Declaration of Authorship

We hereby declare that this thesis titled "Sleep Stage Classification With Machine learning Models (RandomForestClassifier and DecisionTreeClassifier)" is an authentic research carried out by Musa S Jawo, Haddy Jasseh , and Ismahan Abdifatah as requirement for the award of degree Bachelor of Science in Computer Science and Engineering at the Islamic University of Technology, Gazipur, Dhaka, under the supervision of Lutfun Nahar Lota Assistant Professor , Department of Computer Science and Engineering, Islamic University Oof Technology(IUT), Dhaka, Bangladesh and Co-Supervisor Tanjila Alam Sathi Lecturer Department of Computer Science and Engineering, Islamic University Oof Technology(IUT), Dhaka, Bangladesh .

The matter embodied in this thesis has not been submitted in part or full to any other institute for award of any degree.

Authors

....
Musa S Jawo
Student ID: 180041152

....
Haddy Jasseh
Student ID: 180041160

Ismahan
....
Ismahan Abdifatah
Student ID: 180041252

Supervisor:

on behalf of
....
Lutfun Nahar Lota
Assistant Professor,
Department of Computer Science and Engineering ,
Islamic University Of Technology

Acknowledgment

We would like to express our gratitude to God Almighty for guiding us through all of our challenges. Day by day, we have felt your guiding. You are the one who allowed us to complete our degree. We will continue to put our faith in you for our future.

We would also like to express our gratitude to our supervisors Lutfun Nahar Lota and Tanjila Alam Sathi, who enabled us to complete this work. Their instruction, dedication, patience and counsel guided us through all stages of this essay writing. They have been instrumental and exceptional mentors, providing us with insightful feedback, challenging ideas, and pushing us to strive for excellence.

We are equally grateful to Dr. Md. Azam Hossain for the continuous encouragement and support throughout the journey.

Finally, we would like to thank our parents and families as a whole whom the Almighty chooses for their moral support, endless prayers and motivation they have rendered throughout the course of our research. We are always grateful to the committee members and friends for their support.

1 Abstract

Automated classification of sleep stages is in demand to overcome the limitations of manual sleep stage classification. Analyzing sleep stages manually using neurophysiological signals and inspecting visually is very difficult, time-consuming process. Many techniques have been proposed already in the past decades. Sleep experts, physicians do not have assurance with such techniques concerned with accuracy, specificity and sensitivity. Sleep state classification using electroencephalogram (EEG) signals is crucial for understanding sleep patterns and diagnosing sleep disorders. This thesis aims to improve the accuracy and robustness of sleep state classification by employing a voting technique that combines multiple classification models. The research involves preprocessing and feature extraction from EEG signals, training individual classification models, and applying a voting mechanism to make the final sleep state classification decision. The proposed approach aims to enhance the reliability of sleep stage classification and contribute to the field of sleep medicine. Statistical features are extracted and trained with Decision Tree, Support Vector Machine and Random Forest algorithms with different testing dataset percentage. Results show combination of Random forest and decision tree algorithm achieves 90% of accuracy.

Contents

2 Introduction

For humans, sleep is a vitally important physiological phenomenon that helps the body restructure. The majority of a person's physiological functions are inactive while they fall asleep. More growth hormones and prohormones are secreted by the pituitary gland during this time, which encourages cell and tissue repair, gets rid of human exhaustion, and gets ready for physiological activities while people are awake.It is important to keep in mind that multiple sleep periods might be created based on the depth of sleep. According to recent studies, there are three main stages of sleep that may be identified by certain brain waves and their ratios: wake (W), norapideyesmovement (NREM), and rapid eye movement (REM). The Rechtstaffen and Kamp (R&K) recommendations further separated the NREM stage into four stages: 1, 2, 3, and 4. (also referred to as S1, S2, S3, and S4). W, S1, S2, S3, S4, and REM are the six stages that make up the typical RK sleep cycle

Traditional sleep monitoring, also known as polysomnography (PSG), has the drawback of requiring the application of a wide range of potentially sleep-disturbing sensors to the body and requiring only highly qualified sleep technologists or scientists to interpret the findings. Traditional PSG, while important in the diagnosis of sleep disorders, is relatively unsuitable for routine, non-diagnostic sleep monitoring and will simply cause more sleep disturbances when used on a daily basis by untrained people. This situation highlights the necessity for inconspicuous sleep monitoring techniques, ideally ones that are inexpensive and don't require special training to use. Cardiorespiratory monitoring is a potential technology for private, continuous, and unobtrusive sleep monitoring since it can be unobtrusive and the data can be evaluated by a computer. We used machine learning models alongside a z-normalizer. All these implemented models were used to see which model will give higher performance base on F1-score and Accuracy

2.1 Research Objectives

Sleep is a fundamental need of the human body. In order to maintain health, sufficient sleep is a must. Efficiency of sleep is based on sleep stages. Sleep stage classification is required to identify sleep disorders. Sleep stage classification identifies different stages of sleep and helps in the easy diagnosis of sleep disorders in patients. To automatically diagnosis sleep disorder in order to save time, resources, reliable and affordable healthcare to masses the objectives of this research is as follows:

1. Review existing literature on sleep state classification using EEG signals and voting classifiers to identify gaps and challenges in the field.
2. Preprocess and analyze EEG data to extract relevant features that capture distinct characteristics of different sleep stages.

3. Explore and select a set of diverse classification models suitable for sleep state classification based on their performance and compatibility with the voting classifier.
4. Train and optimize individual classifiers using the extracted features and annotated sleep state labels.
5. Implement a voting classifier that combines the predictions of the individual classifiers to make the final sleep state classification decision.
6. Evaluate the performance of the proposed sleep state classification system using appropriate evaluation metrics and compare it with existing methods.
7. Assess the robustness and generalizability of the system by testing it on diverse datasets and validating its effectiveness across different populations.
8. Analyze the impact of the voting classifier on the accuracy and reliability of sleep state classification, highlighting its advantages over individual classifiers

2.2 Models Used

1. **RandomForest Classifier** : The random forest classifier is a supervised learning algorithm which you can use for regression and classification problems. It creates a set of decision trees from a randomly selected subset of the training set
2. **Decision Tree Classifier**: Decision trees are an intuitive supervised machine learning algorithm that allows

you to classify data with high degrees of accuracy.This means that they use pre labeled data in order to train an algorithm that can be used to make a prediction
3. **KNeighborsClassifier**: K-NN algorithm stores all the available data and classifies a new data point based on the similarity.
4. **SuperVectorMachine**:A support vector machine (SVM) is a supervised machine learning model that uses classification algorithms for two-group classification problems. After giving an SVM model sets of labeled training data for each category, they're able to categorize new text.It can handle both classification and regression on linear and non-linear data.
5. **xgboost Classifier**: which stands for Extreme Gradient Boosting, is a scalable, distributed gradient-boosted decision tree (GBDT) machine learning library. It provides parallel tree boosting and is the leading machine learning library for regression, classification, and ranking problems.
6. **Voting Classifier** : A voting classifier is an ensemble learning technique that combines the predictions of multiple individual classifiers to make a final prediction or decision. It is commonly used in machine learning for tasks such as classification or regression, where combining the outputs of multiple models can lead to improved performance and robustness. The voting classifier works by aggregating the predictions of its constituent classifiers and selecting the class or value that receives the most votes.

