance to errors. Finally, several parameters were selected for testing to determine the optimal parameters.

When using the SVM algorithm, the data was manually divided into segments of 10s for algorithm processing. If a staging result violates the normal sleep physiological situation, it can be corrected by referring to "Clinical Electroencephalography".[21,22]

Extreme Gradient Boosting Algorithm

The extreme gradient boosting algorithm (XG Boost) is a variant of the gradient-boosting machine. Its purpose is to breakthrough the computing power limit of the gradient boosting tree algorithm and improve the performance and computing speed of the model. The algorithm was easy to parallelize, had high accuracy, and supported regression and classification. Additionally, it is not sensitive to the dimension, unit of the feature, and missing values; thus, it is not required to standardize the feature.[23,24]

The XG Boost algorithm also requires setting and adjusting parameters. This project aims to use the grid search and K-fold cross-validation methods to optimize the parameters. Grid search is used to traverse all the parameters to determine the best groups: num_round = 300, eta = 0.275, max_depth = 6, γ=15, colsample=1, $\lambda = 10$, $\alpha = 0$. The K-fold cross-validation method is used to train and test the complementary subsets, which were obtained by dividing the sample set multiple times. Its purpose was to avoid the overfitting caused by reusing a part of the data in the training model.[25,26]

The results were also required to be logically adjusted according to the physiological situation and regularities of the sleep cycle.

The convergence speed of the SVM and the accuracy of the model is closely connected with standardization. If the data is not normalized, it will take several hours to run the algorithm. Therefore, eigenvalue standardization must be performed for the support vector machine. On the contrary, for XG Boost, if the features are standardized, some data information will be lost, which affects the model accuracy.

SVM and XG Boost were applied to the database. Then, 40 sets of data were selected indiscriminately in SHHI. As shown in Table 1, the average accuracy of the XG Boost algorithm is 76.77%, and the average accuracy of the SVM algorithm is 63.37%. The SVM algorithm requires approximately 20s to predict a set of data, whereas the XG Boost algorithm requires less than 1 s. In terms of accuracy and time, XG Boost was superior.

Table 1 The Accuracy of the First 40 Test Sets on Different Algorithms

Symbol	XG Boost	SVM	Symbol	XG Boost	SVM
1–200121	78.25%	72.90%	1–200141	74.84%	62.46%
1–200122	82.91%	79.57%	1–200142	80.68%	66.44%
1–200123	84.97%	77.26%	1–200143	74.21%	60.37%
1–200124	84.16%	68.31%	1–200144	82.31%	70.56%
1–200125	81.03%	77.13%	1–200145	79.30%	60.18%
1–200126	85.54%	54.36%	1–200146	82.97%	57.66%
1–200127	69.18%	45.14%	1–200147	83.01%	75.18%
1–200128	83.57%	72.75%	1–200148	43.27%	59.48%
1–200129	76.54%	54.20%	1–200149	85.72%	71.25%

(*Continued*)

Table 1 (Continued).

Symbol	XG Boost	SVM	Symbol	XG Boost	SVM
1–200130	70.47%	57.44%	1–200150	68.10%	60.56%
1–200131	81.52%	71.05%	1–200151	73.34%	62.92%
1–200132	77.49%	68.57%	1–200152	80.69%	59.76%
1–200133	83.22%	65.92%	1–200153	89.80%	34.92%
1–200134	81.54%	74.09%	1–200154	73.89%	61.09%
1–200135	67.49%	55.68%	1–200155	75.52%	60.87%
1–200136	70.80%	67.70%	1–200156	78.52%	76.98%
1–200137	80.92%	58.99%	1–200157	66.83%	61.82%
1–200138	61.19%	46.65%	1–200158	73.34%	51.06%
1–200139	78.10%	68.20%	1–200159	66.29%	58.37%
1–200140	78.20%	60.83%	1–200160	78.24%	66.10%

Deeper Insights from Classification Models

Algorithm Optimization

For the XG Boost algorithm, the generalization error of the model can generally be evaluated and adjusted through experimental tests. For the XG Boost algorithm, the generalization error of the model can generally be evaluated and adjusted through experimental tests. Owing to the problem of sample imbalance in sleep staging, stratified sampling was used to ensure that the proportion of sample categories was balanced. This study used the 5-fold cross-validation method.

Sample from SHHS1 sizes varies with sleep stage. The Wake stage accounted for 31.59%, N1 stage accounted for 3.24%, N2 stage accounted for 39.69%, N3 stage accounted for 13.27%, and REM stage accounted for 12.21%.

If the problem of sample imbalance is ignored, poor specificity and sensitivity of this category, which contains fewer samples, occur. XG Boost algorithm can set the weight of each sample to solve the problem of sample imbalance. First, the weight of each category is directly set as the reciprocal of the proportion of samples in this category, so that the weight of samples in each category can be regarded as equal. Then, an exhaustive test was carried out near this set of parameter combinations, and the weights of N1, N3 and REM stages were traversed with a certain step length, and the average accuracy of stratified sampling 5-fold cross-validation was used as an evaluation index to select the parameter model with the best score. The weight of wake is finally determined to be 3.2, N1 is 16.0, N2 is 2.5, N3 is 7.0, and REM is 8.0. After the weight setting, the recall rate of stage N1 increased 23%, and the F1 score has increased 20%.

