

# Accurate Machine Learning-Based Automated Sleep Staging Using Clinical Subjects with Suspected Sleep Disorders



Santosh Kumar Satapathy, Ravisankar Malladi,  
and Hari Kishan Kondaveeti

**Abstract** Sleep is a basic requirement of human life. It is one of the vital roles in human life to maintain the proper mental health, physical health, and quality of life. In this proposed research work, we conduct an automated sleep staging classification system to a proper investigation of irregularities that occurred during sleep based on single-channel electroencephalography (EEG) signal using machine learning techniques. The major advantage of this proposed research work over the standard polysomnography method are: (1) It measures the sleep irregularities during sleep by considering three different medical condition subjects of different gender with different age groups. (2) One more important objective of this proposed sleep study is that here we obtain different session recordings to investigate sleep abnormality patterns, which can help to find better diagnosis toward the treatment of sleep-related disorder. (3) In the present work, we have obtained two different time-framework epochs from individual subjects to check which window size is more effective toward identification on sleep irregularities. The present research work based on a two-state sleep stage classification problem based on a single channel of EEG signal was performed in a different stepwise manner such as the acquisition of data from participated subjects, preprocessing, feature extraction, feature selection, and classification. We obtained the polysomnographic data from the ISRUC-Sleep data repository for measuring the performances of the proposed framework, where the sleep stages are visually labeled. The obtained results demonstrated that the proposed methodologies achieve a high classification accuracy of 99.46 and 97.46% using SVM and DT classifiers, respectively, and which support sleep experts to accurately measure the irregularities that occurred during sleep and also help the clinicians to evaluate the presence and criticality of sleep-related disorders.

---

S. K. Satapathy (✉)

Department of CSE, Pondicherry Engineering College, Puducherry, India

e-mail: [Santosh.satapathy@pec.edu](mailto:Santosh.satapathy@pec.edu)

R. Malladi

Department of CSE, Gudlavalleru Engineering College (A), Andhra Pradesh, India

H. K. Kondaveeti

School of Computer Science Engineering, VIT-AP University, Amaravati, Andhra Pradesh, India

**Keywords** Sleep staging · Electroencephalogram · Feature screening · Machine learning

## 1 Introduction

Sleep is one of the important physiological activities for the human body, which directly controls memory consolidation, and it also decides the performance of the daily activities. Sleep plays an important role in the human body because it represents the primary functions of the human brain. One human individual is spending one-third of its duration as sleep. Proper quality of sleep maintains the physical and mental fitness of the human body, which is alternatively helpful to perform well in workplaces, control emotions, and be able to take proper decisions [1, 2]. Nowadays, it has seen that sleep diseases (SD) are becoming one of the major causes of death across the world. The main reason for this serious health issue is an imbalance of sleep patterns, and it has occurred due to job pressure, and rapid changes in lifestyles. Across the globe. It has been observed that the prevalence of sleep diseases has significantly increased over the past years. According to the report of the Center for Control of Disease and Prevention (CDC) of the US Government, around 9 million populations have difficulty maintaining good quality sleep [3]. According to a survey of the National Highway Traffic Safety Administration in the USA, it has been found that due to the drowsiness factor, around 56,000 to 100,000 car accidents have happened, which directly reported that more than 1500 have died and 71,000 are affected with injuries annually [4]. It has been found that sleep diseases are considered to be the most predominant death cause with the different age groups of populations across the globe. In general, different types of sleep disorders are categorized such as obstructive sleep apnea, insomnia, hypersomnia, narcolepsy, breathing-related disorders, stroke, stress, and cardiovascular diseases [5]. All these diseases progressively increased with age. So early diagnosis is helpful for the human being to prevent the severity of these diseases, and it helps to improve the subject's quality of life. The first most important step for sleep diseases is sleep scoring. The most popular test for analyzing sleep quality is the polysomnography (PSG) test. PSG tests include the signals such as electroencephalogram (EEG), electrocardiogram (ECG), electromyogram (EMG), and electrooculogram (EOG). The entire sleep staging procedures are analyzed according to two available sleep standards such as the Rechtschaffen and Kales (R&K) [6] and the American Academy of Sleep Medicine (AASM) [7]. According to R&K sleep guidelines, the whole sleep cycle is categorized into six sleep stages such as wake stage(W), non-rapid eye movement (NREM stage1 (N1), NREM stage2 (N2), NREM stage3 (N3), and NREM stage4 (N4)) and rapid eye movement (REM) stage. The only changes reflected with the AASM manual incomparable to R&K standards are NREM sleep stages. According to the AASM guidelines, the total sleep stages are five, the NREM stage 3 (N3) and the NREM stage 4 (N4) are combined into one sleep stage called the NREM stage3. Traditionally the sleep scoring procedure was conducted through the visual inspection method, where one