Z-score optimization New value:

$$\frac{(x-\mu)}{\sigma}$$

Where:

$$x : Originalvalue$$

$$\mu : Meanofdata$$

$$\sigma : Standarddeviationofdata$$

2.3 Research Challenges

The follow where major challenges that were faced in sleep stage classification using EEG signals:

1. **Variability and Complexity of EEG Signals:** EEG signals exhibit significant inter- and intra-subject variability, making it challenging to identify consistent patterns for sleep stage classification. The signals are influenced by factors such as age, gender, health conditions, and electrode placement, which can introduce additional complexity.

2. **Imbalanced Class Distribution:** Sleep stages are often imbalanced, with some stages occurring more frequently than others. This class imbalance can impact the performance of classification models, as they may become biased towards the majority class, leading to reduced accuracy in classifying minority classes.

3. **Individual Differences and Personalized Classification:** Sleep patterns and EEG characteristics can

vary significantly among individuals. Developing personalized sleep stage classification models that consider individual differences and adapt to specific individuals' sleep profiles poses a challenge. Accounting for subject-specific features and preferences can improve classification accuracy and facilitate personalized sleep analysis.

4. **Artifact Detection and Removal:** EEG signals are susceptible to various artifacts, such as eye movements, muscle activity, and electrical interference. Accurate detection and removal of these artifacts are crucial for reliable sleep stage classification. Developing robust artifact detection techniques and incorporating them into the classification pipeline is a significant challenge

3 Problem Statement

With respect to huge medical technological advancement, we identified that many people have been affected and suffering from sleeping problems without detecting it at its starting stage. This is due to the low level of analysis and built in models incorporated in the medical systems designed to collect and analyze signals from patients. We will use Machine learning models(Random forest and Decision tree) techniques to analyze EEG data from patients to accurately classify them to the 5 sleep stages named W,RM, W1,W2 and W3 and put forward which model is perfect for dealing with EEG data with good results. This paper of ours will go a long way to help people detect and diagnose sleeping disorders and its related concerns at its early stage rather than later, to save more lives.

4 Literature Review

The following papers were reviewed and briefly discussed in relation to our study project.
Tzimourta et al (2018) (1) This paper focuses on the EEG-based automatic sleep stage classification. The dataset used is from the ISRUC-sleep dataset, that consists of 100 adults with sleep disorders evidence and one recording per subject, 8 adults with sleep disorders evidence and two recordings per subject and, and 10 healthy subjects and one recording per subject. The proposed methods consist of two stages: the feature extraction and the classification of the machine learning model, which are Naive Bayes, Decision Trees, K-Nearest Neighbors, and Support Vector Machines. The results reached in terms of accuracy were 75.29 percent with Random Forests.
The limitations of this method are that only a few features were examined; they weren't able to use several feature selection methods. Compared with other classification methods, random forests have higher accuracy.

Sant et al (2020) (2) This paper work is focusing on the analysis of EEG signal to classify sleep stages using machine learning algorithms by considering 10s of epochs and using a band-pass filter, EEG signals are filtered and divided into frequency sub-bands. the dataset of this paper was obtained from the Sleep-EDF database and from Dr. Chandrasekhar's clinic with a sample size presented in Figure 1.

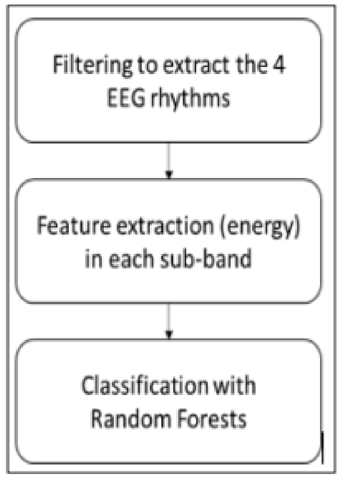
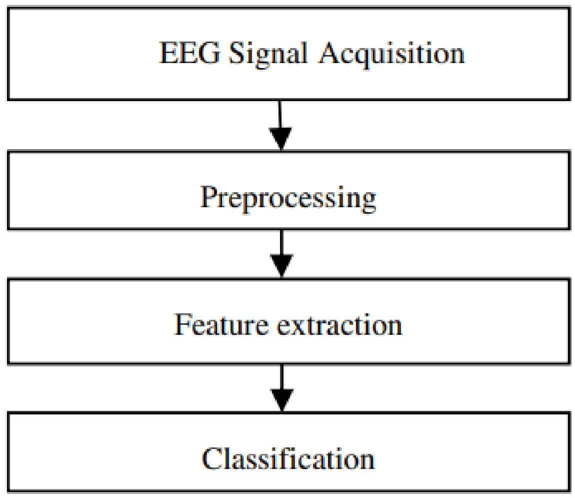


Figure 1: Flowchart of the proposed methodology.

of 125 patients from all age groups between 20-50. The methods in this paper are EEG dataset, preprocessing feature extraction and others,and the machine learning algorithms used in this paper is Decision Tree, support vector machine and Random Forest for sleep stage classification. Compared to SVM and DT algorithms, RF offers greater accuracy of 97.8%. The suggested approach is contrasted with the methods used by other researchers. It is implied that while many of these strategies have relied on a small number of test participants to determine the accuracy of sleep stages, our approach is workable and offers simple application. Comparing the suggested method to the most recent studies on the classification of sleep stages, there is no doubt that it is more accurate and practical.



l Proposed flowchart of sleep stage classification

Smith et al (2021)(3) The work in this paper is focusing on the application of machine learning to sleep stage classification The data from adult Wistar rats (n = 8) were used in experiments in a facility fully accredited by the American Association for the Accreditation of Laboratory Animal Care. The methods used are sleep EEG/EMG data collection, data processing and annotation, input formalization, evaluation of machine learning classifiers, and others. The highest performing classifier is the Random Forest when compared to other classifiers, with 95.78%. A domain expert will also reevaluate categorization pairs in future research that the models in this work found confusing.

