International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 11, No.1 (Jan-2022)

http://dx.doi.org/10.12785/ijcds/110159

Machine Learning Approaches with Automated Sleep Staging System based on Two-Layer Heterogeneous Ensemble Learning Stacking Model

Santosh Kumar Satapathy[1,*] and D Loganathan[1]

[1,*] *Department of CSE, Pondicherry Engineering College, Puducherry, India*

Received 2 Jan. 2021, Revised 8 Jul. 2021, Accepted 31 Dec. 2021, Published 31 Jan. 2022

Abstract: Sleep is an essential requirement for human health and well-being, but many people face sleep problems. These problems can lead to several neurological and physical disorders and adversely affect the overall quality of life. Artificial intelligence (AI)-based methods for automated sleep stage classification is a fundamental approach to evaluating and treating this public health challenge.The main contribution of this research work is to develop an Automated Sleep Staging System based on Two-Layer Heterogeneous Ensemble Learning Stacking Model (ASSS-TL-HELSM) for sleep staging under the American Academy of Sleep Medicine (AASM) sleep scoring rules. The main aim of this model is to enhance sleep staging accuracy, reduce overfitting and handle overdrift. For signal preprocessing, we use two different feature selection techniques, Fisher Score (FS), and ReliefF (ReF). For feature extraction, we obtain a total of 28 features.The proposed model analyzes the sleep behavior of the subject using the seasonal and trend components. Sleep recordings from two different subgroups of Institute of Systems and Robotics University of Coimbra (ISRUC-Sleep) were obtained for our experiments.Compared with recent studies using single-channel electro encephalogram (EEG) signals, our proposed ASSS-TL-HELSM model shows the best sleep staging classification accuracy performance on a five sleep stages classification (SC-5) task. The overall classification accuracy is 97.93%, and 97% for features selected through FS and ReF respectively, with the subgroup-I(SG-I) data; similarly, for the subgroup-III(SG-III) data, the features selected through FS, and ReF show a classification accuracy of 98.16% and 98.78% respectively. The comparisons between the proposed model and the existing model show that the proposed model gives better sleep staging accuracy for the five-sleep state's classification.

Keywords: Sleep Stages, Electroencephalography, Feature Selection, Ensemble Learning, Time Series Classification.

1. Introduction

Sleep is one of the important elements which regulates the fitness and healthiness of the human body. In this state, the subject state of mind and brain unconscious from the surrounding environment. Generally, one human being can spend its two-third timing of lifetimes as sleeping. To maintain overall health and wellbeing, it's essential to maintain the proper sleep patterns on daily basis. It has been observed that the lack of sleep patterns affects directly our physical and mental healthiness [1-2]. and mainly the physical and mental fitness depends upon their sleep quality. The regular disturbances in the sleep patterns not only affect the day-to-day activities but their impacts also visible in the later period of life such as reducing memory cognition, learning ability, and memory attention. On a long-term basis, these regular sleep problems may cause several types of sleep-related disorders. Sleep disorders not only reduce the physical movements in the daytime but also raises so many negative impacts such as cardiovascular diseases, diabetes, depression, and neurological disorders. The major step for diagnosis of different types of sleep-related disorders is studying the sleep patterns including the entire sleep cycles over the entire night and the total sleep time. In general, the normal sleep cycle continues for 90-110 minutes, which typically begins with the wake stage and gradually shifting into the deeper sleep [3]. One of the most important gold standard treatments is called as polysomnography test (PSG), which is specifically obtained during various types of sleep-related disorders. It is one of the multi-parametric tests which monitor the behavior from the different parts of the body through recording electrophysiological signals such as Electroencephalogram (EEG), Electromyogram (EMG), Electrooculogram (EOG), and Electrocardiogram (ECG) signals [4].The whole sleep staging procedure is executed under the two standard guidelines that are Rechtschaffen

E-mail:santosh.satapathy@pec.edu,drloganathan@pec.edu

and Kales (R&K) [5] and American Academy of Sleep Medicine (AASM) rules [6].As per the R&K sleep manuals, the NREM stage is further divided into four sub-sleep stages that are NREM N1, NREM N2, NREM N3, and NREM N4 stages [7-8].

In general, the entire sleep is broadly divided into two phases that is non-rapid eye movement (NREM) and rapid-eye-movement (REM). According to the R&K rules, the entire sleep cycle is segmented into seven distinct sleep stages such as Wake (Wakefulness), S1 (drowsiness), S2 (light sleep), S3 (deep sleep), and S4 (deep sleep), and REM (rapid eye movement).In the year 2007, AASM modified the R&K sleep standards and developed new sleep guidelines for sleep scoring. This major revision is reflected in NREM sleep stages, instead of four sub-sleep stages, only three stages are considered as NREM stage1 (N1), NREM stage2 (N2), and NREM stage3 (N3). According to the AASM rule NREM Stage3 (N3) is obtained with combinations of stages S3 and S4. [9].The other main challenge is variations in the sleep scoring results because of expertise in the field of the sleep behavior assessment. Therefore designing an accurate and robust sleep staging system can able to reduce the scoring time and also helps to analyze the sleep behavior accurately over the individual sleep stages. In the recent developments, there are different sleep staging studies proposed by the different researchers to automate classify the sleep stages. Most of the research work is based on the brain EEG signal from the different medical conditioned subjects. Because the EEG signal patterns provide meaningful information directly from the brain which helps to analyze the changes in sleep behavior over the individual sleep stages [10-11].

A. Contribution

The main objective of this proposed research work is to develop an automated sleep staging system using a heterogeneous ensemble learning stacking model to handle the drawbacks of the ensemble approaches that as overfitting problems, bias problems, and concept drift problems. The proposed research work analysis the sleep behavior of the subject in two ways that are seasonal and trend patterns of the sleep data. For trend modeling, we have obtained three base-layer classification model and the final predictions are average predictions of all these base-layers models, which is to be forwarded into the meta-classification layer for final predictions. Similarly, the seasonal data modeled using Classification and Regression Tree (CART).It helps to overcome the overfitting issues and finally the median of all the predictions is to be considered as the final prediction. Here we consider median value, which helps to reduce the impacts of outliers and it is not to be sensitive with regards to the major changes in the sleep behavior but the mean value which is highly sensitive with the large variations of the sleep behaviour during sleep stages.

Therefore we propose an improved ASSS-TL-HELSM which helps to overcome the drawbacks of the existing ensembling approaches and helps to improve on the sleep stages classification accuracy for five to two sleep states classification (SC-5 to SC-2) problems.

B. Significance

Sleep behavior of the subject continuously changes concerning the different time instants. During time-series predictions, we require statistical models to predict the sleep behavior patterns. These models help to understand the changes in sleep behavior of sleep and classify the sleep stages properly. However, it has been found that these models also suffered to properly identify the unexpected sleep behavior of the subject during sleep like sleep spindles and K-complexes. Sometimes the model also failure to understand the behavioral aspects of the subjects' sleep records, which are directly sometimes, influenced by the environmental factors (or) particular some sudden stress on professional social lives. This entire scenario to be treated as concept drift. Implementing the concept drift significantly to improvements in identifying the hidden relationship between sleep characteristics and its sleep stages. Though the sleep behavior is continuously changes unexpectedly concerning the time. Therefore it should highly essential to introduce the concept of drifting during sleep staging. Further ensemble techniques are widely applied in various applications for improving the sleep stages classification accuracy. This proposed model associates combinations of multiple predictions from the obtained base-layer classification models and avoids the over fitting problem. The significance of this proposed research work is to propose heterogeneous-based ensemble techniques for sleep staging classification and to validate its effectiveness on the three benchmark datasets under the AASM and R&K sleep staging rules.

C. Related Work

Most of the proposed studies rely on EEG signals. These studies recommend the extraction of features from the representative input signals. Moreover, these studies also suggest the use of feature reduction approaches for identifying the most relevant features. Finally, different classification techniques have been used to analyze EEG signals considering two to six sleep stages. Multiple studies propose automatic sleep stage classification systems based on single-channel EEG signals. It has also been observed that deep neural models are widely used in different fields of the biomedical research area. In recent research developments, it has found that notable increases happened with the use of the deep neural network in the field of biomedical signals [12-15]. Some of the recent contributions conducted by different researchers using Machine Learning (ML) and Deep

Learning (DL) models for classifying sleep stages are described below here.

Oboyya et al. [16] obtained the wavelet transform techniques for feature extraction from EEG signals and a fuzzy c-means algorithm for sleep stage classification. The reported overall classification accuracy was 85%.

In [17], the author was proposed based on a feature weighting method using K-means clustering.The selected features were forwarded to K-means and decision tree classifiers. The study reported an overall accuracy of 83%.

Aboalayon [18] also designed a sleep stage classification model based on EEG signals.The authors used a Butterworth bandpass filter to segment the EEG signals. The extracted features were used with Support Vector Machine (SVM) classifier.The work reported 90% classification accuracy.

In [19] proposed the use of bootstrap aggregation for classification.This method was applied on two benchmark datasets that are the Sleep dataset in European Data Format (EDF) and Dreams subject's datasets. The proposed system presented an accuracy of 92.43%.

In [20] the author extracted the graph properties and the experimental work executed based on the EEG signals. The proposed SVM classifier presents an average classification accuracy of 95.93%.