Correction of Sleep Staging Results

The staging results obtained using the XG Boost algorithm are only the results of machine learning, often violating the physiological characteristics. Therefore, three characteristics were followed through the regularity of the sleep cycle to revise the results of sleep staging in this study:

The initial state of sleep data is generally the Wake stage, and the duration will be no less than 1 min;

The REM stage is generally transformed from the N2 and N3 stages, and then, the REM stage is generally transformed into the N2 and N3 stages;

Generally, from the N1 and N2 stages of the two light sleep stages instead of the N3 and REM stages into the Wake stage.

Conclusion and Discussion

Table 2 is the confusion matrix. The sum of each row of the confusion matrix is the actual number of samples of that category, and the sum of each column is the predicted number of samples of that category. Nij is the number of samples that are actually category i but predicted to be category j. i and j are 0, 1, 2, 3, 5, which correspond to the five sleep stages of Wake, N1, N2, N3, and REM, respectively.

Table 2 Confusion Matrix

		Predict				
		Wake	**NREM1**	**NREM2**	**NREM3**	**REM**
True	Wake	N_{00}	N_{01}	N_{02}	N_{03}	N_{05}
	NREM1	N_{10}	N_{11}	N_{12}	N_{13}	N_{15}
	NREM2	N_{20}	N_{21}	N_{22}	N_{23}	N_{25}
	NREM3	N_{30}	N_{31}	N_{32}	N_{33}	N_{35}
	REM	N_{50}	N_{51}	N_{52}	N_{53}	N_{55}

This study used precision, recall, accuracy, and F1 score to comprehensively evaluate the model.

$$Pre_i = \frac{N_{ii}}{\sum_{j\in(0,1,2,3,5)} N_{ji}} * 100\%$$

$$Rec_i = \frac{N_{ii}}{\sum_{j\in(0,1,2,3,5)} N_{ij}} * 100\%$$

$$F - Score_i = \left(1+\beta^2\right) * \frac{Pre_i * Rec_i}{\beta^2 * Pre_i + Rec_i}$$

$$Acc = \frac{\sum_{i\in(0,1,2,3,5)} N_{ii}}{\sum_{i\in(0,1,2,3,5)} * \left(\sum_{j\in(0,1,2,3,5)} N_{ij}\right)} * 100\%$$

Tables 3–6 lists the result that the algorithm applied to SHHS1 test samples. Tables 7–11 list the result that the algorithm applied to clinical data.

Table 3 The Precision of 80 Sets of Test Data

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200121	97.42	57.14	87.50	85.25	96.64
shhs-200122	96.84	66.67	98.41	93.33	73.36
shhs-200123	99.35	35.29	95.12	86.73	96.08
shhs-200124	99.56	13.64	90.71	60.00	59.53
shhs-200125	97.60	60.00	92.29	98.66	68.18
shhs-200126	99.29	29.41	77.97	0.00	0.00
shhs-200127	59.61	13.64	93.84	28.57	81.71
shhs-200128	97.57	57.14	96.30	84.21	76.62
shhs-200129	87.38	18.46	87.46	82.58	89.83
shhs-200130	64.23	27.06	97.89	0.00	63.70
shhs-200131	95.51	25.00	87.20	98.17	97.03
shhs-200132	98.75	20.00	80.42	86.73	96.99
shhs-200133	90.70	70.00	95.13	93.75	85.96

(*Continued*)

Table 3 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200134	97.78	18.75	93.71	67.21	78.95
shhs-200135	83.33	30.77	78.30	94.59	93.88
shhs-200136	94.55	25.00	76.36	96.30	72.97
shhs-200137	97.67	50.00	88.89	82.81	86.08
shhs-200138	64.95	1.37	75.79	83.33	75.51
shhs-200139	95.81	30.56	82.22	93.06	90.57
shhs-200140	97.13	20.00	89.31	100.00	68.15
shhs-200141	82.54	0.00	92.54	88.17	90.50
shhs-200142	95.27	40.00	86.78	91.85	97.13
shhs-200143	97.58	0.00	66.86	97.03	100.00
shhs-200144	82.25	0.00	98.88	93.98	85.55
shhs-200145	75.85	0.00	97.81	0.00	0.00
shhs-200146	99.60	42.11	96.07	97.24	91.47
shhs-200147	100.00	50.00	95.78	83.09	78.40
shhs-200148	72.25	0.00	79.83	10.00	80.30
shhs-200149	97.59	62.50	97.11	91.04	85.59
shhs-200150	96.04	0.00	82.35	83.33	57.27
shhs-200151	99.16	15.38	65.26	22.08	86.67
shhs-200152	84.22	71.43	96.11	80.88	66.67
shhs-200153	100.00	8.00	85.66	41.18	96.55
shhs-200154	96.69	80.00	86.30	97.85	70.89
shhs-200155	93.18	21.15	84.77	81.69	93.98
shhs-200156	96.73	0.00	76.34	82.04	95.35
shhs-200157	98.76	6.90	46.90	87.88	82.81
shhs-200158	93.51	37.50	80.28	88.76	71.43
shhs-200159	97.47	27.27	58.10	100.00	65.63
shhs-200160	97.75	75.00	85.28	100.00	81.54
shhs-200161	95.37	0.00	73.31	98.25	92.31
shhs-200162	97.13	83.33	95.02	93.68	81.44
shhs-200163	98.81	27.27	86.57	88.89	86.73
shhs-200164	98.65	83.33	90.69	71.07	56.05
shhs-200165	99.01	60.00	91.89	93.46	57.77
shhs-200166	33.50	0.00	82.49	99.36	74.33