Zhao et al. (2022)(4) The work in this paper is focused on the evaluation of a single-channel EEG-based sleep staging algorithm in which 57 features were extracted from three different aspects. The methods include a screening of datasets from the expanded Sleep-EDF (ES-EDF) database, which was taken from 12 healthy subjects aged between 21-34, consisting of 5 males and 7 females. The methods include feature extraction, time domain feature selection, frequency domain feature selection, rank-based feature selection, and many others. The classification models used in this paper are the support vector machine (SVM), backpropagation neural network (BPNN), random forest (RF), and decision tree (DT) algorithms. The random forests achieved the highest accuracy of 94.85% when all of the features were used. The limitations of this paper are: First, the REM and N1 phases have reduced recognition rates. Second, the N1 phase had the highest concentration of wrong W phase predictions. These two issues' primary causes are as follows: The REM and N1 stages cannot be easily distinguished because this study only extracted features based on EEG, and the second point is due, in part, to the slow eye movements that occur in both the closed-eye W and N1 stages, which are characterized by low voltage mixed frequency waves in the EEG of the REM and N1 stages. · On the other hand, the experts' interpretation becomes more arbitrary while moving from the W stage to the N1 stage, making it challenging to ensure the accuracy of the outcomes of the sleep staging. Therefore, increasing the rate at which the REM and N1 stages are recognized is still an area where research on sleep staging should be directed.

Smith et al (2021)(5) This paper focuses on a comparative study on the classification of sleep stages based on EEG signals using feature selection and classification algorithms. University College Dublin and St. Vincent's University Hospital contributed the data set that was used in the study, with a sample size of 25 individuals consisting of 21 males and four females aged between 50 and 28-68 years.

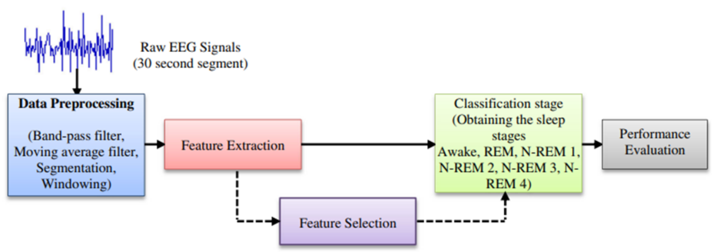
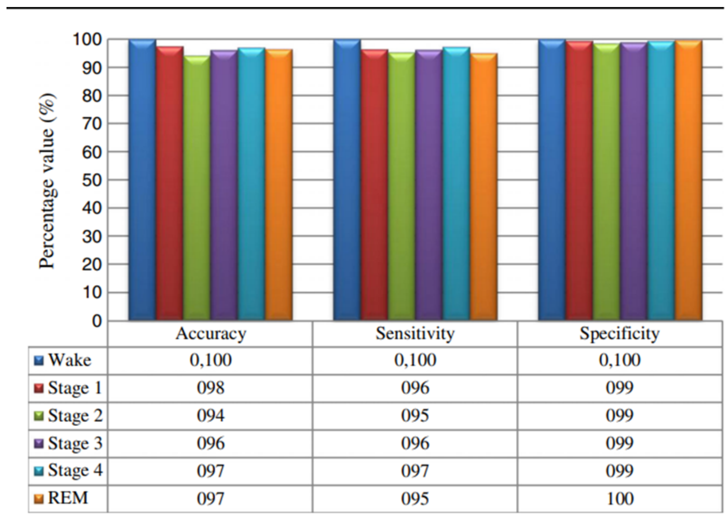
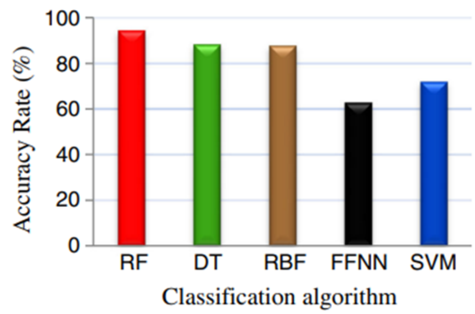


Figure 1: Architecture of the model

The methods used are feature extraction, which was divided into 4 categories (time domain, non-linear, frequency-based, and entropy-based features), feature selection, and classifier algorithms (decision tree, feed-forward neural network, radial basis network, SVM, and random forest). The performance evaluation methods consist of classification accuracy, confusion matrix, analysis of sensitivity and specificity analysis, and k-fold cross-validation. The random forests achieved the highest accuracy of 97.03 percent.



	Accuracy	Sensitivity	Specificity
Wake	0,100	0,100	0,100
Stage 1	098	096	099
Stage 2	094	095	099
Stage 3	096	096	099
Stage 4	097	097	099
REM	097	095	100



Qureshi et al (2017)(6) The work in this paper is focused on evaluating different machine learning techniques for classifying sleep stages on single-channel EEG. A whole-night polysomnogram from 25 subjects was recorded using R and K standard. The methods used in this study were obtaining raw EEG signals, pre-

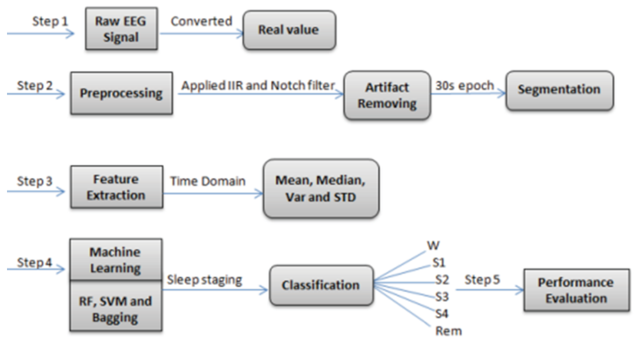


Figure 2: Qureshi's work Flow

processing, feature extraction, machine learning and performance evaluation. The machine learning algorithms used were Random Forest, support vector machine and bagging classifier. The performance evaluation methods consist of classification accuracy, confusion matrix, sensitivity, specificity, and k-fold cross-validation. For the classification of sleep stages, it was discovered that RF had the greatest accuracy rate (97.73%) across all outcomes.