In [21] used PSG data for automated sleep staging. The extracted properties are classified using decision table classifiers. The reported overall accuracy was 80.70%.

Sriraam, N. et al. [22] used multi-channel EEG signals from ten healthy subjects in a proposed automatic sleep scoring model. In this study, spectral entropy features were extracted from input channels to identify irregularities in different sleep stages. The selected features were fed into the multi-layer perceptron (MLP) model and the overall accuracy of the model with 20 hidden units was reported as 92.9%. Moreover, using 40, 60, 80, and 100 hidden units, the proposed method reported 94.6%, 97.2%, 98.8%, and 99.2% accuracy, respectively.

Memar, P. et al. [23] proposed a system to classify sleep and wake stages. The authors selected 25 suspected sleep disorder subjects and 20 healthy subjects for the experimental tests. In total, 13 features were extracted from each of eight (alpha, theta, sigma, beta1, beta2, gamma1, and gamma2) sub-band epochs. The extracted features were validated using the Kruskal-Wallis test, then classified with a random forest classifier. The overall accuracy obtained was 95.31% and 86.64% with 5-Fold cross-validation and subject-wise validation respectively.

Da Silveira et al. [24] used discrete wavelet transform to decompose the signals. The random forest classifier was tested for its ability to discriminate the various sleep stages and reported an overall accuracy of 90%.

Xiaojin Li et al. [25] designed the hybrid sleep staging model using EEG signal.The overall classification accuracy of the proposed method was 85.95% using Random Forest (RF) classifier.

Zhu, G et al. [26] considered the visibility graphs and a horizontal graph to detect gait-related movements. Finally, nine features that were extracted from the input signals were forwarded to SVM classifiers considering multiple sleep stages. The proposed method presented an accuracy of 87.50% for two-state sleep stage classification problems.

Eduardo T. Braun et al. [27] developed a portable sleep staging system using a different combination of features from EEG signals. The proposed method presented the best classification accuracy of 97.1 % for the two-state classification.

In [28], the author proposed deep learning model with single-channel EEG signal. The proposed model reported 74% sleep staging classification accuracy.

Sors A et al. [29] obtained CNN model for five sleep states classification and the results reported for the proposed model is 87%.

Chambon et al. [30] proposed automated sleep staging system based on multi-modality signal fusions and K-nearest neighbhor (KNN).The accuracy performance of the proposed model reported as 80%.

In [31] the authors considered two-channels of the EEG signal and the EOG signal and one-channel of the EMG signal and the model performance reported as 83%.

Tripathy et al. [32] obtained EEG data and RR interval information using deep neural model. The reported average accuracy for NREM+REM as 95.71%, for deep sleep vs light sleep as 94.03%, and for sleep vs wake as 85.51% respectively.

Zhihong Cui et al. [33] proposed multi-modal signal data and obtained the CNN model and the model achieved an accuracy of 92.2%.

Supra Tk A et al. [34] proposed the CNN model with LSTM concept for sleep staging and the performance of the proposed scheme reported as 86.2%.

D. Structure of the paper

Section 2 describes the experimental data preparation in detail. In Section 3, we describe the proposed methodology for sleep staging evaluation. Section 4 discusses the experimental results obtained from the proposed methodology from two subgroups of subjects. In Section 5, we briefly discuss the results of our proposed methodology, as well as its advantages and limitations, and compare the results with those of state-of-the-art methods. Section 6 ends with concluding remarks and a description of future work.

2. EXPERIMENTAL DATA

A. ISRUC-Sleep Dataset

In this study, we use data from two categories of subjects who are healthy or have different medical conditions. The recorded data was collected from ISRUC-Sleep database. This dataset was prepared under the monitoring of sleep technicians in the sleep centre of Hospital of Coimbra University (CHUC) [35]. In this proposed work, we used two subgroups (SG-I/SG-III) of the ISRUC-Sleep dataset for our experimental work.The first subgroup-I of the ISRUC-sleep repository (SG-I), was contained the subjects, the most of the subjects were affected with the sleep-related diseases. In this subgroup, the recorded data from 100 subjects were considered. This data includes 55 male subjects and 45 female subjects between 20 and 85 years of age (mean age ± standard deviation, 51 ± 16 years). Most of the subjects experienced sleep apnea events, and all of the individuals were under medication. Moreover, the subjects usually breathed without the aid of any type of respiratory support machine.

The second set of recordings used in this study corresponds to the subgroup-III of the ISRUC-Sleep dataset (SG-III). This subgroup includes ten subjects (9 male and 1 female). All subjects are healthy. The data was collected in one recording session. These subjects range in age between 30 – 58 years (mean age ± standard deviation, 40 ± 10 years).The sleep behavior of the subjects with sleep problem and completely health controlled is shown in Figure 1 and Figure 2 respectively. Table I provides detailed information on the distributions of the sleep epochs into different sleep stages. Table II presents the details of the subjects participated in this proposed work.

TABLE I. DISTRIBUTION OF SLEEP EPOCHS

Subject Number/ Subgroups	W	N1	N2	N3	R
Subject-1(SG-I)	165	63	173	231	118
Subject-2 (SG-I)	231	72	226	147	74
Subject-9 (SG-I)	72	143	315	136	84
Subject-16 (SG-I)	128	125	280	120	97
Subject-1 (SG-III)	149	91	267	158	85
Subject-2 (SG-III)	89	120	274	149	118
Subject-5 (SG-III)	67	65	287	251	80
Subject-6 (SG-III)	54	111	261	247	77

TABLE II. BRIEF INFORMATION ON THE STUDY SUBJECTS

Subject	Age	Sex	Diagnosis	Disease
1	64	M	Small airways obstruction syndrome (SAOS)	Depression
2	52	M	SAOS	Restless leg syndrome
9	61	M		Cheyne-Stokes
16	50	M		No
1	30	M	HEALTHY	No
2	41	M		
5	49	F		
6	38	M		

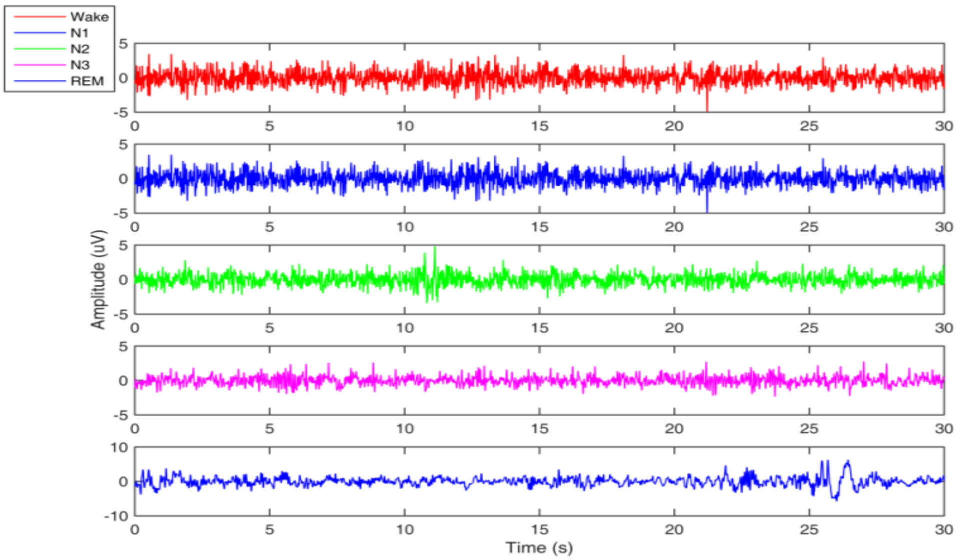

Figure 1. Sleep EEG signal: Sample recording from a subject with small airways obstruction syndrome, Subgroup-I subject-16 (session-1 recording).

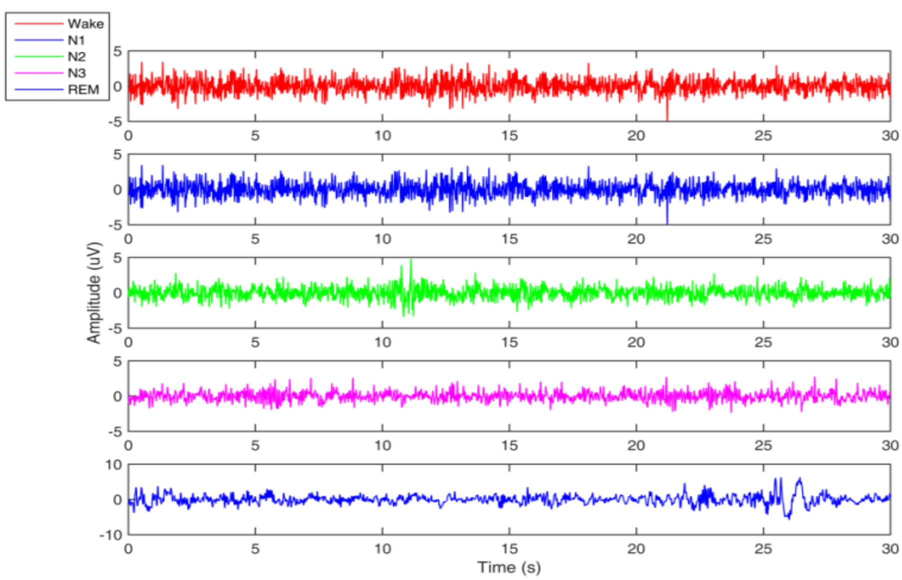

Figure 2. Sleep EEG signal: Sample from a healthy subject with no sleep problem symptoms, Subgroup-III subject-5 (session-1 recording)