clinician was monitoring the sleep behavior of the subject for 6–8 h. of sleep. This traditional sleep analysis method requires more human resources for monitoring the whole sleep recordings, and also, it consumes more time for analysis, due to more human interpretation, sometimes the results are erroneous [6]. Sometimes it is also one of the major causes of not achieving higher classification accuracy in the classification of sleep stages. With consideration of all these above-mentioned facts, the automated sleep scoring approach has gained a lot of attention in recent researches [7, 8]. Automated sleep scoring not only causes accuracy improvements but also provides quick diagnosis [9]. It has been observed that the PSG test is one of the costly experiments, and it also gives so many unpleasant scenarios for the subjects because of its so much connectivity of wires in the different parts of the body [10, 11]. Henceforth instead of PSG signals, most of the researchers preferred to EEG signal, because it directly provides the brain activities during sleep hours. This helps a lot for analyzing the sleep abnormality, and it is also more popular for its easier recording facility. In general, EEG signals are combinations of different waveforms, which help to characterize the different sleep stages with different frequency bands such as delta band (0–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), spindle (12–14 Hz), sawtooth (2–6 Hz), and k-complex (0.5–1.5 Hz). Finally, the scoring and decisions are taken by the sleep experts through proper interpretation of the quantitative and visual analysis of collected sleep recordings. In some cases, the sleep experts use an algorithm for pre-scoring the entire sleep recordings, and these successive representations of the sleep stages information called hypnograms, which is highly required during the diagnosis of the different types of sleep disorders. Sleep staging is generally a tedious job, which requires highly experienced technicians and experts. This other limitation with subject to sleep staging is variations on sleep scoring from experts to experts, which is also one of the major causes for diagnosing sleep diseases [12, 13].

In this paper, we have obtained a single-channel EEG signal for sleep staging analysis; this approach makes it more interesting because of its ease of operational deployments on mobile devices. It also makes more comfortable situations for the patients due to less cabling used during recordings. It has been observed that most of the contributions with single-channel EEG signals were executed two-step methodology. In the first step, the different hand-engineered features are extracted from the different waveforms, and in the second step, the extracted features are forwarded to a classifier for classifying the sleep stages based on the feature characteristics. In general, it has been seen that most of the authors obtained one of the three following domains of the features [14] (a) time-domain features, (b) frequency-domain features, (c) nonlinear features. It is a very difficult part for sleep experts to manually monitoring the recorded EEG signals, and it raises so many errors because during long 7–8 h. EEG recordings, its hectic situation for sleep experts to monitoring within the 30 s framework and fix the labeling of sleep stages. This approach consumed more time and required more manpower for hours of sleep recordings. To overcome difficulty from the manual approach, nowadays automated sleep stage classification is obtained to analyze the sleep-related disorder and real-time diagnosis, and the most important step is designing sleep stage classification. Currently, overnight sleep study

through polysomnography is one of the standard procedures for measuring sleep irregularities during sleep [14].

## 1.1 Related Work

Several sleep analysis studies were proposed for characterization the sleep-related abnormalities based on the sleep standards recommended by R&K and AASM manuals. Various computational methodologies were proposed by researchers to support sleep experts for assisting sleep staging. Those carried steps were on the information extraction (polysomnography channel selection), on the preprocessing (removing the data artifacts and data normalization), on the feature extraction step (transformation of time- and frequency-domain features), on the feature selection technique (identifying the most relevant features) and finally on the classification algorithm. Here we have presented some comparative studies regarding sleep staging. In [15], the authors have obtained wavelet concept techniques for feature extraction and classified the selected features using the fuzzy algorithm. The classification model provided 85% accuracy. Güneş, K et al. [16] used K-means clustering and feature weighting techniques to design an ASSC system. Welch spectral transform was considered for feature extraction, and those selected features were forwarded to a decision tree (DT) and obtained with an overall accuracy of 83%. Aboalayon [17] used EEG signal and obtained Butterworth bandpass filters and used SVM classifiers and reported 90% classification accuracy.

In [18], extracted features form an empirical mode decomposition of the signal and use bootstrap-aggregating techniques for multi-class sleep staging classifications.

In [19] applied the EEMD algorithm for signal enhancement from single-channel EEG signal, and extracted statistical features are forwarded into boosting techniques, and the reported accuracy for two-six sleep stages is reported as 98.15%, 94.23%, 92.66%, 83.49%, and 88.07%, respectively.

Kristin M. Gunnarsdottir et al. have designed an automated sleep stage scoring system with overnight PSG data. Here the authors extracted both time- and frequency-domain properties from PSG signal, and considered healthy individual subjects with no prior sleep diseases and the extracted properties were classified through DT classifiers. The overall accuracy for test set data was reported as 80.70% [20].

Sriraam, N. et al. used a multi-channel EEG signal from ten healthy subjects. In this study, the author has proposed the automated sleep stage scoring in between wake and stage 1 of sleep. In this research work, spectral entropy features are extracted from the input channels to distinguish the irregularities among the sleep states. The extracted features processed through a multilayer perceptron feedforward neural network and the overall accuracy with 20 hidden units were reported as 92.9%, and subsequently for 40,60, 80, and 100 hidden units in MLP, it was reported as 94.6, 97.2, 98.8, and 99.2, respectively [21].