Aboalayon, K. A., Almuhammadi, W. S., amp; Faezipour, M. (2015, May) (7) This paper focuses on A comparison of different machine learning algorithms using single channel EEG signal for classifying human sleep stages. Due to data similarities, this study mixes REM and Stage 1 NREM. Following that, performance is contrasted using 20 healthy subjects39; single channel EEG data. Numerous supervised machine learning classifiers were employed, including multi-class Support Vector Machines (SVM), Decision Trees (DT), Neural Networks (NN), K-

Nearest Neighbors (KNN), and Naive Bayes (NB). According to the findings, the suggested method for differentiating sleep stages successfully achieves high accuracy of 97.30% using DT classifier.

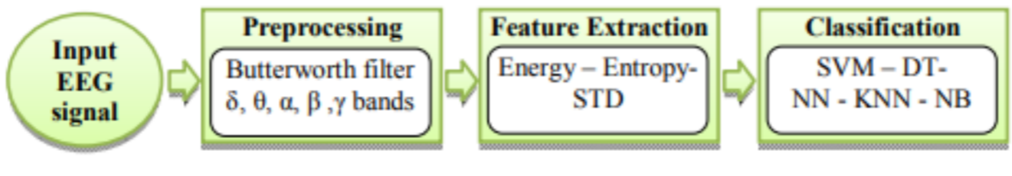


Figure 3: Aboalayon, K. A's model

Chriskos, P., Kaitalidou, D. (2017) (8) This paper focuses on Automatic sleep stage classification applying machine learning algorithms on EEG recordings. The suggested methodology makes use of modern mathematical techniques like graph theory metrics and synchronization likelihood applied to sleep EEG data. The resulting features are then included into three different machine learning methods, including neural networks, support vector machines, and k-nearest neighbors. A graph theoretical analysis was performed on the EEG recordings of 23 healthy young men between the ages of 20 and 45. According to their accuracy, the evaluation of their comparative performance is looked into. It's interesting that the support vector machine gets the highest accuracy achievable, of 89.07making it a good tool for classifying sleep stages. It may be inferred from the experimental findings in this study that the graph metrics generated and extracted from sleep epochs, as well as the synchronization likelihood values, are acceptable for classifying sleep stages. The goal of future research in this area is to improve classification accuracy by using more data samples, more precise feature extraction, and finer feature selection.

Qureshi, S., amp; Vanichayobon, S. (2017, July). (6) This paper focuses on Evaluate different machine learning techniques for classifying sleep stages on single-channel EEG. In this study, they offer three distinct machine learning methods—Random Forest, Bagging, and Support Vector Machines—along with a time domain feature for categorizing different stages of sleep based on single-channel EEG. 25 patients39; polysomnograms throughout the entire night were recorded using the Ramp;K standard. According to the results, Random Forest classifiers have overall accuracy, specificity, and sensitivity levels of 97.73%, 96.3%, and 99.51%, respectively. the five phases of the suggested method are: collecting raw EEG signals; filtering the raw EEG data; feature extraction; machine learning; and performance assessment.

Uçar, M. K., Bozkurt, M. R., Bilgin, C., amp; Polat, K. (2018) (9) This paper focuses on Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques. The goal of the study was to identify sleep and wakefulness using a useful and usable method. The signal of heart rate variability (HRV) has been obtained for this purpose from photoplethysmography (PPG).

PPG and HRV signals have been used to extract features. Then, using the F-score feature selection method, the features that will accurately reflect sleep and wakefulness were chosen. The k-nearest neighbors classification technique and support vector machines were used to categorize the chosen features.

Satapathy, S. K., amp; Loganathan, D. (2021) (10) This paper focuses on A study of human sleep stage classification based on dual channels of EEG signal using machine learning techniques In order to increase the accuracy of sleep staging, they suggest in this paper a useful automated system. They used the input signal in their work to extract both linear and non-linear features. The generated feature vector was then reduced using a feature selection method based on the Relief weight algorithm to identify a set of best features. Four machine learning methods, including the support vector machine, K-nearest neighbor algorithm, decision tree, and random forest, were used to categorize the chosen features. In this study, the 10-fold cross validation strategy was considered. Our suggested methods produced the highest classification accuracy results, 91.67% with the C4A1 channel and 93.8% with the O2-A1 channel when utilizing the Random forest classification model. Four fundamental processes were used to carry out the proposed research project: (1)preprocessing the signal, (2) feature extraction, (3) feature screening, and (4) classification.

Satapathy, S. K., Kondaveeti, H. K., Sreeja, S. R., Madhani, H., Rajput, N., amp; Swain, D. (2023) (10) This paper focuses on A Deep Learning Approach to Automated Sleep Stages Classification Using Multi-Modal Signals. Sleep-EDF was the 2013 version of the Sleep Cassette (SC) subset in the Sleep-EDF Expanded dataset, consisting of 30 healthy subjects aged 25 to 34. This experimental study compares the two methods and selects the most effective course of action. A neural network made up of CNN and LSTM as well as the three major machine learning classifiers Random Forest (RF), K-Nearest Neighbors (KNN), and Support Vector Machines (SVM) have all been trained on a large base of heterogeneous data. With CNN + LSTM, the suggested model has an accuracy of 87.4%, while the baseline ML algorithms ranged from 74.07% to 83.65%. The proposed methodology