3. PROPOSED HETEROGENOUS ENSEMBLE LEARNING STACKING MODEL

In this study, we proposed an Automated Sleep Staging System using Two-Layer Heterogeneous Ensemble Learning Stacking Model (ASSS-TL-HELSM).It has been seen that sleep diseases have a strong relationship between the different sleep parameters like age factor, pre-medication, professional stress, continuous changes of lifestyles, and seasonal dependencies. Therefore we develop an automat ed sleep staging system with consideration of all these dependencies which are directly or indirectly influencing sleep-related diseases. Further to monitor the effects of these dependencies during sleep, in a regular interval of time for a subject either on a monthly, quarterly, or yearly basis conduct the sleep staging. The complete workflow of this proposed model is shown in Figure 2. To observe the sleep behavior, the entire dataset is segmented into two sections that change sleep behavior concerning the trend data and seasonal data. During trend component analysis, we use the ensemble stacking model in which we have obtained the base models namely Gradient Boosting Tree (GBDT), Random Forest (RF), and eXtreme Gradient Boosting (XGBoost). Our proposed model supports the parallel structure since the base-layers models are trained independently from each other. Here the predictions of the base-layers models become the inputs for the meta-layer. Similarly for monitoring the changes in sleep behavior concerning seasonal parameters, we have obtained the bootstrapping and CART as base-layers classifiers. The output of the base-layers of the ensemble model is computed as the median of all the bootstrapped predictions. The median factor is very less prone to causes the outliers. Therefore we have used the median value for all the bootstrapped classification. The final sleep staging classification accuracy is reported with the aggregation of classification results of the seasonality and trend parameters. Finally, the meta-classifier layer provides the final sleep-staging predictions results with new testing dataset. In summary, the proposed model obtained a two-layer heterogeneous ensemble learning stacking model to improve the classification performance of the sleep staging. The complete pseudocode of the proposed model is given below here.

Algorithm 1: Pseudocode of the Heterogeneous Ensemble Learning Stacking Model.

Input: Trend Data (*TD*), Seasonal Data (*SD*)
Output: Final predictions from the proposed ASSS-TL-HELSM model
Step 1:
Obtain seasonality trend decomposition is performed on the time series sleep data X_i divided into its two components SD_i and TD_i

Step 2:
For *P=1 to M* **do**
Base layer classification models are applied into the segmented time series sleep data of the trend part $(TD_1, TD_2, TD_3 \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots TD_N)$
For *Q=1 to N* **do**
Learn the classifier PD_{MN} from the trend decomposed data
End For
Calculate the final value of the trend predictions using average of all the base models predictions

$$TP = \sum_{i=1}^{N} \frac{BaseModels_i}{N}$$

% TP=Final trend predictions results
End For

Step 3:
For *P=1 to M* **do**
Obtain the seasonal $(SD_1, SD_2, SD_3 \ldots \ldots \ldots \ldots \ldots \ldots \ldots SD_N)$ data is sampled N times randomly and applying the bootstrap techniques and generates the bootstrapping samples $(SD_1{}^*, SD_2{}^*, SD_3{}^* \ldots \ldots \ldots \ldots \ldots \ldots \ldots . SD_N{}^*)$
For *Q=1 to N* **do**
Learn the classifier PD_{MN} from the seasonal data
%N=Number of times bootstrap sample constructing
%M=Number of forecasts taking to take the final predictions
End For
Compute the final assessment predictions by taking the median of all the M predicates

$$SP = Median \sum_{i=1}^{N} \frac{BaseModels_i}{N}$$

%SP=Final seasonal predictions results
End For

Step 4:
Final predictions is computed by aggregating the result of the trend predictions and the seasonality prediction using meta-classification model
For P=1 to M **do**
For each $X_i \in TP + SP$ do
Train the meta-classifier on new instances (Xi', Yi')
End For
End For
Return $Y_i = \{Y_1, Y_2, Y_3 \ldots \ldots \ldots \ldots \ldots \ldots \ldots Y_N\}$ from meta-classifier

A. Feature Extraction from EEG data

The different existing contributions by the authors show that the importance of the different features during sleep staging classification. The feature values are estimated for each individual of epoch the input signal and discriminate the sleep stages characteristics from each epoch. According to the R&K and AASM sleep scoring rules, the general recommended length of the epoch is 20-30s. There are many ways to analyze the sleep data using different feature parameters such as linear(time-domain, frequency-domain, time-frequency domain) and non-linear features.

1) Time-domain features

Generally, the features of this category are directly retrieved from the signal itself.From each sleep stages epoch, we extracted the statistical features and morpho logical features.The statistical attributes help to analyze the signal distribution statistics, similarly, the morphological properties were used to measure the waveform behavior of the entire segment. One other important feature, which has widely used during time-series analysis, that is Hjorth parameters.

2) Frequency-domain features

The Sleep behavior of the subjects are continuous changes in the different sleep stages with concerns to the changes in frequency levels and it's highly important to identify the irregularities of the sleep during sleep stages for sleep staging. Though EEG signals are compositions of the different signal sub-bands (α, θ, δ, β, γ) and all the changes characteristics of this signal sub-bands need to be analyzed by extracting the time and frequency domain features. The main aspects to use the frequency-domain parameters are to describe the changes sleep behavior with respect to the individual sleep stages.In this research work, we have extracted a set of 13 time-domain and 15 frequency-domain features.The details about the extracted features from the input signals are described in Table III.

TABLE III. EXTRACTED FEATURES

Time-domain		
Feature No.	**Extracted Feature**	**Feature Descriptions**
1	$Mean(\bar{x}) = \frac{1}{N}\sum_{i=1}^{N} x_i$	The mean electrical potential of an epoch is calculated.
2	$Maximum = \mathrm{Max}[x_i]$	It is used to quantify the range of data and it helps to find the magnitude of signal baseline.
3	$Minimum = \min[x_i]$	
4	$Variance(Var) = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N-1}$	It helps to determine how the data is dispersion with respect to the value of the mean.
5	$Standrad\ deviation(SD) = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\bar{x})^2\right)^2$	It provides the information regarding how far the signal fluctuates from the mean.
6	$Median = \left(\frac{N+1}{2}\right)^{th}$	It helps to get the information about the centre and spread of the signal data. The median is the value separating the upper half set from the lower half of the set.
7	$75^{th} percentile(75_P^{th}) = Max(x_i) < P\{75\}$ x_i is the signal data $p\{75\}$ = the 75^{th} percentile of the signal data	Mainly it provides the information about the data distribution of the values. It defines the value below which 75% of the random variables values data is located
8	$Signal\ Skewness(skew) = \sum_{i=1}^{N}\frac{(E(x_i-\bar{x})^3)}{\sigma^3}$ E is the expected mean value $E(x) = \sum_{i=1}^{N} p_i x_i$	The skewness (3^{rd} order) and kurtosis (4^{th} order) were helps to measure the central tendency, peakedness, asymmetry and degree of dispersion of the EEG signals distribution.
9	$Signal\ kurtosis(kurt) = \sum_{i=1}^{N}\frac{(E(x_i-\bar{x})^4)}{\sigma^4}$	
10	$Signal\ Activity = Var(x_i)$	It provides the dynamic temporal information of the polysomnography signals. These parameters are computed based on the variance of the derivatives of the signal data.
11	$Signal\ Mobility = \sqrt{Var(x_i{}')/Var(x_i)}$	
12	$Signal\ Complexity = \sqrt{Var(x_i{}'')\ X\ Var(x_i)/Var(x_i{}')^2}$	

13	$Zerocrossing\ Rate(ZC) = \sum_{i=1}^{N} ZC'(i)$	It helps to count the number of times the EEG signal crosses the reference lines. Additionally it also effective towards characterization of the sleep spindles.
Frequency-domain		
14-17	$Relative\ Spectral\ Power(\delta,\theta,\alpha,\beta) = \frac{\int_{-f1}^{f0} x_i(f)df + \int_{f0}^{f1} x_i(f)df}{\int_{-\infty}^{\infty} x_i(f)df}$	Mainly it measures the distribution of power into different frequency sub-bands.
18-24	$Power\ Ratios = \frac{\delta}{\beta}, \frac{\delta}{\theta}, \frac{\theta}{\alpha}, \frac{\theta}{\beta}, \frac{\alpha}{\beta}, \frac{(\theta+\alpha)}{(\alpha+\beta)}$	It is used for computing the power from the current and background of the epochs with the different frequency ranges for interpreting the behavior of the sleep stages.
25-28	$Power\ spectrum(\alpha) = \sum_{i=1}^{N} x_{i(\alpha,\beta,\theta,\delta)}$	It helps to retrieve the information on how the intensity of a time-series signal data is distributed in the frequency domain.