(*Continued*)

Table 3 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200167	90.77	38.10	83.87	71.08	81.56
shhs-200168	98.14	46.67	91.52	80.22	78.31
shhs-200169	87.70	34.78	89.58	80.00	86.61
shhs-200170	75.56	50.00	88.25	81.22	90.00
shhs-200171	96.49	33.33	90.91	98.06	80.30
shhs-200172	89.60	5.71	74.51	95.12	95.45
shhs-200173	86.05	23.08	55.02	97.96	100.00
shhs-200174	96.64	63.64	91.61	92.20	99.44
shhs-200175	97.17	40.00	93.01	100.00	53.64
shhs-200176	50.00	50.00	85.67	96.36	24.56
shhs-200177	74.87	17.19	97.64	27.50	76.88
shhs-200178	85.53	18.75	83.96	96.13	65.46
shhs-200179	98.29	61.54	82.81	89.22	87.50
shhs-200180	95.63	100.00	97.97	45.83	68.31
shhs-200181	92.23	18.75	88.14	0.00	76.85
shhs-200182	100.00	22.22	46.58	93.75	90.91
shhs-200183	98.15	66.67	67.09	0.00	46.15
shhs-200184	96.40	33.33	74.17	74.30	86.05
shhs-200185	49.60	5.88	61.70	50.00	100.00
shhs-200186	84.75	60.87	94.39	37.14	66.92
shhs-200187	91.24	51.28	96.90	53.13	63.24
shhs-200188	87.16	0.00	91.74	6.06	66.04
shhs-200189	99.22	0.00	75.26	97.41	80.58
shhs-200190	99.62	38.10	67.58	88.57	80.43
shhs-200191	47.24	47.22	71.28	96.77	68.64
shhs-200192	78.81	32.14	94.04	85.71	77.40
shhs-200193	98.68	42.31	69.03	89.58	94.20
shhs-200194	99.10	18.18	78.12	80.30	82.52
shhs-200195	83.74	0.00	87.84	96.36	0.00
shhs-200196	99.64	71.43	80.87	75.51	85.19
shhs-200197	75.51	66.67	73.43	79.03	74.34
shhs-200198	90.99	5.41	48.78	100.00	0.00
shhs-200199	97.27	16.67	96.94	82.22	76.17

(*Continued*)

Table 3 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200200	99.68	100.00	91.13	78.62	96.88
Average Precision	89.82	34.26	83.81	76.17	75.98

Table 4 The Accuracy of 80 Sets of Test Data

Sample	Accuracy(%)	Sample	Accuracy(%)	Sample	Accuracy(%)
shhs-200121	90.08	shhs-200122	91.6	shhs-200123	93.81
shhs-200124	85.2	shhs-200125	91.46	shhs-200126	88.77
shhs-200127	73.87	shhs-200128	91.58	shhs-200129	83.04
shhs-200130	74.46	shhs-200131	92.02	shhs-200132	86.7
shhs-200133	90.65	shhs-200134	90.08	shhs-200135	79.58
shhs-200136	78.68	shhs-200137	89.57	shhs-200138	67.06
shhs-200139	87.9	shhs-200140	88.57	shhs-200141	90.9
shhs-200142	90.94	shhs-200143	82.6	shhs-200144	92.32
shhs-200145	68.76	shhs-200146	95.23	shhs-200147	92.63
shhs-200148	73.79	shhs-200149	93.83	shhs-200150	78.58
shhs-200151	79.14	shhs-200152	87.25	shhs-200153	93.55
shhs-200154	85.51	shhs-200155	84.75	shhs-200156	85.96
shhs-200157	70.55	shhs-200158	81.84	shhs-200159	77.59
shhs-200160	94.29	shhs-200161	81.2	shhs-200162	92.7
shhs-200163	85.49	shhs-200164	76.84	shhs-200165	83.79
shhs-200166	73.9	shhs-200167	80.88	shhs-200168	91.56
shhs-200169	86.78	shhs-200170	84.47	shhs-200171	90.77
shhs-200172	79.51	shhs-200173	70.63	shhs-200174	93.83
shhs-200175	78.58	shhs-200176	56.97	shhs-200177	69.11
shhs-200178	79.6	shhs-200179	88.12	shhs-200180	89.12
shhs-200181	85.41	shhs-200182	67.26	shhs-200183	85.82
shhs-200184	81.09	shhs-200185	51.73	shhs-200186	86.25
shhs-200187	89.17	shhs-200188	84.36	shhs-200189	89.17
shhs-200190	84.85	shhs-200191	60.86	shhs-200192	85.93
shhs-200193	83.26	shhs-200194	82.23	shhs-200195	86.06
shhs-200196	85.62	shhs-200197	75.39	shhs-200198	63.1
shhs-200199	90.11	shhs-200200	92.65	Average Accuracy	83.23