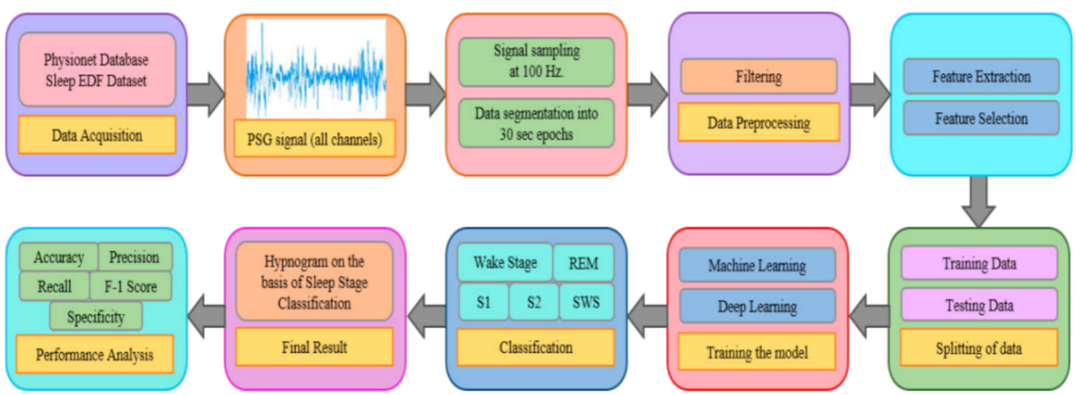
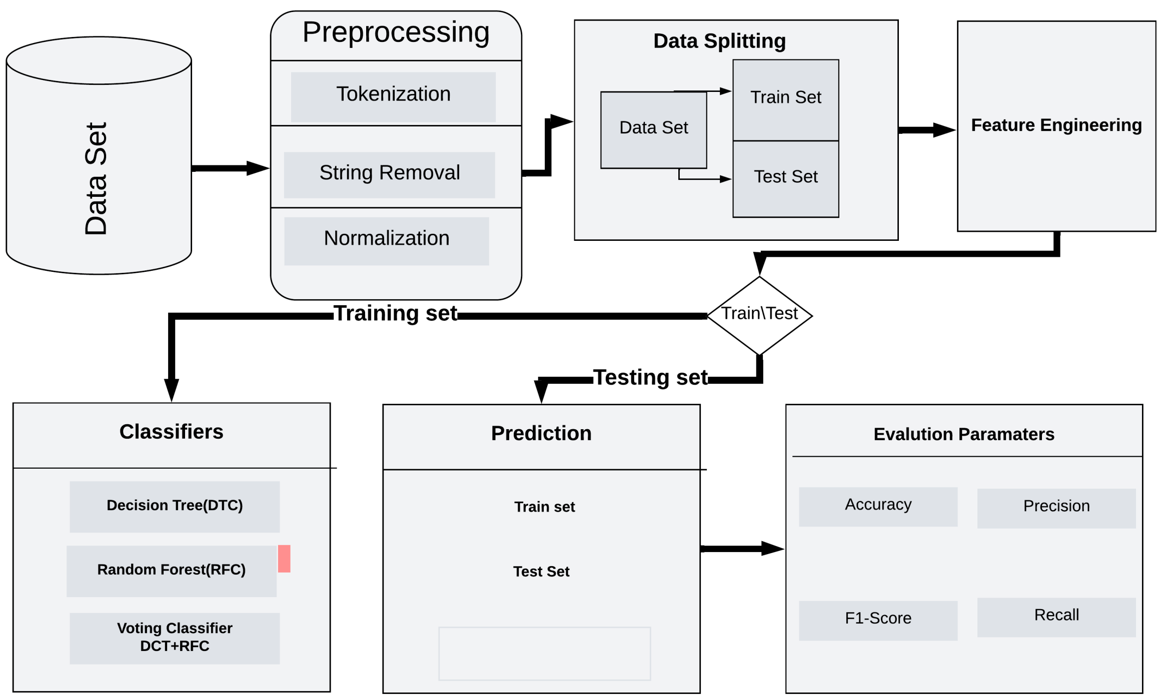


Fig. 1. Model Workflow

5 Design Methodology

Dataset acquisitions, , pre-processing, and classifiers of machine learning models used are well emphasis in this sections

Work Flow model of our thesis



5.1 Data source

Collect or obtain suitable EEG datasets containing labeled sleep state information from Refsan Janey's github that was used in this investigation by extracting it from the Research_5_Stages_Sleep_Classification_HMC database

. The information included 78 columns of records. The population consisted of 80 participants between the ages of 27 and 63 (57 men and 13 women; 53 to 135 kg), with a sample frequency of 100 Hz and a duration between 401 and 578 minutes. Records were classified into five classes: W, N1, N2, N3, and R.

5.2 Data pre-processing

Preprocess the EEG signals by removing artifacts, filtering noise, and segmenting them into relevant time intervals. Extract informative features from the preprocessed EEG signals using signal processing techniques, time-domain analysis, or frequency-domain analysis. Z-score regularization and oversampling techniques were used in the pre-processing. The purpose of the Z-score was to put the data in a range(-2.5 t0 2.5) for easy working with the data set. The EEG signal data set was considered the input of models, in which class imbalance was removed by a SMOTE and oversampling. A better performance was achieved after making sure all the columns had an equal number of data points when compared to the results obtain on the raw data set before pre-processing.

5.3 Classification

Select a set of diverse classifiers suitable for sleep state classification, such as decision tree (DT), random forest (RF), support vector machine (SVM), and the K-neighbor were applied for classification and compared with each other, which was suggested for evaluating the performance of sleep stages classification. Train and optimize the individual classifiers using the extracted features and the corresponding sleep state labels. The decision tree consists of a number of tests for making decisions that have a tree structure and operate using the division and conquest approach. Each non-leaf node has a connection to the bifurcation feature test. which establishes a threshold or split point on a feature to divide the data into subgroups. Each node's data is broken down into many subcategories based on the differences in the values. Each leaf node has an associated tag (class) for the samples it contains. A series of feature tests are run starting at the root node in the prediction stage, and results are gathered as they progress to the leaf nodes. The most recent ensemble approach, random forest, is regarded as the advancement of the bagging method, and the primary distinction between the two is the use of random features. At each stage of choosing the branch when creating a decision tree, RF chooses a set of features at random before continuing to choose the typical branch according to the feature test. In the RF technique for classification, a self-starter sample T is taken from the training data, and from each of those samples, an unpruned classification and regression tree (CART) is built. Finally, a majority vote for predicting all trained single-trees is used to create the classification. According to the findings, random forest and decision tree performed better than the other classifiers. Finally we Implemented a voting classifier that combines the predictions of the individual classifiers, either using majority voting or weighted voting based on classifier performance. Evaluate the performance of the sleep state classification system using cross-validation or independent testing, considering metrics such as accuracy, macro avg , weighted avg , and F1 score. Compare the results with existing sleep state classification methods to demonstrate the effectiveness and improvements achieved with the voting classifier.