B. Feature Selection

This section also equally important during sleep staging because it has been noticed that all the extracted features may not be effective during classification process. As we have discussed earlier that the sleep stages characteristics are highly dynamic in nature. To discriminate the sleep patterns, an appropriate set of features were required. To identifying the set relevant features different feature reduction techniques obtained by the different researchers. From the above feature extraction, it has been found that each extracted features present different sleep stages information's for the example power spectrum of the sigma band captured information about shallow sleep characteristics in the N2 stage according to AASM, and in the same way it has found for S2 stage according to R&K scoring rules. Similarly, the beta spectrum band identified the high-frequency wave-forms, which occurred during the Wake and REM sleep stages. Hjorth parameters obtained the changes in characteristics during the N3 and N4 sleep stages. So it is more important is what certain number of features to be used further during sleep scoring. The second most important thing is that what combinations of the features could provide the entire full description with regards to changes in sleep characteristics during sleep hours without having any types of redundancy.

1) ReliefF(ReF) algorithm

This section also equally important during sleep staging because it has been noticed that all the extracted features may not be effective during classification process. As we have discussed earlier that the sleep stages characteristics are highly dynamic in nature. To discriminate the sleep patterns, an appropriate set of features were required. To identifying the set relevant features different feature reduction techniques obtained by the different researchers. In this study, we have considered the ReliefF feature selection algorithm used to sort the features according to its individual weight factor. At first it identifies the individuals randomly and after that it finds the similar class from its nearest neighbhor and on the other hand separates the non-similar class from the nearest neighbhor individuals. In each round the weight of the feature updated. At last we selected as the best feature which feature scored the higher weights (maximal relevance) [36].

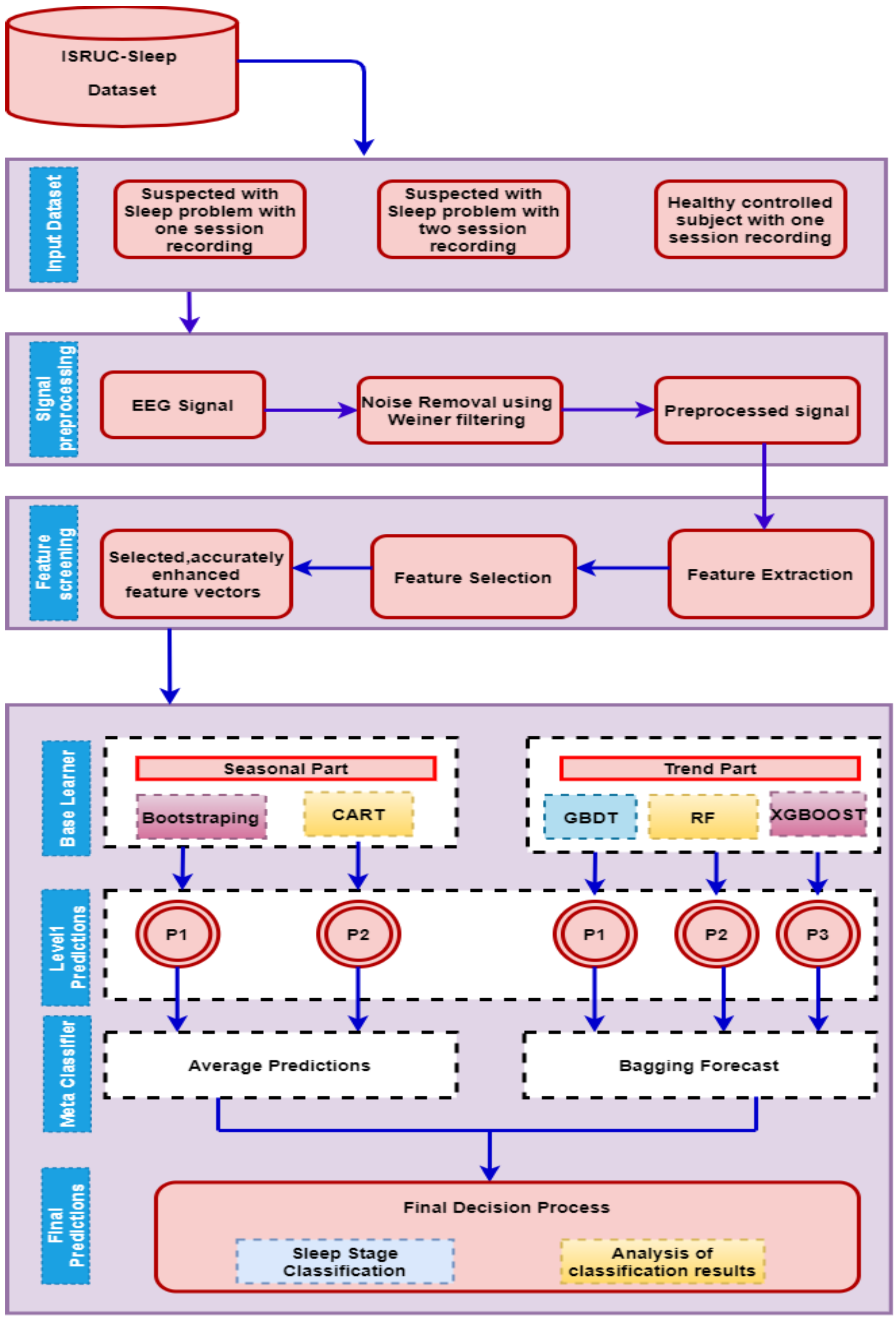

Figure 2. The complete workflow of the proposed ASSS-TL-HELSM

2) Fisher(FS) Score

This algorithm assigned the ranking against each feature which can be used to identify the suitable features using feature vectors and eliminate redundant features. The main purpose of the Fisher score algorithm is to find common correlation criteria. The feature scoring algorithm first compares the Fisher score of each feature (F) and then considers the threshold value as θ. If F >θ, the feature is selected; otherwise, it is not selected. This algorithm is very effective in terms of feature ranking because of its simple structure [37].

C. Classification Evaluation Metrics

To analysis the effectiveness of the proposed sleep staging methodology, we have obtained certain performa nce metrics such as accuracy [38], recall [39], specificity [40], Precision [41] and F1score [42].

$$Accuracy = \frac{True_{Po} + True_{Ne}}{True_{Po} + True_{Ne} + False_{Po} + False_{Ne}} \% \quad (1)$$

$$Recall = \frac{True_{Po}}{True_{Po} + False_{Ne}} \% \quad (2)$$

$$Specificity = \frac{True_{Ne}}{True_{Ne} + False_{Po}} \% \quad (3)$$

$$Precision = \frac{True_{po}}{True_{Po} + False_{Po}} \% \quad (4)$$

Where:

$True_{Po}$ = No's of True +ve
$True_{Ne}$ = No's of True -ve
$False_{Po}$ = No's of False +ve
$False_{Ne}$ = No's of False -ve

$$F1\ Score = \frac{2 * Recall * Precision}{Recall + Precision} \quad (5)$$

4. EXPERIMENTAL RESULTS

The proposed method is implemented using MATLAB R2015a with windows 10 operating system. To evaluate the performance of the proposed method, a set of experiments were conducted using two different categories of sleep data of ISRUC-Sleep dataset. The first experiment was conducted using data from subjects who were affected with different types of sleep problems and who had undergone a single sleep recording session. The second experiment was conducted using data from the subjects who were completely healthy and no history of any diseases related to sleep. In this research, we used only C3-A2 single-channel EEG signals to examine the subjects' sleep behavior.

According to recent sleep staging studies, the C3-A2 channel is the most useful in terms of classification accuracy because it provides central information that represents the brain's behavior during sleep [42-57]. In other words, it conveys a good deal of information from the central part of the brain. Thus, we used C3-A2 channel recordings in the present work.

A. Feature Selection Results

The extracted features were screened through two different feature selection techniques, FS and ReF.The selected features are sorted according to feature weight (weights from high to low).The final selected features using FS and ReF are presented in Tables IV and V using SG-I and SG-III sleep data respectively.

TABLE IV. SELECTED FEATURES WITH SG-I DATA

Screening Algorithm	Final Suitable Features
FS	18,8,13,10,5,15,17,4,1,14,9,11,7,2, 19,23,16,24,22,26,12,25,3,21,20,6, 28,27
ReF	8,18,5,14,10,9,1,4,13,15,7,6,3,11,2,22,16, 24,17,26,27,25,21,28,23,19,12,20

TABLE V. SELECTED FEATURES: RESULTS WITH SG-III DATA

Selection Algorithm	Final Suitable Features
FS	1,17,16,24,22,2,15,14,3,10,28,26,6,11,21, 20,27,19,23,25
ReF	13,18,8,12,1,5,7,4,9,17,15,2,10,14,3,16,24, 22,6,28,11,23,19,20,27,21,25,26

B. *Performance of Sleep staging performance using the SG-I data*

In this experiment, we considered the single-channel C3-A2 EEG signal recordings of four sleep-disordered subjects. All four subjects received a diagnosis of sleep syndrome disease, and some of the subjects also suffered from insomnia and other types of sleep problems. The length of each epoch is 30s. A total of 3000 epochs with 6000 sample points were obtained from the four subjects, along with their sleep stage class labels. The recorded signals are cleaned using Butterworth band-pass filter of order 10. Then, the FS and ReF feature screening techniques were used to screen the relevant features. Finally the screened were forwarded into the three base layer classifiers, RF, GBDT, and XGBoost, to produce the initial sleep staging predictions. The Classification

Accuracies (CAs) achieved for two–five sleep stages classification (SC-2 to SC-5) using the RF, GBDT and XGBoost classifiers are presented in Tables VI, IX, and XII.