In [22] proposed two-state sleep staging and the acquired signal decomposed into eight sub-bands, finally 13 features are extracted from each sub-band epoch. The

suitable features are identified through the mRMR feature selection algorithm. The model achieved an overall accuracy of 95.31% through a random forest classifier. Da Silveira et al. used discrete wavelet transform (DWT) for signal segmentation. Skewness, kurtosis, and variance features were extracted from respective input channels. The extracted features were applied to a random forest classifier, and overall accuracy was reported as 90% [23]. Prochaska et al. [24] used polysomnography data features to identify the sleep abnormalities from three different medical conditions of subjects and used an SVM classifier for two-state classifications in between the Wake versus NREM stage and another one in between the Wake versus REM stage. The proposed study achieved an overall accuracy of classification between the Wake-NREM and Wake-REM stage as 85.6% and 97.5%, respectively. Xiaojin Li et al. introduced the hybrid model for identifying the irregularities that occurred in different stages of sleep during the night, and extracted features were forwarded into random forest classifiers. It has been reported that overall classification accuracy has reached 85.95% [25].

## 2 Experimental Data

To analyze the proposed methodology effectiveness, we have obtained the session-1 and session-2 sleep recordings from the subjects who were already affected by the different types of sleep-related disorders. These required recordings were retrieved from the subgroups of ISRUC-Sleep dataset, which is one of the public databases specifically available for sleep research. These recordings were prepared by the groups of domain experts at sleep center in the Hospital of Coimbra University. This dataset contained the recorded subject details from different age groups, gender categories, and medical conditions. All recordings were recorded through the sleep experts in the sleep laboratory at the Hospital of Coimbra University (CHUC). As per our proposed research objective, the first subject used for experimental work from Subgroup-I of ISRUC-sleep repository. The second category of a subject is taken for our proposed experimental work from Subgroup-II of the ISRUC-sleep database. The distribution of sleep stages epochs per individual subjects is presented in Table 1.

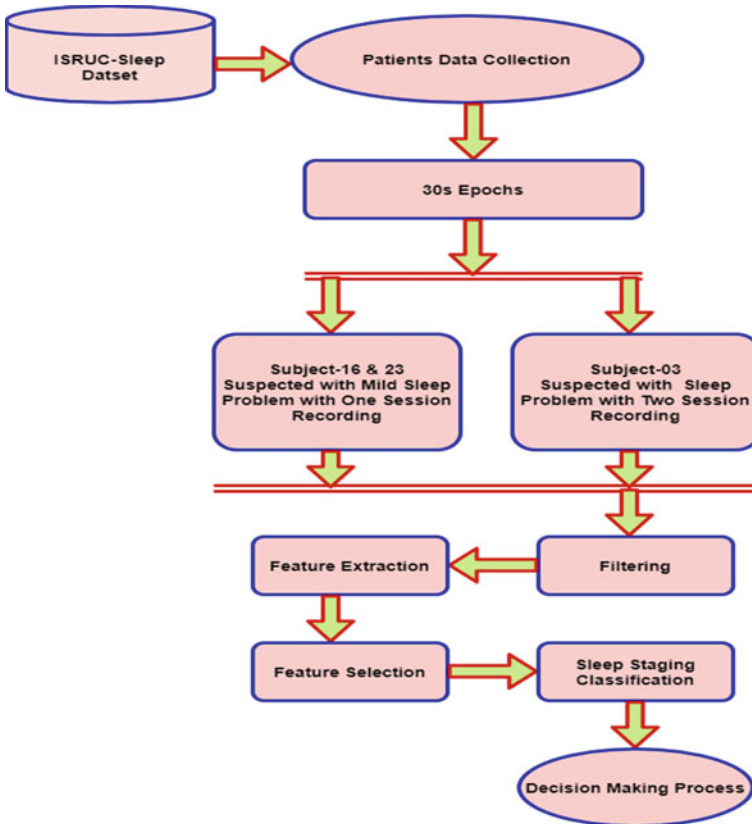
## 3 Methodology

In this work, we proposed a machine learning-based sleep scoring system using a single channel with subjects having different medical conditioned subjects. The main objective of this proposed work is to study the sleep stages behavior of the subjects who were already had some types of sleep diseases symptoms. Additionally, in this research work we also analyzed the sleep quality of the subjects by obtaining

**Table 1** Detailed information of each subject sleep dataset records used in this study

Enrolled subjects	Wake	N-REM1	N-REM2	N3-REM3	REM
Subject-16 Subgroup I/ One Session	128	125	280	120	97
Subject-23 Subgroup I/ One Session	212	99	270	65	104
Subject-3 Subgroup-II/ One Session	68	126	271	175	110
Subject-3 Subgroup-II/ Two Session	76	127	236	168	143

the different session recordings on two different dates. The complete layout of this proposed work is described in Fig. 1.



**Fig. 1** Complete layout of the proposed sleep staging system

**Table 2** Short explanation of the extracted features for this proposed study

Feature set	Feature no.	Feature set	Feature no.
<i>Time-based</i>			
Mean	1	Minimum	3
Maximum	2	Standard Deviation	4
Median	5	Variance	6
Zero Crossing Rate	7	75percentile	8
Signal Skewness	9	Signal Kurtosis	10
Signal Activity	11	Signal Mobility	12
Signal Complexity	13		
<i>Frequency-based</i>			
RSP in delta, theta, beta and alpha bands	14,15,16,17	Power Ratios	18,19,20,21 22,23,24
Band powers	25,26,27,28		

### 3.1 Feature Extraction

The selection of inputs for the classifier is the most valuable for identifying sleep pattern abnormality. Even if obtained highly effective classification model performed very poor performance, if proper inputs are to be identified. It can be found that the different classifiers performed different results for the same set of features; it indicates matching both may found results. On the other part, sometimes we have given some set of features that favors the classification process. It has been found that the sleep behavior of the subjects is highly unstable and non-stationary because the changes characteristics are directly linked with the time and frequency ranges. To properly discriminate the sleep stages, we need to analyze the signals by obtaining the time- and frequency-based parameters. In this study, we have as whole extracted 28 features from the input signal, out of those 13 features are time-based and the other 15 features are frequency-based [26–28]. The obtained features are described in Table 2.