Table 5 The F1-Score of 80 Sets of Test Data

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200121	95.7	11.59	91.93	91.23	94.43
shhs-200122	93.87	16.67	95.02	90.91	84.63
shhs-200123	96.67	29.27	94.57	92.89	90.74
shhs-200124	96	14.63	84.23	31.58	71.83
shhs-200125	97.27	33.33	95.99	77.98	81.08
shhs-200126	92.54	37.04	87.29	0	0
shhs-200127	73.16	16.51	78.79	44.44	84.28
shhs-200128	97.83	34.41	93.91	91.43	85.79
shhs-200129	73.77	25.81	87.24	90.46	82.6
shhs-200130	78.22	30.67	78.2	0	75.85
shhs-200131	96.08	16.67	93.16	89.17	87.5
shhs-200132	87.45	19.05	89.15	72.96	86
shhs-200133	95.01	40.38	94.18	72.29	92.24
shhs-200134	95.52	22.22	92.11	71.3	84.11
shhs-200135	84.75	5.88	87.83	58.33	56.44
shhs-200136	95.63	30	84.31	23.74	84.38
shhs-200137	89.55	18.18	94.12	84.8	83.44
shhs-200138	78.75	2.25	80.6	12.2	27.01
shhs-200139	94.49	29.33	84.92	95.37	84.96
shhs-200140	91.6	11.11	93.72	87.47	81.06
shhs-200141	78.2	0	96.12	76.64	90.5
shhs-200142	95.27	25	89.65	95.75	84.58
shhs-200143	94.85	0	80.14	98.49	24.12
shhs-200144	89.39	0	93.54	96.53	92.21
shhs-200145	83.98	0	71.4	0	0
shhs-200146	93.28	51.61	97.99	91.43	95.54
shhs-200147	95.98	23.08	92.85	90.76	87.89
shhs-200148	83.17	0	76.61	7.06	57.61
shhs-200149	95.61	58.82	94.38	89.71	92.23
shhs-200150	66.21	0	86.44	89.34	60.29
shhs-200151	81.76	14.81	78.98	33.01	89.27
shhs-200152	90.91	16.95	92.25	89.43	69.35
shhs-200153	96.23	9.52	92.27	58.33	87.5

(*Continued*)

Table 5 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200154	94.81	12.31	92.65	75.21	82.97
shhs-200155	94.8	28.95	86.85	89.92	58.43
shhs-200156	96.34	0	81.76	86.83	57.75
shhs-200157	90.66	9.76	63.85	39.19	62.35
shhs-200158	75	17.39	89.06	81.87	35.71
shhs-200159	88.03	16.67	73.5	2.44	74.67
shhs-200160	98.86	26.09	91.63	89.69	88.33
shhs-200161	96.26	0	84.6	67.67	32
shhs-200162	94.68	22.22	96.16	95.19	88.31
shhs-200163	77.93	31.58	82.91	94.12	89.91
shhs-200164	83.43	43.48	77.14	83.09	67.56
shhs-200165	91.16	7.5	92.7	80.97	73.23
shhs-200166	50.19	0	75.48	78.59	85.28
shhs-200167	73.29	42.67	87.09	78.67	76.41
shhs-200168	97.68	38.89	91.18	85.88	78.79
shhs-200169	93.45	24.24	91.49	54.79	84.72
shhs-200170	83.95	21.43	87.44	89.64	40
shhs-200171	94.82	14.63	92.95	89.38	89.08
shhs-200172	93.09	6.56	85.39	94.66	31.23
shhs-200173	92.5	14.63	70.99	46.38	38.36
shhs-200174	97.13	57.14	95.62	89.04	92.23
shhs-200175	89.05	14.81	80.74	100	69.83
shhs-200176	65.68	8.7	62.92	80.92	30.63
shhs-200177	81.71	26.83	70.32	43.14	85.26
shhs-200182	81.62	5.88	63.56	92.72	8.77
shhs-200183	97.18	34.78	80.31	0	63.16
shhs-200184	97.28	25	83.76	78.81	39.15
shhs-200185	62.2	9.62	66.37	1.65	2.68
shhs-200186	88.89	33.33	93.23	47.27	71.49
shhs-200187	95.42	33.33	94.81	54.84	71.07
shhs-200188	92.75	0	87.06	8	46.67
shhs-200189	93.64	0	84.39	92.84	80.98
shhs-200190	96.04	49.23	80.65	44.6	60.66

(*Continued*)

Table 5 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200191	64.17	23.61	64.76	38.71	72.97
shhs-200192	87.82	22.5	93.35	36.36	85.41
shhs-200193	93.02	33.85	81.68	78.9	62.2
shhs-200194	83.71	23.53	80.13	89.08	71.13
shhs-200195	87.2	0	93.53	79.7	0
shhs-200196	94.08	14.29	81.21	86.05	88.46
shhs-200197	80.43	23.53	74.74	84.12	59.36
shhs-200198	77.59	7.92	54.37	100	0
shhs-200199	90.68	8	94.77	75.51	86.47
shhs-200200	93.99	23.26	94.63	88.03	95.88
Average F1	88.17	19.78	85.21	68.28	67.49

Table 6 The Recall of 80 Sets of Test Data

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200121	94.03	6.45	96.83	98.11	92.31
shhs-200122	91.07	9.52	91.85	88.61	100
shhs-200123	94.14	25	94.02	100	85.96
shhs-200124	92.68	15.79	78.61	21.43	90.53
shhs-200125	96.94	23.08	100	64.47	100
shhs-200126	86.65	50	99.14	0	100
shhs-200127	94.69	20.93	67.9	100	87.01
shhs-200128	98.1	24.62	91.63	100	97.47
shhs-200129	63.83	42.86	87.02	100	76.44
shhs-200130	100	35.38	65.11	0	93.72
shhs-200131	96.67	12.5	100	81.68	79.67
shhs-200132	78.48	18.18	100	62.96	77.25
shhs-200133	99.74	28.38	93.26	58.82	99.49
shhs-200134	93.37	27.27	90.57	75.93	90
shhs-200135	86.21	3.25	100	42.17	40.35
shhs-200136	96.74	37.5	94.09	13.54	100
shhs-200137	82.68	11.11	100	86.89	80.95
shhs-200138	100	6.25	86.06	6.58	16.44