TABLE VI. ACCURACY OF THE RANDOM FOREST CLASSIFIER WITH THE SG-I DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	99.22%	98.01%	97.86%	97.46%
ReF	99.01%	97.92%	97.79%	97.3%

As shown in Table VI, the highest classification accuracy was reported for SC-2 (99.22%), SC-3 (98.01%), SC-4 (97.86%) and SC-5 (97.46%) with FS selected features. The confusion matrix results for the SC-5 task using the RF classification model for the two feature selection algorithms FS and ReF are shown in Tables VII, and VIII.

TABLE VII. CONFUSION MATRIX WITH FS-SELECTED FEATURES USING THE RF CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	621	11	10	10	5
N1(1)	6	458	4	0	3
N2(2)	7	1	1052	3	1
N3(3)	4	1	2	478	0
R(5)	0	2	5	1	315

TABLE VIII. CONFUSION MATRIX WITH RF-SELECTED FEATURES USING THE RF CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R(5)
W(0)	621	11	10	10	5
N1(1)	6	458	4	0	3
N2(2)	7	1	1052	3	1
N3(3)	4	1	7	473	0
R(5)	0	2	5	1	315

TABLE IX. ACCURACY OF THE GBDT CLASSIFIER WITH THE SG-I DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	98.32%	97.91%	97.12%	95.2%
ReF	98.20%	97.79%	96.22%	94.86%

As shown in Table IX, the highest classification accuracy was reported for SC-2 (98.32%) with FS, SC-3 (97.91%) with FS, SC-4 (97.12%) with FS and SC-5 (97.46%) with FS. The confusion matrix results for the CT-5 task using the GBDT classifier are shown in Tables X and XI.

TABLE X.CONFUSION MATRIX WITH FS-SELECTED FEATURES USING THE GBDT CLASSIFIER

True/Predicated	W(0)	N1 (1)	N2 (2)	N3 (3)	R(5)
W(0)	601	31	10	10	5
N1(1)	6	458	4	0	3
N2(2)	7	1	1032	13	11
N3(3)	4	11	20	450	0
R(5)	0	2	5	1	315

TABLE XI.CONFUSION MATRIX WITH REF-SELECTED FEATURES USING THE GBDT CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	589	27	13	11	7
N1(1)	6	456	4	2	3
N2(2)	7	1	1052	3	1
N3(3)	4	21	40	420	0
R(5)	0	2	5	1	315

TABLE XII.ACCURACY OF THE XGBOOST CLASSIFIER WITH THE SG-I DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	98.72%	97.87%	96.86%	92.25%
ReF	98.31%	97.92%	96.79%	93.2%

From Table XII, it has been found that the highest classification accuracy was reported for the XGBoost classifier using two selected features from FS, and ReF for the SC-2 to SC-5 classification problems. The highest classification accuracy was reported for SC-2 (98.72%) with FS, SC-3 (97.92%) with ReF, SC-4 (96.86%) with FS and SC-5 (93.2%) with ReF. The confusion matrix results for the SC-5 task are shown in Tables XIII and XIV.

TABLE XIII.CONFUSION MATRIX WITH FS-SELECTED FEATURES USING THE XGBOOST CLASSIFIER

True/Predicated	W(0)	N1 (1)	N2 (2)	N3 (3)	R(5)
W(0)	591	51	20	12	3
N1(1)	14	450	4	0	3
N2(2)	7	21	1000	15	21
N3(3)	4	11	20	430	21
R(5)	0	2	5	1	315

TABLE XIV.CONFUSION MATRIX WITH REF-SELECTED FEATURES USING THE XGBOOST CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	581	51	10	10	5
N1(1)	14	450	4	0	3
N2(2)	7	21	1000	15	21
N3(3)	4	11	20	450	0
R(5)	0	2	5	1	315

C. Sleep staging performance using the SG-III data

This experiment looked at recordings from, four healthy control subjects who were not affected by any type of sleep-related problems. The same methodology was implemented for sleep staging with this dataset. The same properties were extracted to discriminate between the sleep classes. The properties selected using the FS and ReF are forwarded into the three base classifiers and the sleep staging performances for SC-2 to SC-5 are shown in Tables XV, XVIII and XXI respectively.

TABLE XV. ACCURACY OF THE REF CLASSIFIER WITH THE SG-III DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	99.52%	98.87%	97.36%	98.13%
ReF	99.11%	98.92%	96.79%	97.2%

The highest accuracy was 99.52% with FS, 98.92% with ReF, 97.36% with FS, and 98.13% with FS for the SC-2 to SC-5 classification problems. The confusion matrix for SC-5 is shown in Tables XVI and XVII for features selected by the FS, and ReF algorithms, respectively.

TABLE XVI. CONFUSION MATRIX WITH THE SG-III DATASET USING FS-SELECTED FEATURES WITH THE RF CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	614	0	1	4	0
N1(1)	4	380	4	1	0
N2(2)	3	3	1034	3	3
N3(3)	3	3	8	580	1
R(5)	1	2	8	4	336

TABLE XVII. CONFUSION MATRIX WITH THE SG-III DATASET USING REF-SELECTED FEATURES WITH THE RF CLASSIFIER

True/Predicated	W(0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	610	4	1	4	0
N1(1)	4	360	14	11	0
N2(2)	3	3	1030	7	3
N3(3)	3	3	8	580	1
R(5)	1	2	8	4	336

TABLE XVIII. ACCURACY OF THE GBDT CLASSIFIER WITH THE SG-III DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	98.72%	98.07%	97.56%	94.00%
ReF	98.21%	97.92%	96.69%	95.56%

From Table XVIII, the highest classification accuracy was: SC-2 (98.72%) with FS, SC-3 (98.07%) with FS, SC-4 (97.56%) with FS, and SC-5 (95.56%) with ReF.The confusion matrix for the three feature selection algorithms is presented in Tables XIX and XX.

TABLE XIX. CONFUSION MATRIX WITH THE SG-III DATASET USING FS- SELECTED FEATURES WITH THE GBDT CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	587	8	6	4	4
N1(1)	12	320	14	21	22
N2(2)	6	13	1010	7	10
N3(3)	3	3	8	572	9
R(5)	1	2	8	4	336

TABLE XX. CONFUSION MATRIX WITH THE SG-III DATASET USING REF-SELECTED FEATURES WITH THE GBDT CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R(5)
W(0)	601	9	5	4	0
N1(1)	12	320	14	21	22
N2(2)	3	3	1030	7	3
N3(3)	3	3	8	580	1
R(5)	1	2	8	4	336

TABLE XXI. ACCURACY OF THE XGBOOST CLASSIFIER WITH THE SG-III DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	98.82%	97.07%	95.56%	93.50%
ReF	98.27%	98.02%	96.89%	94.03%

The XGBoost classifiers achieved the following highest accuracy scores: SC-2, 98.82% using FS, SC-3, 98.02% using ReF, SC-4, 96.89% using ReF, and SC-5, 94.03% using ReF.The confusion matrix for the SC-5 task using the FS and ReF algorithms is shown in Tables XXII and XXIII.

TABLE XXII. CONFUSION MATRIX WITH THE SG-III DATASET USING FS-SELECTED FEATURES WITH THE XGBOOST CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	609	9	5	4	8
N1(1)	27	305	20	15	22
N2(2)	15	16	990	17	8
N3(3)	3	3	8	580	1
R(5)	1	2	8	4	336

TABLE XXIII. CONFUSION MATRIX WITH THE SG-III DATASET USING REF-SELECTED FEATURES WITH THE XGBOOST CLASSIFIER

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R(5)
W(0)	609	9	5	4	8
N1(1)	27	300	16	15	22
N2(2)	8	10	996	12	13
N3(3)	3	3	8	580	1
R(5)	1	2	8	4	336

D. Sleep staging performance using the proposed ASSS-TL-HELSM algorithm with subgroup-I/III datasets

We applied the same feature selection technique combinations, which we finally obtained by performing the feature selection analysis.The classification results of this proposed ASSS-TL-HELSM using the SG-I and SG-III data for the SC-5 problem are presented in Tables XXIV and XXVIII respectively.

TABLE XXIV.ACCURACY OF THE ASSS-TL-HELSM CLASSIFIER WITH THE SG-I DATASET

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	99.07%	98.02%	97.98%	97.93%
ReF	98.12%	97.42%	96.68%	97.00%

The proposed ensemble classification model showed the highest classification accuracy for SC-2 (99.07%), SC-3 (98.02%), SC-4 (97.88%), and SC-5 (97.93%) using FS algorithm-selected features. The confusion matrix for the SC-5 problem is presented in Tables XXIV and XXV for ASSS-TL-HELSM using FS and ReF selected features respectively.The reported performance metrics results with SG-I and SG-III data using ASSS-TL-HELSM presented in Tables XXV and XXVI respectively.