### 3.2 Feature Selection

Next to feature extraction, the other important task with regard to classification problem is screening the relevant parameters which help to model for properly classifying the sleep stages. Sometimes it has been found that the extracted fall the features may not be more effective with respect to analyze of the sleep behavior, which directly put impacts on the classification performance of the models. In our study, we adopted the feature screening techniques as online streaming feature selection (OSFS) techniques to screen the suitable features from the pool of extracted features [29]. The selected features concerning the individual subject are presented in Table 3.

**Table 3** Screened Feature Lists

Name/Gender	Selected features
Subject-16 MALE	F1 <sub>16</sub> ,F2 <sub>16</sub> ,F3 <sub>16</sub> ,F4 <sub>16</sub> ,F5 <sub>16</sub> ,F7 <sub>16</sub> ,F9 <sub>16</sub> ,F10 <sub>16</sub> ,F11 <sub>16</sub> ,F13 <sub>16</sub> ,F14 <sub>16</sub> , F15 <sub>16</sub> ,F22 <sub>16</sub> ,F25 <sub>16</sub> ,F27 <sub>16</sub> ,F31 <sub>16</sub> (16 Features)
Subject-23 FEMALE	F1 <sub>23</sub> ,F2 <sub>23</sub> ,F3 <sub>23</sub> ,F4 <sub>23</sub> ,F5 <sub>23</sub> ,F7 <sub>23</sub> ,F11 <sub>23</sub> ,F12 <sub>23</sub> ,F15 <sub>23</sub> ,F16 <sub>23</sub> ,F17 <sub>23</sub> , F20 <sub>23</sub> ,F24 <sub>23</sub> ,F26 <sub>23</sub> ,F31 <sub>23</sub> ,F33 <sub>23</sub> ,F37 <sub>23</sub> (17 Features)
Subject-03 MALE (Session-1) Recording	F1 <sub>03S1</sub> ,F2 <sub>03S1</sub> ,F3 <sub>03S1</sub> ,F4 <sub>03S1</sub> ,F5 <sub>03S1</sub> ,F7 <sub>03S1</sub> ,F8 <sub>03S1</sub> ,F9 <sub>03S1</sub> ,F11 <sub>03S1</sub> , F12 <sub>03S1</sub> ,F13 <sub>03S1</sub> ,F14 <sub>03S1</sub> ,F21 <sub>03S1</sub> ,F22 <sub>03S1</sub> ,F25 <sub>03S1</sub> (15 Features)
Subject-03 MALE (Session-2) Recording	F1 <sub>03S2</sub> ,F2 <sub>03S2</sub> ,F3 <sub>03S2</sub> ,F4 <sub>03S2</sub> ,F5 <sub>03S2</sub> ,F7 <sub>03S2</sub> ,F8 <sub>03S2</sub> ,F9 <sub>03S2</sub> ,F11 <sub>03S2</sub> , F12 <sub>03S2</sub> ,F13 <sub>03S2</sub> ,F14 <sub>03S2</sub> ,F21 <sub>03S2</sub> ,F22 <sub>03S2</sub> ,F25 <sub>03S2</sub> (15 Features)

## 4 Experimental Results and Discussion

The main intention behind this research is to analyze the changes sleep stages and classifying the sleep stages using machine learning classification models. This entire procedure is called as sleep scoring. In this work, the entire experiments were performed on two different subgroups of sleep recordings one from ISRUC-Sleep subgroup-I and the other from ISRUC-Sleep subgroup-II dataset. The entire sleep staging experiments followed according to the AASM sleep standards. The proposed sleep scoring methodology is executed through four basic steps that are signal preprocessing, feature extraction, feature screening, and finally classification. In this work, we have considered only the single channel of EEG signal for acquisition of the signal recordings. Next to the acquisition, the required signals need to be processed for further eliminating the irrelevant noises and artifacts which are contaminated during recordings in the raw signal and eliminated these muscle artifacts and noisy portions from recorded signals through the Butterworth band-pass filter. In the next phase, a set of experiments were conducted to extract the features from both the time and frequency domains. As a whole 28 features were extracted from recorded signals of the subjects, and the same details are mentioned in Table 2. The size of the feature vectors for all enrolled subjects for 30 s epoch length is  $28 \times 750$ . Matrix dimension for feature vector is feature number  $\times$  epoch number. The next task is the selection of the most efficient features from among the feature vector. To work out this selection experiment, we have applied OSFS feature selection techniques. The matrix representation for feature selection vectors is selected feature number  $\times$  epoch number. These matrixes are  $16 \times 750$ ,  $17 \times 750$ ,  $15 \times 750$ ,  $15 \times 750$  for subject-16, subject-23, subject-03 (session-1 recording), and subject-03 (session-2 recording), respectively, for input length of epoch is 30 s. By implementations of tenfold cross-validation techniques on the SVM [30, 31] and DT [32] classifiers, the selected best features are fed as input to the model. We also conducted a comparative analysis with all these enrolled subjects and their session recordings, and finally,



comparison experimental results are presented according to the single channel of EEG signals and two sleep classes (wake versus sleep). In this proposed study, we have used some criteria of evaluation metrics for measuring the performances of the proposed sleep scoring study. Here, we have considered six performance metrics for analyzing the performance of the proposed methodology such as classification accuracy [33], recall [34], specificity [35], precision [36], F1-score [37], and Kappa score [38]. Analysis of the comparative results from conducted experiments, and obtained results are presented below.