(*Continued*)

Table 6 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200139	93.21	28.21	87.8	97.81	80
shhs-200140	86.67	7.69	98.6	77.73	100
shhs-200141	74.29	0	100	67.77	90.5
shhs-200142	95.27	18.18	92.71	100	74.91
shhs-200143	92.26	0	100	100	13.71
shhs-200144	97.89	0	88.76	99.21	100
shhs-200145	94.07	100	56.22	100	100
shhs-200146	87.72	66.67	100	86.27	100
shhs-200147	92.27	15	90.08	100	100
shhs-200148	97.99	0	73.64	5.45	44.92
shhs-200149	93.71	55.56	91.8	88.41	100
shhs-200150	50.52	0	90.94	96.27	63.64
shhs-200151	69.55	14.29	100	65.38	92.04
shhs-200152	98.75	9.62	88.68	100	72.27
shhs-200153	92.74	11.76	100	100	80
shhs-200154	92.99	6.67	100	61.07	100
shhs-200155	96.47	45.83	89.02	100	42.39
shhs-200156	95.95	0	88	92.2	41.41
shhs-200157	83.79	16.67	100	25.22	50
shhs-200158	62.61	11.32	100	75.96	23.81
shhs-200159	80.26	12	100	1.23	86.6
shhs-200160	100	15.79	98.99	81.31	96.36
shhs-200161	97.17	0	100	51.61	19.35
shhs-200162	92.35	12.82	97.33	96.74	96.45
shhs-200163	64.34	37.5	79.55	100	93.33
shhs-200164	72.28	29.41	67.11	100	85.02
shhs-200165	84.45	4	93.53	71.43	100
shhs-200166	100	0	69.57	65	100
shhs-200167	61.46	48.48	90.56	88.06	71.88
shhs-200168	97.23	33.33	90.84	92.41	79.27
shhs-200169	100	18.6	93.48	41.67	82.91
shhs-200170	94.44	13.64	86.65	100	25.71
shhs-200171	93.21	9.38	95.09	82.11	100

(*Continued*)

Table 6 (Continued).

Sample	Wake(%)	NREM1(%)	NREM2(%)	NREM3(%)	REM(%)
shhs-200172	96.88	7.69	100	94.2	18.67
shhs-200173	100	10.71	100	30.38	23.73
shhs-200174	97.63	51.85	100	86.09	85.99
shhs-200175	82.18	9.09	71.33	100	100
shhs-200176	95.71	4.76	49.72	69.74	40.69
shhs-200177	89.94	61.11	54.95	100	95.68
shhs-200178	78.16	25	78.22	69.3	100
shhs-200179	96.63	12.5	100	83.49	63.64
shhs-200180	95.17	4.08	91.69	100	100
shhs-200181	91.75	12	90.87	100	71.55
shhs-200182	68.95	3.39	100	91.7	4.61
shhs-200183	96.22	23.53	100	0	100
shhs-200184	98.17	20	96.22	83.91	25.34
shhs-200185	83.39	26.32	71.81	0.84	1.36
shhs-200186	93.46	22.95	92.1	65	76.72
shhs-200187	100	24.69	92.8	56.67	81.13
shhs-200188	99.1	0	82.84	11.76	36.08
shhs-200189	88.66	0	96.05	88.68	81.37
shhs-200190	92.72	69.57	100	29.81	48.68
shhs-200191	100	15.74	59.33	24.19	77.88
shhs-200192	99.17	17.31	92.66	23.08	95.27
shhs-200193	87.97	28.21	100	70.49	46.43
shhs-200194	72.46	33.33	82.26	100	62.5
shhs-200195	90.95	0	100	67.95	0
shhs-200196	89.1	7.94	81.54	100	92
shhs-200197	86.05	14.29	76.1	89.91	49.41
shhs-200198	67.62	14.81	61.4	100	0
shhs-200199	84.92	5.26	92.69	69.81	100
shhs-200200	88.92	13.16	98.4	100	94.9
Average Recall	89.30	19.75	89.35	71.48	72.35

Table 7 The Precision of 20 Clinical Data

Sample	Wake	N1	N2	N3	N4
1	0.97	0.44	0.83	0.99	0.18
2	0.94	0.33	0.64	0.88	0.43
3	0.56	0.50	0.46	1.00	0.01
4	0.98	0.15	0.91	0.98	0.59
5	0.97	1.00	0.76	0.95	0.43
6	0.89	0.55	0.85	0.92	0.67
7	0.95	NA	0.77	0.36	0.00
8	0.96	1.00	0.90	0.99	0.71
9	0.97	0.00	0.29	0.45	0.22
10	0.89	0.67	0.89	0.99	0.64
11	0.89	NA	0.77	0.97	0.00
12	0.96	0.43	0.68	0.94	0.44
13	0.92	NA	0.57	1.00	0.07
14	0.98	0.00	0.63	1.00	0.04
15	0.88	0.33	0.78	0.84	0.17
16	0.91	0.00	0.20	0.04	0.58
17	0.79	0.00	0.76	1.00	0.21
18	0.61	0.00	0.32	0.81	0.64
19	0.97	0.48	0.73	1.00	0.38
20	0.97	0.28	0.56	1.00	0.01
Average Precision	0.90	0.36	0.67	0.86	0.32