6 Findings/Results

The effectiveness and performance of the classifiers are the main topics of this section. Accuracy is defined as the ratio of positive and negative values that the test accurately marks as positive and negative. Thus, the results obtained with Random forest from RafsanJany-44 implemented model and data set was (acc=0.74%) The result of the accuracy before 50 epochs was non-linear but after 50 to 100 epochs the accuracy remains constant as shown in fig.2

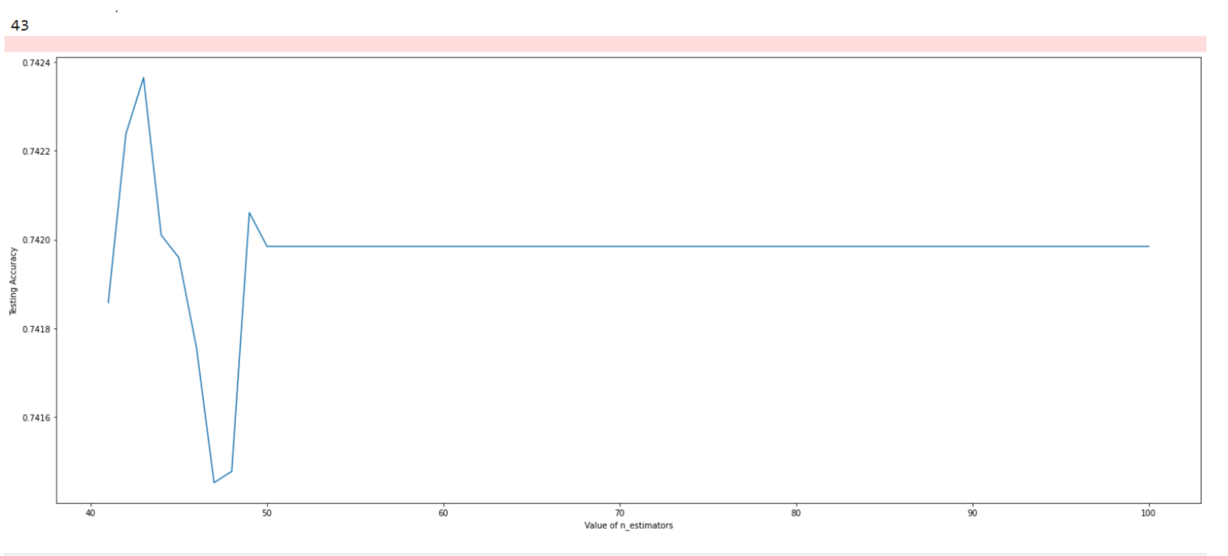


Figure 4: Accuracy graph of RafsanJany-44 models

6.1 Performance Analysis

In our implemented models of Random forest and decision tree we got higher accuracy for training and testing

- RandomForestClassifier(RFC)
- DecisionTreeClassifier(DTC)
- K-NeighborsClassifier(K-NC)
- SupportVectorClassifier(SVC)
- XgboostClassfier

1. Before applying regularization technique(SMOTE): Before applying any normalization technique these were some of the results obtained by using the various classifiers models as name above

Training Set					
Results	RFC(%)	DTC(%)	K-NC(%)	SVC(%)	Xgboost(%)
Accuracy	0.99	0.99	0.78	0.63	0.74
Macro Avg	0.99	0.99	0.75	0.53	0.69
Weighted Avg	0.99	0.99	0.78	0.59	0.73

Class wise analysis RANDOM FOREST CLASSIFIER RFC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	1.00	0.98	0.99	3915
1	0.97	1.00	0.98	12219
2	1.00	0.97	0.99	6531
3	1.00	0.99	0.99	5439
4	1.00	0.99	0.99	6448

Class wise analysis DECISION TREE CLASSIFIER DTC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	1.00	0.98	0.99	3915
1	0.97	1.00	0.98	12219
2	1.00	0.97	0.99	6531
3	1.00	0.99	0.99	5439
4	1.00	0.99	0.99	6448

From the above tables its is show that on the training dataset Random forest and Decision Tree classifiers has higher performance of 99% than the rest

The below chat shows that 99% result performance were obtained using Random forest and Decision Tree classifier on the raw data

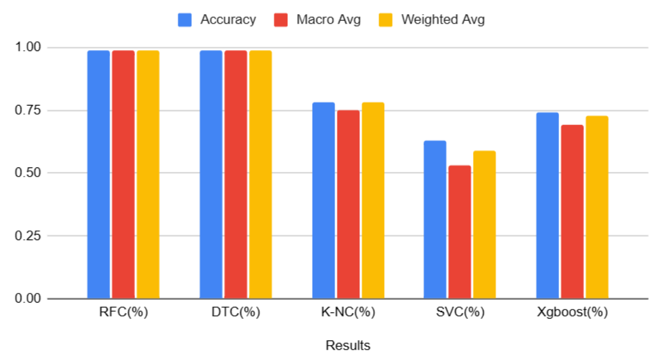


Figure 5: Accuracy graph for the models before SMOTE

On this training set the Random forest and Decision Tree classifiers out performed the rest but with low accuracy this is due to the class imbalance problem in the unstructured dataset

Testing Set					
Results	RFC(%)	DTC(%)	K-NC(%)	SVC(%)	Xgboost(%)
Accuracy	0.80	0.80	0.72	0.63	0.74
Macro Avg	0.76	0.76	0.68	0.53	0.69
Weighted Avg	0.80	0.80	0.71	0.59	0.73

Class wise analysis RANDOM FOREST CLASSIFIER RFC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	0.64	0.36	0.47	952
1	0.76	0.88	0.82	2970
2	0.90	0.81	0.85	1713
3	0.83	0.83	0.83	1416
4	0.83	0.90	0.86	1588

Class wise analysis DECISION TREE CLASSIFIER DTC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	0.64	0.36	0.47	952
1	0.76	0.88	0.82	2970
2	0.90	0.81	0.85	1713
3	0.83	0.83	0.83	1416
4	0.83	0.90	0.86	1588

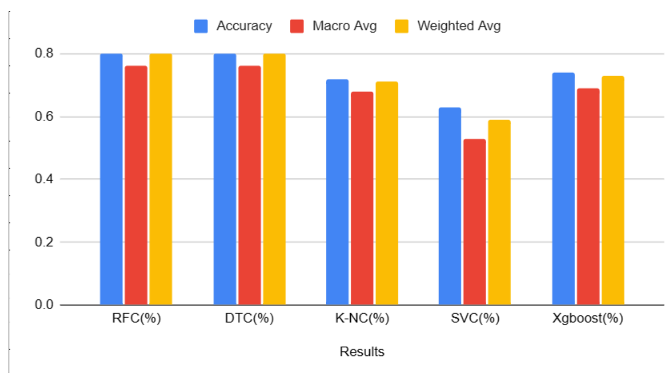


Figure 6: Accuracy graph for the models before SMOTE

2. After Applying the SMOTE Technique

After the application of normalization technique to make sure that all the classes have equal amount of dataset the follow result were obtained on the train and test set of the data