### 4.1 Classification Accuracy of Category-I Subject ISRUC-Sleep Database

In this experimental part, we have obtained two subjects who have been affected by some kind of sleep-related disorders and here from subject session-1 recording recorded by sleep experts for diagnosing the irregularities that happened during sleep hours. Table 4 presents the confusion matrix for two-state sleep stage classification problems for both the subjects 16 and 23 with time length of epoch is 30 s. It has been observed that the SVM depicts an overall classification accuracy of 95.62 and 91.20% achieved through DT classifiers for subject-16. For subject-23, the same classifiers SVM and DT reached overall accuracy of 91.46% and 87.73%, respectively, for epoch length 30 s.

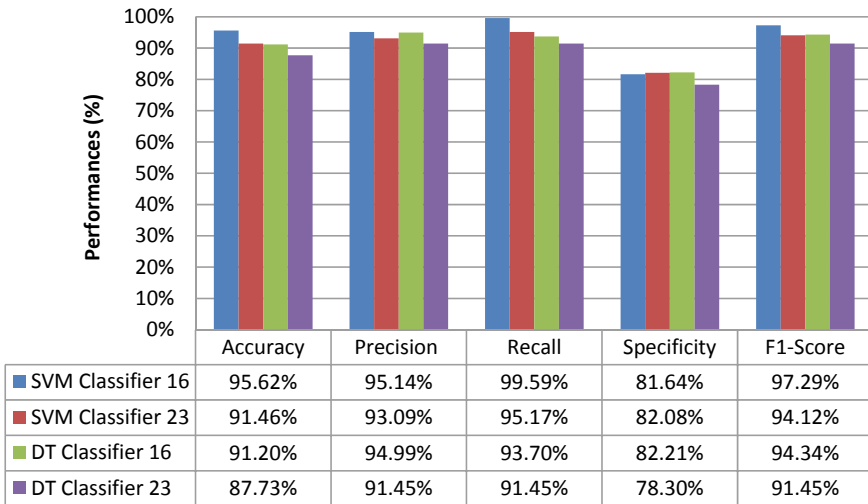
The results achieved from the input of 30 s length epoch for subject-16 and subject-23 are specified in Table 5. Figure 2 displays performance statistics for 30 s epoch length for subject-16 and subject-23.

**Table 4** Confusion matrix of subjects 16 and 23 according to AASM guidelines

<i>Subject-16</i>			
		W	S
SVM	W	133	30
	S	3	584
		W	S
DT	W	133	29
	S	37	551
<i>Subject-23</i>			
		W	S
SVM	W	174	38
	S	512	26
		W	S
DT	W	167	45
	S	29	509

**Table 5** Performance of the proposed SleepEEG study using SVM and DT classifiers

Channel	Subject enrolled epochs length-30s			
	SVM classifier		DT classifier	
C3-A2	16	23	16	23
Accuracy (%)	95.62	91.46	91.20	87.73
Precision (%)	95.14	93.09	94.99	91.45
Recall (%)	99.59	95.17	93.70	91.45
Specificity (%)	81.64	82.08	82.21	78.30
F1-Score (%)	97.29	94.12	94.34	91.45



**Fig. 2** Performance statistics of model with 30 s epoch duration for subject-16 and subject-23 using SVM and DT classifiers

The overall performance value of the proposed Category-I subject ISRUC-Sleep database is measured through the evaluation parameters that are recall, specificity, precision, and F1-score, and it reported for subject-16 as 99.59%, 81.64%, 95.14%, and 97.29% through SVM, 93.70%, 82.21%, 94.99%, and 94.34% through DT, respectively; similarly, the same parameters reached for subject-23 through SVM and DT are 95.17%, 82.08%, 93.09%, and 94.12%, and 91.45%, 78.30%, 87.73%, and 91.45%. The computation of score is of six levels of agreements: 0.81–1, 0.61–0.80, 0.41–0.60, 0.21–0.4, 0.00–0.20, and less than 0 correspond to excellent, substantial, moderate, fair, slight agreement, and poor agreement [38]. Table 6 gives the kappa coefficient score concerning obtained classification techniques for both the subjects 16 and 23, and it has been found from results that all classification techniques are found excellent agreement with subject to best accuracy for investigation on sleep irregularities.

**Table 6** Performance of the accuracy and Kappa score based on the two-state sleep classification problem for subjects 16 and 23

Classifiers	Subject-16		Subject-23	
	Accuracy (%)	Kappa score	Accuracy (%)	Kappa score
SVM	95.62	0.92	91.46	0.78
DT	91.20	0.73	87.73	0.69

### 4.2 Classification Accuracy of Category-II Subject ISRUC-Sleep Database

In the ISRUC-Sleep subgroup-II dataset experiment, the proposed sleep stage classification model has experimented based upon only a single channel with two different session recordings from one gender enrolled subject with suspected sleep-related disorder symptoms. Table 7 represents the confusion matrix for both session recordings of subject 03 with the duration of epoch 30 s.