Table 8 The Accuracy of 20 Clinical Data

Sample	Wake	N1	N2	N3	N4
1	0.94	0.55	0.85	0.80	0.91
2	0.87	0.54	0.82	0.79	0.95
3	0.67	0.51	0.60	0.52	0.95
4	0.95	0.55	0.95	0.96	0.95
5	0.94	0.52	0.92	0.67	0.94
6	0.90	0.54	0.93	0.79	0.87
7	0.94	0.50	0.88	0.94	NA
8	0.94	0.53	0.96	0.92	0.93
9	0.88	0.50	0.86	0.55	0.96

(*Continued*)

Table 8 (Continued).

Sample	Wake	N1	N2	N3	N4
10	0.93	0.51	0.93	0.88	0.95
11	0.82	0.50	0.81	0.71	NA
12	0.94	0.54	0.88	0.94	0.73
13	0.87	0.50	0.82	0.67	0.53
14	0.94	0.50	0.83	0.69	0.93
15	0.82	0.55	0.88	0.65	0.95
16	0.72	0.50	0.62	0.54	0.61
17	0.80	0.50	0.87	0.54	0.88
18	0.56	0.50	0.53	0.79	0.59
19	0.95	0.63	0.89	0.68	0.88
20	0.94	0.59	0.84	0.64	0.95
Average Accuracy	0.87	0.53	0.83	0.73	0.86

Table 9 The F1-Score of 20 Clinical Data

Sample	Wake	N1	N2	N3	N4
1	0.94	0.18	0.79	0.74	0.30
2	0.87	0.15	0.71	0.71	0.59
3	0.68	0.05	0.40	0.09	0.02
4	0.95	0.13	0.91	0.96	0.72
5	0.94	0.06	0.84	0.50	0.59
6	0.90	0.14	0.89	0.71	0.72
7	0.93	NA	0.84	0.51	NA
8	0.94	0.12	0.93	0.91	0.79
9	0.89	NA	0.44	0.20	0.35
10	0.92	0.03	0.90	0.85	0.77
11	0.93	NA	0.71	0.59	NA
12	0.96	0.13	0.74	0.91	0.47
13	0.94	NA	0.64	0.50	0.08
14	0.94	NA	0.70	0.56	0.08
15	0.91	0.15	0.79	0.46	0.29
16	0.90	NA	0.24	0.07	0.34
17	0.85	NA	0.78	0.14	0.33

(*Continued*)

Table 9 (Continued).

Sample	Wake	N1	N2	N3	N4
18	0.73	NA	0.18	0.69	0.29
19	0.96	0.35	0.79	0.53	0.52
20	0.95	0.23	0.66	0.43	0.02
Average F1	0.90	0.14	0.69	0.55	0.41

Table 10 The Recall of 20 Clinical Data

Sample	Wake	N1	N2	N3	N4
1	0.91	0.11	0.76	0.59	0.93
2	0.81	0.09	0.81	0.60	0.97
3	0.87	0.02	0.36	0.05	1.00
4	0.93	0.11	0.91	0.93	0.94
5	0.91	0.03	0.95	0.34	0.94
6	0.92	0.08	0.92	0.58	0.79
7	0.92	0.00	0.91	0.90	NA
8	0.93	0.06	0.95	0.84	0.89
9	0.82	0.00	0.90	0.13	1.00
10	0.94	0.02	0.92	0.75	0.95
11	0.98	0.00	0.66	0.43	NA
12	0.97	0.08	0.81	0.88	0.50
13	0.96	0.00	0.73	0.33	0.09
14	0.91	0.00	0.78	0.38	1.00
15	0.95	0.10	0.81	0.31	0.94
16	0.90	0.00	0.31	0.15	0.24
17	0.92	0.00	0.80	0.07	0.83
18	0.89	0.00	0.13	0.60	0.19
19	0.94	0.27	0.86	0.36	0.84
20	0.93	0.20	0.79	0.28	1.00
Average Recall	0.92	0.06	0.75	0.48	0.78

As shown in Tables 3–6, in the database,

1. The average accuracy of 80 groups of samples is 83.24%.
2. For the Wake and N2 stages, the precision and recall rates are both over 80%. The algorithm has the strongest staging ability in the Wake and N2 stages.

Table 11 The AUC of 20 Clinical Data

Sample	AUC
1	0.8047
2	0.8219
3	0.7861
4	0.8917
5	0.819
6	0.8166
7	0.7873
8	0.8929
9	0.7962
10	0.8942
11	0.7071
12	0.8178
13	0.7154
14	0.8383
15	0.7896
16	0.6455
17	0.753
18	0.6362
19	0.7849
20	0.7984
Average Recall	0.78984

3. The algorithm has the second highest staging ability in N3 and REM. For the N3 and REM stages, the precision and recall rates are above 70%. Few samples were misjudged and missed.
4. For the N1 stage, the algorithm staging ability is low.