TABLE XXV. CONFUSION MATRIX WITH THE SG-I DATASET USING FS-SELECTED FEATURES WITH THE ASSS-TL-HELSM

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	640	4	8	2	3
N1(1)	6	458	4	0	3
N2(2)	7	1	1052	3	1
N3(3)	4	1	7	473	0
R(5)	0	2	5	1	315

TABLE XXVI. CONFUSION MATRIX WITH THE SG-I DATASET USING REF-SELECTED FEATURES WITH THE ASSS-TL-HELSM

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	629	4	8	7	9
N1(1)	5	452	4	6	4
N2(2)	7	1	1049	3	4
N3(3)	4	1	7	465	8
R(5)	0	2	5	1	315

TABLE XXVII.PERFORMANCE EVALUATION WITH INPUT FROM THE SG-I DATASET USING THE ASSS-TL-HELSM

Performance Metrics	SG-I Data (SC-5)	
	FS	ReF
Accuracy	97.93%	97%
Precision	98.34%	96.66%
Sensitivity	97.71%	96.74%
F1-Score	98.03%	96.68%

As shown in Table XXVII, the highest classification accuracy of the proposed model was 97.93% with the FS algorithm.Similarly, the best performance in terms of precision (98.34%), recall (97.71%), and F1-score (98.03%) were reported with FS selected features

TABLE XXVIII.ACCURACY OF THE STACKING MODEL CLASSIFIER WITH THE SG-III DATA

Feature Selection Algorithm	SC-2	SC-3	SC-4	SC-5
FS	99.67%	99.02%	98.98%	98.16%
ReF	99.42%	99.22%	98.99%	98.78%

The proposed stacking model achieved the highest classification accuracy of: SC-2, 99.07%, SC-3, 98.02%; SC-4, 97.98%; and SC-5, 97.93% with input from the SG-III dataset. The confusion matrix for the SC-5 task is presented in Tables XXIX and XXX.

TABLE XXIX.CONFUSION MATRIX WITH THE SG-III DATA USING FS-SELECTED FEATURES WITH THE ASSS-TL-HELSM

True/Predicated	W (0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	605	5	1	3	5
N1(1)	4	381	3	1	0
N2(2)	3	3	1034	3	3
N3(3)	3	3	3	585	1
R(5)	1	2	4	4	340

TABLE XXX. CONFUSION MATRIX WITH THE SG-III DATASET USING REF-SELECTED FEATURES WITH THE ASSS-TL-HELSM

True/Predicated	W(0)	N1 (1)	N2 (2)	N3 (3)	R (5)
W(0)	605	5	1	3	5
N1(1)	4	381	3	1	0
N2(2)	5	8	1020	6	7
N3(3)	3	3	3	583	3
R(5)	1	2	4	6	333

TABLE XXXI. PERFORMANCE EVALUATION RESULTS WITH INPUT FROM THE SG-I DATASET USING THE ASSS-TL-HELSM

Performance Metrics	SG-III Data(SC-5)	
	FS	**ReF**
Accuracy	98.16%	98.78%
Precision	97.89%	96.97%
Sensitivity	97.94%	97.48%
F1-Score	97.91%	97.22%

As shown in Table XXXI, the proposed ensemble learning stacking model performed excellently on the five sleep state classification problem compared to the base layer classification algorithms for healthy control subjects. The proposed stacking model showed a classification accuracy of 98.78% using the ReF algorithm.Similarly, the highest precision (97.89%), sensitivity (97.94%), and F1-Score (97.91%) were achieved using the Fisher score algorithm.

E. Summary of Results

In this paper, we used two subgroups (SG-I/SG-III) of the ISRUC-Sleep dataset for automatic sleep staging classification. Four different sleep classification (SC-2 to SC-5) problems were created from these two subgroups. All experiments in this study used single-channel EEG signals. The first three experiments were conducted based on base layer learning classification models, whereas the final experiment was executed with the proposed ASSS-TL-HELSM using two feature selection techniques FS, and ReF. Table XXXII presents a summary of the results that were obtained through the different base layer classification models and the ensemble learning stacking model with the two different categories of subjects and the two feature selection techniques.

TABLE XXXII. CLASSIFICATION ACCURACY OF VARIOUS CLASSIFICATION MODELS AND SUBJECT SUBGROUPS

Data Set	Classification Model	FS	ReF
SG-I	RF	97.16%	97.3%
	GBDT	95.2%	94.80%
	XGBoost	92.25%	93.2%
	Proposed ASSS-TL-HELSM Model	**97.93%**	**97%**
SG-III	RF	98.13%	97.2%
	GBDT	94.00%	95.56%
	XGBoost	93.50%	94.03%
	Proposed ASSS-TL-HELSM Model	**98.16%**	**98.78%**

The proposed ASSS-TL-HELSM reported the highest classification accuracy of 97.93% using FS and 98.78% using ReF selected features for the SC-5 problem with the SG-I and SG-III data of ISRUC-Sleep database.

5. DISCUSSION

An efficient and improved ASSS-TL-HELSM based automated sleep scoring method is suggested for the classification of multiple sleep stages.Several experiment were executed on three datasets to validate the efficacy of the proposed methodology.The proposed sleep staging methodology can automatically learn high-level information directly from single-channel EEG signals.

A. Hypothesis and limitations of the proposed model

The proposed multi-sleep staging results indicate that the ASSS-TL-HELSM showed improved classification accuracy compared to the earlier automated sleep staging and other ML methods. In comparison with other sleep staging classification problems, we mainly focused on three things which helped to improve the classification accuracy of sleep staging: 1.Analysis the trend and seasonal sleep behavior of the subjects, 2. Feature screening using FS and ReF and 3. Ensemble learning techniques. A major advantage of the proposed model is that it was developed using two different groups of subjects with different sleep recordings. Our proposed methodology was executed with 6,000 epochs, each 30 s in length, for SC-2 to SC-5 classification with the two subgroups of sleep recordings.

We designed a heterogeneous-based ensemble learning stacking model which combines two layers for classificat tion.The first layer is considered the base learning layer and the second layer is called the meta-learning layer. The final decisions are obtained from the meta-learning classification model. Despite the improved sleep staging performance; the proposed methodology has some limitations.

First, the data displayed unequal distributions of sleep stage epochs in various sleep stages. Each sleep stage lasts for a different amount of time, and the epochs of each stage are also not the same. This sleep class imbalance limits the classification accuracy of sleep staging. Another problem relates to features. Different extracted features from the signal have different influences on sleep scoring, which may create problems for performance. Despite these limitations; our proposed framework achieved very good accuracy in comparison to other works. Overcoming the class imbalance problem in future may help us reach classification accuracy closer to 100%. Also, we plan to use deep learning techniques, in which automated feature learning will help overcome differences in feature influence, thus further increasing sleep staging classification accuracy.
The class imbalance problem can be solved by implementing data augmentation concepts using deep learning techniques. We compared the performance of the propose d system with that of other available state-of-the-art classification systems. To do so, we selected studies that used similar datasets and single-channel recordings.Table XXXIII compares the features used in the present work to others in related works, all of which relied on single-channel EEG signals from the ISRUC-Sleep dataset t. These comparisons must take into account the use of single-channel EEG signals. This study shows that the proposed ASSS-TL-HELSM model performed well in comparable to the existing contributions for SC-5 classification problem using single-channel EEG signals supported by machine learning techniques. This study validates the effectiveness and usefulness of the proposed method by presenting a detailed comparative analysis with similar research proposals available in the literature. The classification results of the proposed work were compared to those of 8 similar studies that rely on the same datasets, and were shown to be superior.

TABLE XXXIII. COMPARISON OF THE CAs (%) INBETWEEN PROPOSED METHODOLOGY AND STATE-OF-THE-ART CONTRIBUTIONS

Author	**Classification Techniques**	**Data Selection**	**Accuracy**
Hugo Simoes et al.[58] 2010	Bayesian Classifier	10-fold cross validation	83.00%
Khalighi et al.[59] 2011	SVM	LOOCV	95.00%
Khalighi et al.[60] 2013			81.74%
Sousa et al. [61] 2015			86.75%
Khalighi et al.[62] 2016			93.97%
Hassan, A.R et al.[63] 2016	Bagging	10-fold cross validation	90.69%
Da Silveira et al.[24] 2016	RF	10-fold cross validation	91.50%
Najdi et al.[64] 2017	Stacked Sparse Auto-Encoders (SSAE)	10-fold cross validation	82.30%
KD Tzimourta et al.[65] 2018	Random Forest	10-fold cross validation	75.29%
Kalbkhani, H. et al. [66] 2018	SVM	10-fold cross validation	83.33%
Memar, P. et al.[23] 2018	RF	5-fold cross validation	95.31%
Huang, W. et al.[67] 2019	SVM	10-fold cross validation	92.04%
Wang, Q. et al.[68] 2019	Stacking Model	5-fold cross validation	96.6%
Sharma, M. et al.[69] 2019	SVM	10-fold cross validation	91.5%
Dhok, S. et al.[70] 2020	SVM	10-fold cross validation	87.45%
Santaji, S. et al.[71] 2020	RF	10-fold cross validation	
Zhou, J. et al.[72] 2020	Ensemble Stacking Model	5-fold cross validation	91.20%
Proposed	**Proposed ASSS-TL-HELSM (SG-I)**	**10-fold cross validation**	**Fisher Score (FS)**
			97.93%
			ReliefF (ReF)
			97.00%
	Proposed ASSS-TL-HELSM (SG-III)	**10-fold cross validation**	**Fisher Score(FS)**
			98.16%
			ReliefF (ReF)
			98.78%

6. CONCLUSIONS

This paper presents an improved ASSS-TL-HELSM for multiple sleep stage classification using single-channel EEG signals according to AASM sleep standards. This proposed methodology has been analyzed using 6000 epochs from two different datasets.This study makes four important contributions. First of all the sleep behavior is analyzed by obtaining the trend and seasonal components, which overcomes the earlier drawbacks in the existing ensemble technique.Secondly the multi-feature extraction being more helpful to analysis the changes sleep behavior during the different stages of the sleep. Thirdly, the proposed work used feature screening techniques, which are directly applicable to identifying the most relevant features from extracted feature vectors. Finally the proposed ASSS-TL-HELSM classifying the sleep stages by obtaining the two-layer stacking model, which overcomes the overfitting problem and also produced the improved sleep staging classification accuracy. The results of this study show that the method provides an effective mechanism for handling data from subjects with different health conditions with high accuracy. We also compared the proposed research with similar studies from the literature and showed that the proposed model had better performance.