The achieved results for subject-03 for both the sessions are shown in Table 8. The performance graph results for subject-03 for both session recordings of epoch length 30 s are displayed in Fig. 3.

Figure 3 presents the reported performances for the two-state sleep classification model, Subgroup-II with session-2 recordings for subject-03.

**Table 7** Performance of the proposed SleepEEG study using SVM, and DT Classifiers for Subject-03 (session-2 Recordings)

30 s epoch length			
Subject-03			
Session1_Recording			
		W	S
SVM	W	20	47
	S	20	663
		W	S
DT	W	12	63
	S	55	620

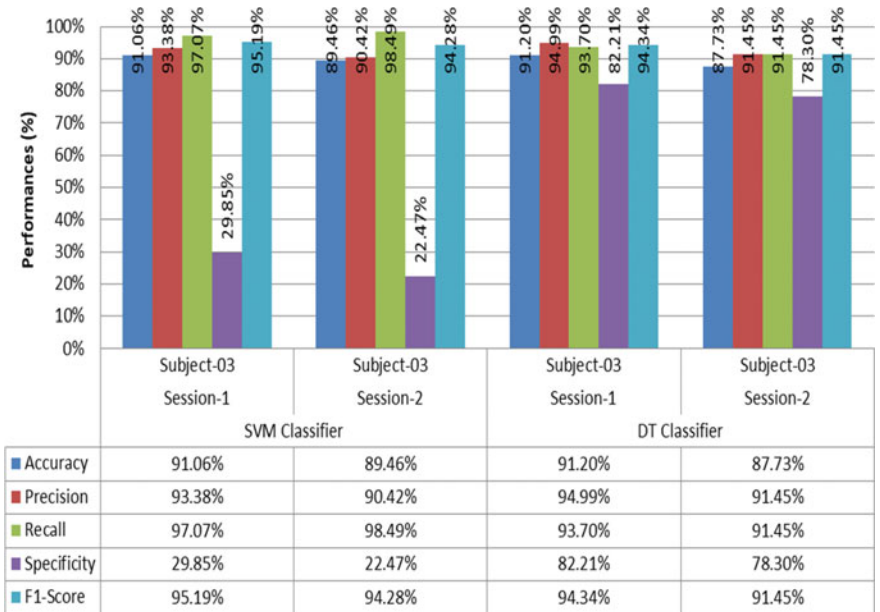
30 s epoch length			
Subject-03			
Session2_Recording			
		W	S
SVM	W	20	69
	S	10	651
		W	S
DT	W	18	63
	S	40	629

**Table 8** Performance of the proposed SleepEEG study using SVM and DT classifiers for Subject-03 (session-2 recordings)

30 s epoch length			
Subject-03			
Session1_Recording			
		W	S
SVM	W	20	47
	S	20	663
		W	S
DT	W	12	63
	S	55	620

30 s epoch length			
Subject-03			
Session2_Recording			
		W	S
SVM	W	20	69
	S	10	651
		W	S
DT	W	18	63
	S	40	629



**Fig. 3** Performance measures using SVM and DT classification techniques for the two-state sleep classification model with session-2 recordings for subject-03 (30 s epochs Length)

**Table 9** Performance of the accuracy and Kappa score based on the two-state sleep classification problem for subject with mild sleep problem having session-2 recording for subject-03

Classifiers	Subject-03 (Session-1 Recording)		Subject-03 (Session-2 Recording)	
	Accuracy (%)	Kappa score	Accuracy (%)	Kappa score
SVM	91.06	0.87	89.46	0.86
DT	84.26	0.67	86.26	0.74

From each subject, here we have acquired two different session recordings; it has been observed that subject 03 with session-1 recording SVM classification model depicts an overall accuracy of 91.06% and 84.26% for DT, respectively. Similarly, it has been found that the classification results of subject 03 with session-2 recordings through SVM and DT were reported as 89.46% and 84.2%. The overall performance of recall, specificity, precision, and F1-score reported with the session-1 recording of ISRUC-Sleep Subgroup-II database of subject-03 through SVM as 97.07%, 29.85%, 93.38%, and 95.19%, similarly for DT classifier, the performances reached 93.70%, 82.21%, 94.99%, and 94.34%. Similarly, the performances with session-2 recordings are reported as 98.49%, 22.47%, 90.42%, and 94.28% through SVM, 91.45%, 78.30%, 91.45%, and 91.45% through DT. The results of the kappa coefficient for subject-03 with both session recordings are presented in Table 9.

For measuring the impact of session recordings for the classification of sleep stages, we have computed the Cohens kappa coefficient; according to session-1 recording for subject-03, the kappa score through SVM and DT is 0.87 and 0.67. From this kappa score, it concludes that DT is not up to the mark performance incomparable to the SVM classification techniques. Similarly for session-2 recording, the kappa performance for SVM and DT is 0.86 and 0.74, respectively.