Training Set					
Results	RFC(%)	DTC(%)	K-NC(%)	SVC(%)	Xgboost(%)
Accuracy	0.99	0.99	0.89	0.72	0.75
Macro Avg	0.99	0.99	0.89	0.72	0.75
Weighted Avg	0.99	0.99	0.898	0.72	0.75

Class wise analysis RANDOM FOREST CLASSIFIER RFC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	0.94	1.00	0.98	31557
1	1.00	0.99	0.99	31567
2	1.00	0.99	0.99	31482
3	1.00	0.99	1.00	31481
4	1.00	0.99	0.99	31625

Class wise analysis DECISION TREE CLASSIFIER DTC				
Class	Precision(%)	recall(%)	f1-score(%)	Support(%)
0	1.00	0.98	0.98	12125
1	1.00	0.98	0.99	12157
2	0.94	1.00	0.99	12145
3	1.00	0.99	1.00	12167
4	1.00	0.99	0.99	12162

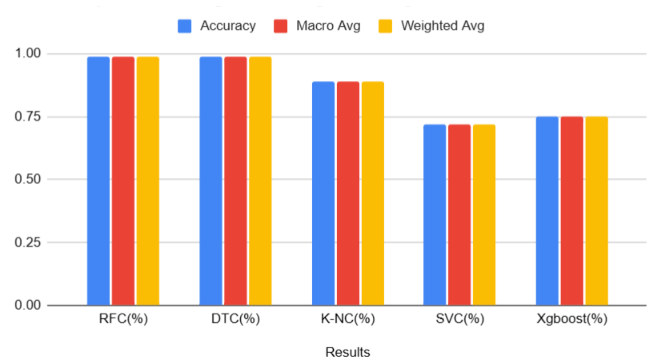


Figure 7: Accuracy graph for the models after SMOTE

The above tables and graph specifically shows that on the training set Random forest and Decision Tree classifiers has attend an accuracy of 99% and good performance on F1 score and other performance related in the table in class wise form analysis

The below tables and graphical representation shows that if the above mentioned classifiers are apply on the raw data set and use again after normalization of the dataset the performance on the normalized dataset will give higher performance

Testing Set					
Results	RFC(%)	DTC(%)	K-NC(%)	SVC(%)	Xgboost(%)
Accuracy	0.89	0.89	0.84	0.72	0.74
Macro Avg	0.88	0.88	0.84	0.72	0.74
Weighted Avg	0.88	0.88	0.84	0.72	0.74

Class wise analysis DECISION TREE CLASSIFIER DTC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	0.85	0.86	0.85	3064
1	0.86	0.78	0.82	3032
2	0.88	0.94	0.91	3044
3	0.92	0.92	0.92	3022
4	0.92	0.92	0.92	3027

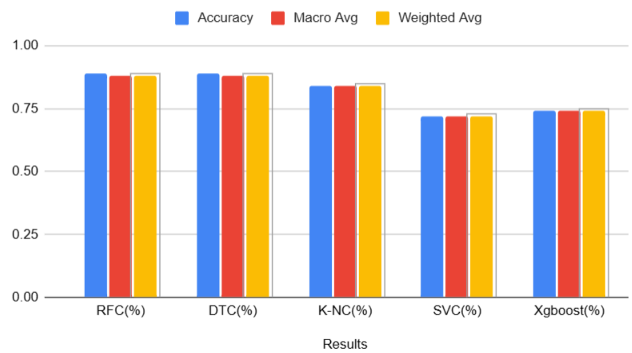


Figure 8: Accuracy graph for the models after SMOTE

3. **Voting Classifier/Ensemble classifiers** In this section we combined the best performing model to increase the performance on the data set this is apply on both the train and test set using 1. Random Forest + Decision Tree Classifier and 2. Decision Tree + K-NeighborsClassifie classifiers

6.1.1 On the Training dataset

The tables and chat below shows that Combinations of Random Forest and Decision Tree classifiers produces higher accuracy of 99% on the training set which out performs when Decision Tree and K-NeighborsClassifie classifiers are combine

Training Set		
Results	RFC+DTC(%)	DTC+K-NC(%)
Accuracy	0.99	0.87
Macro Avg	0.99	0.86
Weighted Avg	0.99	0.87

Class wise analysis Voting Classifier on DTC+RFC				
Class	Precision	recall	f1-score	support
0	1.00	0.98	0.99	12125
1	1.00	0.98	0.99	12157
2	0.94	1.00	0.99	12145
3	1.00	0.99	1.00	12167
4	1.00	0.99	0.99	12162

6.1.2 On the Test Set of the Data

In this section it is visualized that with the combination of Random forest and Decision Tree the accuracy of the model out performed both previous models used singly and the previous paper read

Test Set		
Results	RFC+DTC(%)	DTC+K-NC(%)
Accuracy Macro Avg Weighted Avg	0.90 0.89 0.90	0.85 0.85 0.86

Class wise analysis Voting Classifier on DTC+RFC				
Class	Precision(%)	recall(%)	f1-score(%)	Ssupport(%)
0	1.00	0.98	0.98	12125
1	1.00	0.98	0.99	12157
2	0.94	1.00	0.99	12145
3	1.00	0.99	1.00	12167
4	1.00	0.99	0.99	12162

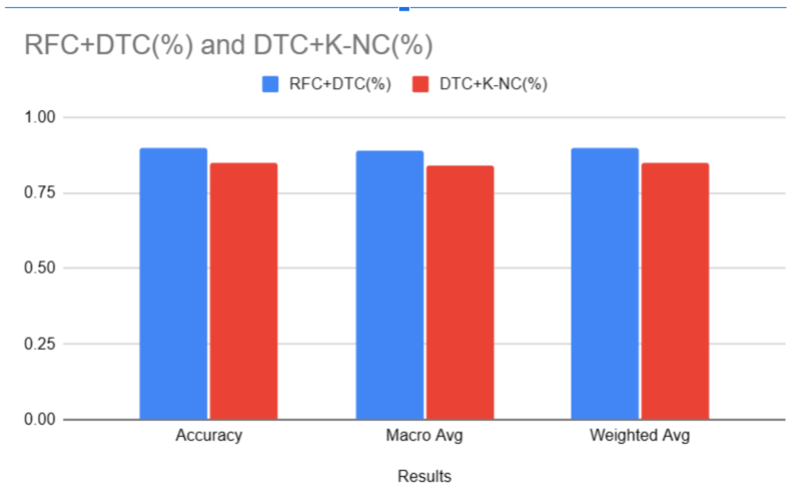


Figure 9: Accuracy graph for the models with Voting Classifier

6.2 Comparison of previous RFC Model of RafsanJany with our current implemented Models

Here all the models built accuracy are compared together with this visualization it shows that when the out performing Models are combine with the help of voting classifeir the accuracy on the EEG data set has increase by 1% which specifies that it do better than the previous research paper and models use