REFERENCES

[1] L.A. Panossian, A.Y. Avidan, “Review of Sleep Disor ders,” Med. Clin. N. Am., 93,pp.407–425, 2009

[2] A. Smaldone, J.C. Honig, M.W. Byrne, “Sleepless in America: inadequate sleep and relationships to health and well-being of our nation’s children,” Pediatrics, 119, 29–37, 2007.

[3] A.R. Hassan, M.I. Hassan Bhuiyan, “Automatic sleep scoring using statistical features in the EMD domain and ensemble methods,” Biocybern. Biomed. Eng. 36(1), pp.248–255, 2016.

[4] Khald. Aboalayon, Helen. Ocbagabir, and Miad. Faezipour, “Efficient Sleep Stage Classification Based on EEG Signals,” 2014 IEEE Long Island Systems, Applications and Technology Conference, LISAT,pp.1-62014.

[5] A. Rechtschaffen, A. Kales, “A Manual of Standardi zed Terminology, Techniques and Scoring Systems for Sleep Stages of Human Subjects”. U.G.P. Office, Public Health Service; Washington, DC, USA 1968.

[6] C. Iber, S. Ancoli-Israel, A.L. Chesson, S.F. Quan, “The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specification”. American Academy of Sleep Medicine 32(2), pp.139-149, 2007.

[7] M. T. Bianchi, “Sleep Devices: wearables and nearables, informational and interventional”. Consumer and clinical Metabolism Clinical & Experimental, 84:99-108, 2018.

[8] D. Cogan, J. Birjandtalab, M. Nourani, J. Harvey, V. Nagaraddi, “Multi-Bio signal Analysis for Epileptic Seizure Monitoring,” Int. J. Neural Syst.,27(1),pp.16500 31, 2017

[9] M. Obayya, F. AbouChadi, “Automatic classification of sleep stages using EEG records based on Fuzzy c-means (FCM) algorithm”. Radio Science Conference (NRSC), pp.265-272, 2014.

[10] E. Alickovic and A. Subasi, "Ensemble SVM Method for Automatic Sleep Stage Classification," in IEEE Transactions on Instrumentation and Measurement, vol. 67, no. 6, pp.1258-1265,2018,

[11] A. Prakash, V. Roy, “An Automatic Detection of Sleep using Different Statistical Parameters of Single Channel EEG Signals”. International Journal of Signal Processing, Image Processing and Pattern Recognition, 9(10), pp.365–374,2017.

[12] M. Rungruanganukul, T. Siriborvornratanakul, “Deep Learning Based Gesture Classification for Hand Physical Therapy Interactive Program,”HCI (18) pp.349-358,2020

[13] Chutisant. Kerdvibulvech, “Recent Multimodal Communication Methodologies in Phonology, Vision, and Touch,”HCI (2),392-400, 2020.

[14] T. Siriborvornratanakul, “An Automatic Road Distress Visual Inspection System Using an Onboard In-Car Camera”. Advances in Multimedia, 2018, pp.1–10.

[15] Ghimatgar, H., Kazemi, K., Helfroush, M. S., & Aarabi, A.An automatic single-channel EEG-based sleep stage scoring method based on hidden Markov model. Journal of Neuroscience Methods, 108320,2019.

[16] M. Obayya. Abou-Chadi, “Automatic classification of sleep stages using EEG records based on Fuzzy c-means (FCM) algorithm”. Radio Science Conference (NRSC), pp. 265-272, 2014.

[17] K. Polat. Güneş, Ş. Yosunkaya, “Efficient sleep stage recognition system based on EEG signal using k-means clustering based feature weighting,” Expert Systems with Applications, 37, pp.7922-7928, 2010.

[18] K. Aboalayon, H. T. Ocbagabir, M. Faezipour, “Efficient sleep stage classification based on EEG signals”. Systems,Applications and Technology Confer ence (LISAT), pp.1-6, 2014.

[19] A. R. Hassan, A. Subasi, “A decision support system for automated identification of sleep stages from single-channel EEG signals,” Knowledge-Based Systems, 128, 115–124,2017.

[20] M. Diykh, Y. Li, P. Wen, “EEG Sleep Stages Classification Based on Time Do-main Features and Structural Graph Similarity,” IEEE Transactions on Neural Systems and Rehabilitation Engineering 24 (11), pp.1159–1168, 2016

[21] K. M. Gunnarsdottir,C. E.Gamaldo, R. M. E. Salas, J. B. Ewen, R. P. Allen, S. V. Sarma, A Novel Sleep Stage Scoring System: Combining Expert-Based Rules with a Decision Tree Classifier. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2018.

[22] N. Sriraam, T. K. Padma Shri, U. Maheshwari, “Recognition of wake-sleep stage 1 multichannel EEG patterns using spectral entropy features for drowsiness detection,”Australasian Physical & Engineering Sciences in Medicine,39(3), pp.797–806, 2018.

[23] P. Memar, F. Faradji, “A Novel Multi-Class EEG-Based Sleep Stage Classification System,” IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(1), pp.84–95,2018.

[24] T. L. T. Da Silveira, A. J. Kozakevicius, C. R. Rodrigues, “Single-channel EEG sleep stage classification based on a streamlined set of statistical features in wavelet domain,” Medical & Biological Engineering & Computing, 55(2), pp.343–352,2016.

[25] B. Wutzl, K. Leibnitz, F. Rattay, M. Kronbichler, M. Murata, “Genetic algorithms for feature selection when classifying severe chronic disorders of consciousness,” PLOS ONE, 14(7), 2019.

[26] G. Zhu, Y. Li, and P. P. Wen, “Analysis and Classification of Sleep Stages Based on Difference Visibility Graphs From a Single-Channel EEG Signal,” IEEE Journal of Biomedical and Health Informatics 2014, 18(6), pp.1813–1821,2014.

[27] E. T. Braun, A. D. J. Kozakevicius, T. L. T. Da Silveira, C. R. Rodrigues, G.Baratto, “Sleep Stages Classification Using Spectral Based Statistical Moments as Features,”Revista de Informática Teórica e Aplicada, 25(1),2018.

[28] O.Tsinalis, PM. Matthews, Y. Guo, S. Zafeiriou “Automatic sleep stage scoring with single-channel eeg using convolutional neural networks,”2016,arXiv preprint arXiv:1610.01683.

[29] A. Sors, S. Bonnet, S. Mirek, L. Vercueil, Payen J-F. A convolutional neural network for sleep stage scoring from raw single-channel eeg. Biomed Signal Process Control,42,pp.107–14,2018.

[30] S. Chambon, M. N. Galtier, P. J. Arnal, G. Wainrib, A. Gramfort, “A Deep Learning Architecture for Temporal Sleep Stage Classification Using Multivariate and Multimodal Time Series”. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 26(4), pp.758 – 769,2018

[31] I. Fernández-Varela,E. Hernández-Pereira, V. Moret-Bonillo,“A Convolutional Network for the Classi fication of Sleep Stages”. Proceedings 2018, 2, 1174.

[32] R. K. Tripathy, 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.2018

[33] Z. Cui, X. Zheng, X. Shao, L. Cui, “Automatic sleep stage classification based on convolutional neural network and fine grained segments. Hindawi Complex 2018:9248410.

[34] A. Supratak, H. Dong, C. Wu, Y. Guo, DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG. IEEE Trans. Neural Syst. Rehabil. Eng. 2017.

[35] S. Khalighi, T. Sousa, J. M. Santos, U. Nunes, “ISRUC-Sleep: A comprehensive public dataset for sleep researchers”,Computer Methods and Programs in Biomedicine 124, pp.180–192,2016.

[36] M. Robnik-Šikonja, I. Kononenko, “Theoretical and Empirical Analysis of ReliefF and ReliefF. Machine Learning” 53, 23–69,2003

[37] J. Xu, B. Tang, H. He, H Man,“Semisupervised Feature Selection Based on Relevance and Redundancy Criteria”IEEE Trans Neural Netw Learn Syst. 28 (9) :pp. 1974-1984, 2017.

[38] T. H. Sanders, M. McCurry, M. A. Clements, “Sleep stage classification with cross frequency coupling”.36th Annu. Int. Conf. IEEE Eng. Med. Biol. (EMBC), pp.4579–4582,2014.