### 4.3 Comparative Analysis in Between Proposed Study and State-Of-The-Art Works

Here we have made a comparison with other similar contributions work to measure the proposed research work effectiveness toward the identification of sleep disorder. Table 10 presents the comparison of the performances based on single-channel EEG acquisition among the proposed research work results with five contributed works.

**Table 10** Comparison of performances of the proposed work with previously published works

Author	Year	Signal type	Method	Accuracy (%)
Eduardo T. Braun et al. Ref. [27]	2018	Single-channel EEG	FFT + Random forest classifier	97.1
Hassan, A. R. et al. Ref. [18]	2017		Tunable-Q wavelet transform (TQWT) +Bootstrap aggregating	92.43
Diykh, M. et al. Ref. [19]	2016		Structural graph similarity K-means (SGSKM) +SVM classifier	95.93
K. Aboalayon et al. Ref. [17]	2014		Frequency sub-bands features extraction + SVM classifier	92.5
Zhu, G. et al. Ref. [23]	2014		Graph domain features + SVM classifier	96.1
<b>Proposed Work</b>	<b>2020</b>		<b>SleepEEG study +SVM and DT classifier</b>	<b>99.46</b> <b>97.46</b>

## 5 Conclusion and Future Directions

The present proposed research work application showed the most effectiveness in the sleep stage scoring by using a single channel of EEG signal. This proposed SleepEEG study would provide an effective mechanism for handling different health conditions of the subjects with high accuracy of sleep abnormality identification from sleep recordings. The main objective of this application is to analyze the irregularities that occurred during sleep hours from various session recordings, and additionally, this application also successfully deals with the specially aged category of subjects with various disease conditions. The major part of this research work is to find the proper solutions based on irregularity's accuracy during sleep. Another important significance of this proposed SleepEEG study is that, according to our best knowledge, this proposed research work considered different session recordings from the participated subjects in these experimental processes.

This experimental research study provides new directions on scoring sleep stages to identify sleep abnormality through the extraction of different features from both domains such as frequency and time. The major changes are shown between the two different session recordings of sleep stages from two different days, and the general sleep stage classification problem is that annotations of sleep stages are another important source of information. These certain things support for discovering new concepts of investigation on sleep irregularities during sleep, and it may get more advantage for predicting the proper diagnosis plan for treating the disorder.

The proposed scheme automated sleep stage classification based on a single channel of EEG signal gives the benefits with the inclusion of different session

recordings and obtained different health condition subjects. It has been observed from the experimental results that the proposed sleep analysis indicated an excellent agreement between automated sleep staging and the gold standard.

The present research work has certain disadvantages that the (1) data used for the experimental purpose from ISRUC-Sleep repository for statistical evaluation was relatively small, (2) only we have included single channel of EEG signal was used for classification, (3) we have not considered the subjects who were effects of diseases, such as narcolepsy and insomnia.

## References

1. Aboalayon KAI, Faezipour M, Almuhammadi WS et al (2016) Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation. *Entropy* 18:272
2. Chung M-H, Kuo TB, Hsu N, Chu H, Chou K-R, Yang CC (2009) Sleep and autonomic nervous system changes? Enhanced cardiac sympathetic modulations during sleep in permanent night shift nurses. *Scand J Work Environ Health*: 180–187
3. Heyat MBB, Akhtar F, Azad S (2016) Comparative analysis of original wave & filtered wave of EEG signal used in the detection of bruxism medical sleep syndrome. *Int J Trend Sci Res Develop* 1(1):7–9
4. Heyat MBB, Akhtar SF, Azad S (2016) Power spectral density are used in the investigation of insomnia neurological disorder. In: *Proceedings of pre-congress symposium of organized Indian Social Science Academy (ISSA) King George Medical University State Takmeelut-Tib College Hospital, Lucknow, Uttar Pradesh, 2016*, pp 45–50
5. Rahman F, Heyat S (2016) An overview of narcolepsy. *IARJSET* 3:85–87
6. Kim T, Kim J, Lee K (2018) Detection of sleep disordered breathing severity using acoustic biomarker and machine learning techniques. *BioMed Eng OnLine* 17:16
7. Siddiqui M, Srivastava G, Saeed S (2016) Diagnosis of insomnia sleep disorder using short time frequency analysis of PSD approach applied on EEG signal using channel ROC-LOC. *Sleep Science* 9(3):186–191
8. Rechtschaffen A (1968) A manual for standardized terminology techniques and scoring system for sleep stages in human subjects. *Brain Inf Serv*
9. Iber C (2007) The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications. *American Academy of Sleep Medicine*
10. Carskadon MA, Dement WC (2017) Normal human sleep: an overview. In: *Principles and practice of sleep medicine*. In: Kryger M, Roth T, Dement WC (eds) 6th edn. Elsevier, Amsterdam, The Netherlands, pp 15–24. <https://doi.org/10.1016/B978-0-323-24288-2.00002-7>
11. Acharya R, Faust O, Kannathal N, Chua T, Laxminarayan S (2008) Nonlinear analysis of EEG signals at various sleep stages. *Comput Methods Progr Biomed* 80(1):37–45
12. Holland JV, Dement WC, Raynal DM (1974) Polysomnography: a response to a need for improved communication. In: *Presented at the 14th annual meeting association of psychophysiology study sleep*
13. Acharya UR et al (2015) Nonlinear dynamics measures for automated EEG-based sleep stage detection. *Eur Neurol* 74(5–6):268–287
14. Spriggs WH (2014) *Essentials of polysomnography*. World Headquarter Jones & Bartlett Publishers, MA, United States
15. Obayya M, Abou-Chadi F (2014) Automatic classification of sleep stages using EEG records based on Fuzzy c-means (FCM) algorithm. In: *Radio science conference (NRSC), 2014 31st national*, 2014, pp 265–272
16. Güneş KP, Yosunkaya Ş (2010) Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting. *Expert Syst Appl* 37:7922–7928