Models wise analysis							
Result	Previous Model RFC	RFC	DTC	K-NC	SVC	Xgboost	RFC+DTC
Accuracy	0.79	0.89	0.89	0.84	0.72	0.74	0.90

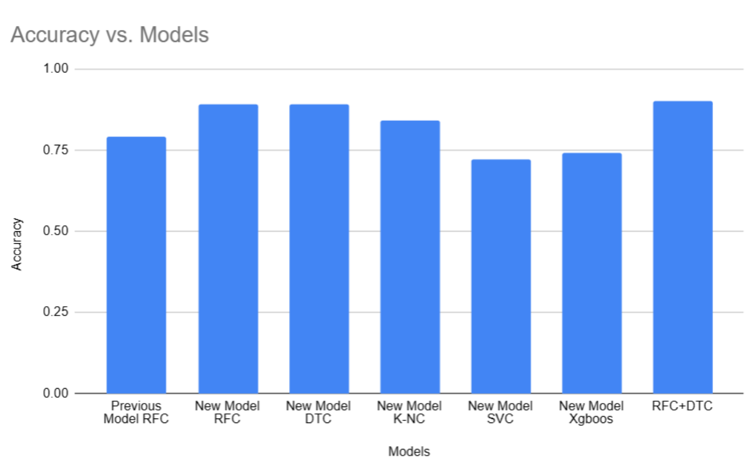


Figure 10: Accuracy graph for the models after SMOTE with Voting classifier

7 Contributions

This thesis work has several contributions to the papers read on the literature, as many of those papers were depending on using one model on the EEG dataset but with this research paper we were able to implement the models singly on the dataset and combination of the out performing models to generate better accuracy of 90% on the test dataset. Here are some other potential contributions that this papers has make on the previous models:

1. Improved predictive performance: Voting classifiers have the potential to enhance the predictive performance compared to individual classifiers. They can combine the strengths of multiple classifiers and mitigate their weaknesses, leading to improved accuracy, robustness, and generalization.

2. Improved wake detection will allow doctors and researchers to derive more accurate sleep quality measures from wireless body-worn equipment, such as sleep start latency, total sleep time, and wake occurrences following sleep onset.

3. Decision fusion strategies: Voting classifiers offer various decision fusion strategies, such as majority voting, weighted voting, and soft voting. This thesis can explore and evaluate different fusion strategies to determine their impact on the overall performance of the voting classifier. This analysis can help identify the most effective fusion strategy for a given problem domain.

4. Handling imbalanced datasets: Imbalanced datasets are common in many real-world applications. This thesis has investigate how voting classifiers can be adapted or enhanced to handle imbalanced data, such as by incorporating sampling techniques or modifying the decision fusion strategy to consider the class distribution

5. Comparative analysis: This thesis can compare the performance of voting classifiers with other classification approaches, such as single classifiers or other ensemble methods. This analysis can help establish the strengths and limitations of voting classifiers and highlight their effectiveness in specific scenarios.

8 Conclusion and Future Works

This work used random forest and decision tree to identify and categorize sleep stages from an EEG signal. To ensure that each class had an equal number of data sets, the SMOTE was utilized to prevent class imbalance. The accuracy was then computed using the decision tree and random forest performances, as well as xghost, SVM. decision tree, random forest, voting classifier (Random Forest and Decision Tree) with 90.0% accuracy, which was considered to be the maximum accuracy possible. Due to the characteristic of oversampling dimensional reduction and the majority vote in forecasting all trained trees connected to the random forest, better results were obtained. The proposed strategy in the current study performed better than the alternatives, assisting doctors in the precise detection of diagnosing sleeping disorders. In the future, deep learning based imaging, such as LSTM and Deep learning Techniques, can be proposed together with an existing model to boost the performance of EEG analysis

References

[1] K. D. Tzimourta, A. Tsilimbaris, K. Tzioukalia, A. T. Tzallas, M. G. Tsipouras, L. G. Astrakas, and N. Giannakeas, "Eeg-based automatic sleep stage classification," *Biomed J*, vol. 1, no. 6, 2018.

[2] S. Santaji and V. Desai, "Analysis of eeg signal to classify sleep stages using machine learning," *Sleep and Vigilance*, vol. 4, no. 2, pp. 145–152, 2020.

[3] A. Smith, H. Anand, S. Milosavljevic, K. M. Rentschler, A. Pocivavsek, and H. Valafar, "Application of machine learning to sleep stage classification," *arXiv preprint arXiv:2111.03085*, 2021.

[4] S. Zhao, F. Long, X. Wei, X. Ni, H. Wang, and B. Wei, "Evaluation of a single-channel eeg-based sleep staging algorithm," *International Journal of Environmental Research and Public Health*, vol. 19, no. 5, p. 2845, 2022.

[5] 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," *Journal of medical systems*, vol. 38, no. 3, pp. 1–21, 2014.

[6] S. Qureshi and S. Vanichayobon, "Evaluate different machine learning techniques for classifying sleep stages on single-channel eeg," pp. 1–6, 2017.

[7] K. A. Aboalayon, W. S. Almuhammadi, and

M. Faezipour, “A comparison of different machine learning algorithms using single channel eeg signal for classifying human sleep stages,” in *2015 Long Island Systems, Applications and Technology*. IEEE, 2015, pp. 1–6.

[8] P. Chriskos, D. S. Kaitalidou, G. Karakasis, C. Frantzidis, P. T. Gkivogkli, P. Bamidis, and C. Kourtidou-Papadeli, “Automatic sleep stage classification applying machine learning algorithms on eeg recordings,” in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*. IEEE, 2017, pp. 435–439.

[9] M. K. Uçar, M. R. Bozkurt, C. Bilgin, and K. Polat, “Automatic sleep staging in obstructive sleep apnea patients using photoplethysmography, heart rate variability signal and machine learning techniques,” *Neural Computing and Applications*, vol. 29, pp. 1–16, 2018.

[10] S. K. Satapathy and D. Loganathan, “A study of human sleep stage classification based on dual channels of eeg signal using machine learning techniques,” *SN Computer Science*, vol. 2, pp. 1–16, 2021.