In the clinical data,

1. The average accuracy of 20 groups of samples is 76.37%.
2. For the Wake and N3 stages, the precision and recall rates are both above 85%. The algorithm has the strongest staging ability in the Wake and N3 stages.
3. The algorithm has the second strongest staging ability in N2 and REM. For the Wake, N2, and REM stages, the recall rates are over 70%. Few samples were misjudged and missed.
4. For the N1 stage, the algorithm staging ability is low.
5. The AUC value of each sleep stage is 0.7637.

Table 12 Performance Comparison of Different Methods (F1-Score)

Methods	Wake	N1	N2	N3	REM
SVM (Sleep-EDF-153)	80.3	13.5	79.5	57.1	58.7
RF (Sleep-EDF-153)	81.6	23.2	80.6	65.8	60.8
XG Boost (SHHS1)	88.2	19.8	85.2	68.3	67.5

Table 13 Performance Comparison of Different Methods (Precision)

Methods	Wake	N1	N2	N3	REM
GAC-SleepNet (SHHS1)	89.7	53.4	81.7	89.0	88.0
XG Boost (SHHS1)	89.8	34.3	83.8	76.2	76.0

The sleep staging result of the algorithms, which run in the database and clinical data, are similar. The wake stage exhibits the best staging results because its brain wave characteristics are significantly different from those of the other stages. The N1 stage is the transition stage between the Wake and N2 stages. It is easy to misjudge the N1 stage as the Wake and N2 stages, so the ability to recognize the N1 stage is unfavorable. The main EEG of the N2 and N3 stages is sleep spindle and K wave. The characteristics of the two stages are similar, so the staging results of N2 and N3 are similar. In clinical trials, there may be more objective factors interfering with the REM stage, so the accuracy of REM stage recognition is weak.

Using a computer to automatically identify sleep stages can reduce the workload of doctors and achieve the effect of assisting diagnosis.

We compare XG Boost in this paper with other 2 traditional machine learning models (Support Vector Machine and Random Forest) on publicly available datasets as shown in Table 12.[27] Among these machine learning methods, XG Boost algorithm performs better than SVM and RF.

We compare XG Boost in this paper with deep learning model GAC-SleepNet in Table 13. GAC-SleepNet uses the characteristic information in the dual structure of the graph structure and the Euclidean structure for the classification of sleep stages. In the graph structure, this study uses a graph convolutional neural network to learn the deep features of each sleep stage and converts the features in the topological structure into feature vectors by a multi-layer perceptron. In the Euclidean structure, this study uses convolutional neural networks to learn the temporal features of sleep information and combine attention mechanism to portray the connection between different sleep periods and EEG signals while enhancing the description of global features to avoid local optima.[28] For stages with less obvious features, the accuracy of the GAC-Sleep NET algorithm is higher than that of the XB Boost. This shows that the results of machine learning are more dependent on features. Although some deep learning models can achieve high precision, GAC-Sleep Net models are difficult to tune and optimize. So both kinds of methods have advantages and disadvantages. However, multi-modal deep learning is the trend of development so how to combine the advantages of machine learning and deep learning is worth studying.

Prospects of Interpretability Methods

The algorithm proposed in this study can be further improved and explored:

1. Optimized feature extraction. In this study, we extracted the time domain, frequency domain, and nonlinear characteristics of the signal and attempted to improve the accuracy of the model by extracting more relevant features.