17. Aboalayon K, Ocbagabir HT, Faezipour M (2014) Efficient sleep stage classification based on EEG signals. In: Systems, applications and technology conference (LISAT), 2014 IEEE Long Island, 2014, pp 1–6
18. Hassan AR, Subasi A (2017) A decision support system for automated identification of sleep stages from single-channel EEG signals. *Knowl-Based Syst* 128:115–124
19. Hassan AR, Hassan Bhuiyan MI (2016) Automatic sleep scoring using statistical features in the EMD domain and ensemble methods. *Biocybern Biomed Eng* 36(1):248–255
20. Gunnarsdottir KM, Gamaldo CE, Salas RME, Ewen JB, Allen RP, Sarma SV (2018) A novel sleep stage scoring system: combining expert-based rules with a decision tree classifier. In: 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC). <https://doi.org/10.1109/embc.2018.8513039>
21. Sriraam N, Padma Shri TK, Maheshwari U (2016) Recognition of wake-sleep stage 1 multi-channel eeg patterns using spectral entropy features for drowsiness detection. *Australas Phys Eng Sci Med* 39(3):797–806. <https://doi.org/10.1007/s13246-016-0472-8>
22. Memar P, Faradji F (2018) A Novel Multi-Class EEG-Based Sleep Stage Classification System. *IEEE Trans Neural Syst Rehabil Eng* 26(1):84–95. <https://doi.org/10.1109/tnsre.2017.2776149>
23. Da Silveira TLT, Kozakevicius AJ, Rodrigues CR (2016) Single-channel EEG sleep stage classification based on a streamlined set of statistical features in wavelet domain. *Med Biol Eng Compu* 55(2):343–352. <https://doi.org/10.1007/s11517-016-1519-4>
24. Pernkopf F, O’Leary P (2001) Feature selection for classification using genetic algorithms with a novel encoding. In: Skarbek W (ed) *Computer analysis of images and patterns*. CAIP 2001. Lecture notes in computer science, vol 2124. Springer, Berlin, Heidelberg
25. Wutzl B, Leibnitz K, Rattay F, Kronbichler M, Murata M et al (2019) Genetic algorithms for feature selection when classifying severe chronic disorders of consciousness. *PLoS ONE* 14(7):e0219683. <https://doi.org/10.1371/journal.pone.0219683>
26. Zhu G, Li Y, Wen PP (2014) Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal. *IEEE J Biomed Health Inform* 18(6):1813–1821
27. Braun ET, Kozakevicius ADJ, Da Silveira TLT, Rodrigues CR, Baratto G (2018) Sleep stages classification using spectral based statistical moments as features. *Revista de Informática Teórica e Aplicada* 25(1):11
28. Khalighi S, Sousa T, Santos JM, Nunes U (2016) ISRUC-sleep: a comprehensive public dataset for sleep researchers. *Comput Methods Programs Biomed* 124:180–192
29. Masaaki H, Masaki K, Haruaki Y (2001) Automated sleep stage scoring by decision tree learning. In: *Proceedings of the 23rd annual EMBS international conference*, October 25–28, Istanbul, Turkey
30. Cortes C, Vapnik V (1995) Support vector networks. *Mach Learn* 20(3):273–297; Kavzoglu T, Colkesen I (2010) Classification of satellite images using decision trees: Kocaeli case. *Electron. J Map Technol* 2(1):36–45
31. Ocak H (2013) A medical decision support system based on support vector machines and the genetic algorithm for the evaluation of fetal wellbeing. *J Med Syst* 37(2):1–9
32. Osman H, Gokhan B (2017) Sleep stage classification via ensemble and conventional machine learning methods using single channel EEG signals. *Int J Intell Syst Appl Eng* 5(4):174–184
33. Bajaj V, Pachori RB (2013) Automatic classification of sleep stages based on the time-frequency image of EEG signals. *Comput Methods Programs Biomed* 112(3):320–328
34. Hsu Y-L, Yang Y-T, Wang J-S, Hsu C-Y (2013) Automatic sleep stage recurrent neural classifier using energy features of EEG signals. *Neurocomputing* 104:105–114
35. Zibrandtsen I, Kidmose P, Otto M, Ibsen J, Kjaer TW (2016) Case comparison of sleep features from ear-EEG and scalp-EEG. *Sleep Sci* 9(2):69–72
36. Berry RB, Brooks R, Gamaldo CE, Hardsim SM, Lloyd RM, Marcus CL, Vaughn BV (2014) *The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications, version 2.1*. American Academy of Sleep Medicine, Darien



37. Sim J, Wright CC (2005) The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther* 85(3):257–268
38. Liang S-F, Kuo C-E, Hu Y-H, Cheng Y-S (2012) A rule-based automatic sleep staging method. *J Neurosci Methods* 205(1):169–176