[39] V. Bajaj, R. Pachori, “Automatic classification of sleep stages based on the time-frequency image of EEG signals” Comput. Methods Programs Biomed.,112, (3), pp.320–328, 2013.

[40] Y.-L. Hsu, Y.-T. Yang, J.-S. Wang, C.-Y. Hsu, “Automatic sleep stage recurrent neural classifier using energy features of EEG signals”, Neurocomputing,104, pp.105–114,2013.

[41] I. Zibrandtsen, P. Kidmose, M. Otto, J. Ibsen, TW. Kjaer, “Case comparison of sleep features from ear-EEG and scalp-EEG”. Sleep Sci., 9(2),pp.69–72,2016.

[42] RB. Berry, R. Brooks, CE. Gamaldo, SM. Hardsim, RM. Lloyd, CL. Marcus, BV. Vaughn, “The AASM manual for the scoring of sleep and associated events: rules, terminology and technical specifications”. American Academy of Sleep Medicine, 2014.

[43] J. Sim, C. C. Wright, “The kappa statistic in reliability studies: use, interpretation, and sample size requirements,” Physical therapy, 85(3), pp.257–26 8,200 5.

[44] S.-F. Liang, C.-E. Kuo, Y.-H. Hu, Y.-H. Pan, and Y.-H. Wang, “Automatic Stage Scoring of Single-Channel Sleep EEG by Using Multiscale Entropy and Autoregressive Models”, vol. 61,2012.

[45] J. Kim, “A Comparative Study on Classification Methods of Sleep Stages by Using EEG,” J. Korea Multimed. Soc., vol. 17, no. 2, pp. 113–123, Feb. 2014.

[46] M. Peker, “A new approach for automatic sleep scoring: Combining Taguchi based complex-valued neural network and complex wavelet transform,” Comput. Methods Programs Biomed., vol. 129, pp. 203–216, Jun. 2016.

[47] A. Subasi, M. K. Kiymik, M. Akin, and O. Erogul, “Automatic recognition of vigilance state by using a wavelet-based artificial neural network,” Neural Comput. Appl., vol. 14, no. 1, pp. 45–55, Mar. 2005.

[48] M. E. Tagluk, N. Sezgin, and M. Akin, “Estimation of Sleep Stages by an Artificial Neural Network Employing EEG, EMG and EOG,” J. Med. Syst., vol. 34, no. 4, pp.717–725, 2010.

[49] A. R. Hassan and M. I. H. Bhuiyan, “An automated method for sleep staging from EEG signals using normal inverse Gaussian parameters and adaptive boosting,” Neurocomputing, vol. 219, pp. 76–87, 2017.

[50] A. R. Hassan and M. I. H. Bhuiyan, “Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting,” Comput. Methods Programs Biomed., vol. 140, pp. 201–210, 2017.

[51] M. Diykh and Y. Li, “Complex networks approach for EEG signal sleep stages classification,” Expert Syst. Appl., vol. 63, pp. 241–248, 2016.

[52]M. Diykh, Y. Li, and P. Wen, “EEG sleep stages classification based on time domain features and structural graph similarity,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 24, no. 11, pp. 1159–1168, 2016.

[53] S. Mahvash Mohammadi, S. Kouchaki, M. Ghavami, and S. Sanei, “Improving time–frequency domain sleep EEG classi-fication via singular spectrum analysis,” J. Neurosci. Methods, vol. 273, pp. 96–106, 2016.

[54] B. Şen, M. Peker, A. Çavuşoğlu, and F. V. Çelebi, “A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms,” J. Med. Syst., vol. 38, no. 3, p. 18, Mar. 2014.

[55] N. Burioka et al., “Approximate entropy in the electroencephalogram during wake and sleep.,” Clin. EEG Neurosci., vol. 36, no. 1, pp. 21–24, 2005.

[56] L. Fraiwan, K. Lweesy, N. Khasawneh, M. Fraiwan, H. Wenz, and H. Dickhaus, “Classification of Sleep Stages Using Multi-wavelet Time Frequency Entropy and LDA,” Methods Inf. Med., vol. 49, no. 3, pp. 230–237.

[57] L. J. Herrera et al., “Combination of Heterogeneous EEG Feature Extraction Methods and Stacked Sequential Learning for Sleep Stage Classification,” Int. J. Neural Syst., vol. 23, no. 03, p. 1350012, Jun. 2013.

[58] H. Simões, G. Pires, U.Nunes,V. Silva "Feature Extraction and Selection for Automatic Sleep Staging using EEG," In Proceedings of the 7th International Conference on Informatics in Control, Automation and Robotics, 3, 128-133.

[59] S. Khalighi, T. Sousa, D. Oliveira, G.Pires, U. Nunes, "Efficient feature selection for sleep staging based on maximal overlap discrete wavelet transform and SVM, "Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2011.

[60] S. Khalighi, T. Sousa, J. M. Santos, U. Nunes, "ISRUC-Sleep: A comprehensive public dataset for sleep researchers," Computer Methods and Programs in Biomedicine, 124, 180–192, 2016.

[61] T. Sousa, A. Cruz, S.Khalighi, G. Pires, U. Nunes, "A two-step automatic sleep stage classification method with dubious range detection". Computers in Biology and Medicine, 59, pp.42–53, 2015.

[62] S. Khalighi, T. Sousa, G. Pires, U. Nunes, "Automatic sleep staging: A computer assisted approach for optimal combination of features and polysomnograp hic channels, "Expert Systems with Applications, 40(17), pp.7046–7059, 2013.

[63] A. R. Hassan, and M. I. H. Bhuiyan, "Computer-aided sleep staging using Complete Ensemble Empirical Mode Decomposition with Adaptive Noise and bootstrap aggregating," Biomedical Signal Processing and Control, 24,pp.1–10, 2016.

[64] S. Najdi, A. A.Gharbali, J. M. Fonseca, "Feature Transformation Based on Stacked Sparse Auto encoders for Sleep Stage Classification," Technological Innovation for Smart Systems,pp.191–200, 2017.

[65] K. D.Tzimourta, A. K.Tsilimbaris, A. T. Tzioukalia, M. G.Tzallas, L. G. Tsipouras, "EEG-Based Automatic Sleep Stage Classification,"Biomed J Sci &Tech Res, 7(4),2018.

[66] H. Kalbkhani, G. Ghasemzadeh, M. Shayesteh, "Sleep Stages Classification from EEG Signal based on Stockwell Transform," IET Signal Processing, 2018.

[67] W. Huang, B. Guo, Y. Shen, X. Tang, T. Zhang, D. Li, & Z. Jiang, "Sleep staging algorithm based on multichannel data adding and multi feature screening, "Computer Methods and Programs in Biomedicine, 2019

[68] Q. Wang, D. Zhao, Y. Wang, and X. Hou, "Ensemble learning algorithm based on multi-parameters for sleep staging,". Medical & Biological Engineering & Computing, 2019

[69] M. Sharma, S. Patel, S. Choudhary, and U. R. Acharya, "Automated Detection of Sleep Stages Using Energy-Localized Orthogonal Wavelet Filter Banks, "Arabian Journal for Science and Engineering, 2019.

[70] S. Dhok, V. Pimpalkhute, A. Chandurkar, AA. Bhurane, M. Sharma, UR. Acharya, "Automated phase classification in cyclic alternating patterns in sleep stages using Wigner-Ville Distribution based features," Comput Biol Med. 119:103691, 2020.

[71] S. Santaji, and V. Desai, "Analysis of EEG Signal to Classify Sleep Stages Using Machine Learning, "Sleep and Vigilance, 2020.

[72] J. Zhou, Y. Tian, G. Wang, J. Liu, D. Wu, W.Xu, … Y. Hu "Automatic Sleep Stage Classification with Single Channel EEG Signal Based on Two-layer Stacked Ensemble Model". IEEE Access,1–1.2020

Mr.Santosh Kumar Satapathy was born in Parlakhemundi, India, in 1987.He received the B.Tech. Degree in Information Technology from the BPUT, Odisha, India, in 2009, and the M.Tech. Degree in Computer Science and Engineering from the BPUT, Odisha, India, in 2012.In 2012, he joined the Department of Information Techn ology, Roland Institute of Techno logy, Berhampur, Odisha, as a Lecturer, and in 2013 he joined as an assistant professor in the Department of Computer Science and Engineering, Gandhi Engineering College, Bhubaneswar, Odisha. He is currently a PhD student in the Department of Computer Science and Engineering at Pondicherry Engineering College, Puducherry. His research interests include data mining, biomedical signal analysis, and machine learning.

Dr.D.Loganathan is a Professor in the Department of CSE, Pondicherry Engineering Colleg e, and Puducherry, India. He has over 32 years of teaching experience. He obtained his PhD from Information & Communication Engineering Faculty, College of Engineering, Anna University, Chennai, INDIA in 2008.He obtained Graduated and Post-Graduated with specialization in Computer Science from Birla Institute of Technology and Science, Pilani, Rajasthan, India in an integrated way from 1978 – 1984. He is a professional member of CSI, IE, IEEE and ISTE. His area of interest is image processing and Information security. He has published more than 30 papers in the different journals and conferences.