2. Feature dimension reduction. In this study, the PCA method was used to select the optimal feature of 20 dimensions. However, it reduced the accuracy of the model. Other methods for feature dimensionality reduction can be attempted to ensure model accuracy.
3. The imbalanced samples. This study increased the weight of minority class samples such as N1. However, the precision and recall rates of the N1 stage are both low. Deep exploration can be performed with the physiological characteristics of N1.
4. More clinical trials. More clinical trials can be conducted to verify the stability of the system and the clinical applicability of the algorithm.
5. Explore deep learning models and continuously optimize the models.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Qian S, Litao Z, Yandan L, Yong G. Application of refined operation analysis in construction planning of sleep medicine center. *Jiangsu Health Ser Manag*. 2019;30(11):1435–1437. doi:10.3969/j.issn.1005-7803.2019.11.024
2. Christensen JA, Koch H, Frandsen R, et al. Classification of iRBD and Parkinson's disease patients based on eye movements during sleep. *Annu Int Conf IEEE Eng Med Biol Soc*. 2013;2013:441–444. doi:10.1109/EMBC.2013.6609531
3. Shunpeng H, Shuang T, Hui C, et al. Clinical features and polysomnography of depression-related insomnia. *Psychol Mthly*. 2021;(3):10–11. doi:10.19738/j.cnki.psy.2021.03.005
4. Yumo D. Effects of targeted sleep nursing on PSG indexes and depression scale in the treatment of patients with depression. *Med Theory Pract*. 2018;31(14):2184–2186. doi:10.19381/j.issn.1001-7585.2018.14.076
5. A.O.S.M. American. *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Vol. 23. Westchester, IL: American Academy of Sleep Medicine; 2007.
6. Sankri-Tarbichi AG. Obstructive sleep apnea-hypopnea syndrome: etiology and diagnosis. *Avicenna J Med*. 2012;2(1):3–8. doi:10.4103/2231-0770.94803
7. Yilmaz A, Mutlu MA, Günay M, Demır AU. Analysis of portable apnea devices: comparison with polysomnographic (psg) records. *IEEE*; 2019: 1–4.
8. Kamble N, Yadav R, Lenka A, Kumar K, Nagaraju BC, Pal PK. Impaired sleep quality and cognition in patients of Parkinson's disease with REM sleep behavior disorder: a comparative study. *Sleep Med*. 2019;62:1–5. doi:10.1016/j.sleep.2019.04.001
9. Duarte FA, Paul KC, Bronstein JM, Keener AM, Ritz B. Clinical progression in Parkinson's disease with features of REM sleep behavior disorder: a population-based longitudinal study. *Parkinsonism Relat Disord*. 2019;62:105–111. doi:10.1016/j.parkreldis.2019.01.018
10. Güneş S, Polat K, Yosunkaya Ş. Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. *Expert Syst Appl*. 2010;37(12):7922–7928. doi:10.1016/j.eswa.2010.04.043
11. da Silveira TLT, Kozakevicius AJ, Rodrigues CR. Single-channel EEG sleep stage classification based on a streamlined set of statistical features in wavelet domain. *Med Biol Eng Comput*. 2017;55(2):343–352. doi:10.1007/s11517-016-1519-4
12. Seifpour S, Niknazar H, Mikaeili M, Nasrabadi AM. A new automatic sleep staging system based on statistical behavior of local extrema using single channel EEG signal. *Expert Syst Appl*. 2018;104:277–293. doi:10.1016/j.eswa.2018.03.020
13. Michielli N, Acharya UR, Molinari F. Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals. *Comput Biol Med*. 2019;106:71–81. doi:10.1016/j.compbiomed.2019.01.013
14. Fang Y, Tao L. Automatic sleep staging based on energy characteristics and fuzzy entropy. *Chin J Med Phys*. 2017;34(9):968–972. doi:10.3969/j.issn.1005-202X.2017.09.022
15. Wenbing F, Xuefeng L, Yanyang Z. Automatic sleep staging based on single-channel EEG signals. *Computer App*. 2017;37(z2):318–321.
16. Goldberger AL, Amaral LAN, Glass L, et al. PhysioBank, physiotoolkit, and physionet: components of a new research resource for complex physiologic signals. *Circulation*. 2000;101(23):E215–E220. doi:10.1161/01.cir.101.23.e215
17. Younes M, Azarbarzin A, Reid M, Mazzotti DR, Redline S. Characteristics and reproducibility of novel sleep EEG biomarkers and their variation with sleep apnea and insomnia in a large community-based cohort. *Sleep*. 2021;44(10). doi:10.1093/sleep/zsab145
18. Yalu W, *Research on Sleep Staging Algorithm Based on Support Vector Machine Classifier*. [M.D thesis]. BME, ZheJiang University; Hang Zhou: 2019.
19. Huang S, Cai N, Pacheco PP, Narrandes S, Wang Y, Xu W. Applications of support vector machine (SVM) learning in cancer genomics. *Cancer Genom Proteom*. 2018;15(1):41–51. doi:10.21873/cgp.20063
20. Reynolds E, Callaghan B, Banerjee M. SVM-CART for disease classification. *J Appl Stat*. 2019;46(16):2987–3007. doi:10.1080/02664763.2019.1625876
21. Chen T. Xgboost: extreme gradient boosting. *R Package Version*. 2015;1(4):1–4.
22. Hassan AR, Bhuiyan MIH. A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features. *J Neurosci Meth*. 2016;271:107–118. doi:10.1016/j.jneumeth.2016.07.012
23. Ogunleye A, Wang QG. XGBoost model for chronic kidney disease diagnosis. *IEEE/ACM Trans Comput Biol Bioinform*. 2020;17(6):2131–2140. doi:10.1109/TCBB.2019.2911071
24. Wang X, You X, Zhang L, et al. A radiomics model combined with XGBoost may improve the accuracy of distinguishing between mediastinal cysts and tumors: a multicenter validation analysis. *Ann Transl Med*. 2021;9(23):1737. doi:10.21037/atm-21-5999

25. Chambon S, Galtier MN, Arnal PJ, Wainrib G, Gramfort A. A deep learning architecture for temporal sleep stage classification using multivariate and multimodal time series. *IEEE Trans Neural Syst Rehabil Eng*. 2018;26(4):758–769. doi:10.1109/TNSRE.2018.2813138
26. Xiaoyan L. *Clinical EEG*. China: People's Health Publishing House; 2006.
27. Jia Z, Lin Y, Wang J, Wang X, Xie P, Zhang Y. Salient Sleep Net: multi-modal Salient Wave Detection Network for Sleep Staging. *ArXiv abs/2105 13864*. 2021;2021:1.
28. Li T, Gong Y, Lv Y, Wang F, Hu M, Wen Y. GAC-SLEEPNET: a dual-structured sleep staging method based on graph structure and Euclidean structure. *Comput Biol Med* 2023;165:107477. doi:10.1016/j.compbiomed.2023.107477

Nature and Science of Sleep

Dovepress

Publish your work in this journal

Nature and Science of Sleep is an international, peer-reviewed, open access journal covering all aspects of sleep science and sleep medicine, including the neurophysiology and functions of sleep, the genetics of sleep, sleep and society, biological rhythms, dreaming, sleep disorders and therapy, and strategies to optimize healthy sleep. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/nature-and-science-of-sleep-journal