© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716558

Applying Machine Learning Algorithms for The Classification of Sleep Disorders

GOLI KARTHEEK[1], VADLA SAI KALYAN[2], MUTHYALA VINAY GOUD[3], DEEPA PANSE[4]
[1,2,3,4]*Sreenidhi Institute of Science & Technology*

Abstract- sleep disorders, particularly sleep apnea, have an impact on people's health, and that reveals the need for a correct diagnosis. However, sleep experts use complex and time-consuming methods for the manual classification of the various sleep stages. Based on the Sleep Disorder Data which is available for public access and consists of 400 records and 13 attributes, this work introduces a machine learning classification model. A number of deep and techniques-based machine learning models are considered and their performance is evaluated in accurately diagnosing sleep disorders. Lifestyle parameters and sleep health characteristics are among the features in the dataset meaningful for discovering patterns, with patterns of which can be indicative of existing sleep-related disorders. Based on the models assessed, it was found that the model with the highest performances are bagged models particularly the Voting Classifier with RF and DT. The accuracy, precision, recall, and F1-score of the algorithm were 0.973, suggesting that the algorithm useful for sleep disorder classification and is reliable. These results imply that the proposed machine learning approaches provides an opportunity to make smarter, faster and more accurate sleep disorders diagnoses in order to enhance the possibilities of the physicians' decision-making process and patients' condition.

Index Terms- Machine Learning Algorithms, Deep Learning, Classification, Sleep Disorder, Voting Algorithm.

I. INTRODUCTION

Sleep is a vital physiological function necessary for physical and mental health. It helps strengthen the body and consolidate the brain and memories. Sleep quality significantly affects cognitive functions, especially in children and older adults, who are at an increased risk of accidents. Poor sleep can lead to various health issues, including heart disease, diabetes, and obesity. Despite its importance, sleep disorders are often underdiagnosed or misclassified due to the complexity of sleep stage evaluation. Physicians, medical professionals, and sleep experts must manually analyze polysomnography (PSG) records to assess sleep stages, a task that is prone to human error and is time-consuming for accurate classification [1].

According to the 2021 World Sleep Day survey by Philips, which polled over 13,000 adults in 13 countries, 55% of adults reported being dissatisfied with their sleep. Factors like the COVID-19 pandemic, sleep apnea, and insomnia were identified as key contributors to poor sleep quality. Specifically, 37% of participants mentioned that the pandemic negatively impacted their sleep, while 37% experienced insomnia, 29% snored, 22% had shift-work sleep disorder, and 12% suffered from sleep apnea [2]. These statistics highlight the widespread nature of sleep-related issues and the need for better diagnostic and classification systems.

Medical professionals classify sleep into five distinct stages: wakefulness, N1, N2, N3, and rapid-eye movement (REM). Wakefulness represents the alert state when individuals are aware of their surroundings, with fast and irregular brain waves. N1 is the lightest sleep stage where brain waves slow down, and muscles relax. N2 is a deeper stage, while N3 represents the deepest sleep stage, where awakening is difficult. REM sleep is characterized by rapid eye movements and brain waves similar to those during wakefulness. Each of these stages plays a critical role in the body's restoration and cognitive processes. PSG allows doctors to observe these stages by recording electroencephalogram (EEG) and electrocardiogram (ECG) signals to monitor the brain and body's activity during sleep [3], [4], [5].

To reduce human intervention in the classification and prediction of sleep stages, several researchers have developed automated techniques using machine learning algorithms (MLAs). These methods can be broadly categorized into conventional machine learning and deep learning algorithms. Traditional MLAs, such as support vector machines and decision trees, are typically applied to smaller datasets, requiring manual feature extraction to classify sleep stages based on signals like entropy and energy. In contrast, deep learning algorithms, inspired by the structure of the human brain, utilize neural networks to learn complex patterns from data automatically. These algorithms are especially beneficial for large, complex datasets and are seen as a potential replacement for traditional MLAs [6], [7]. The most common technique for sleep-stage classification involves the application of EEG signals as input for both traditional and deep learning models [8].

II. RELATED WORK

Many research have utilised ML to detect and classify sleep disorders as OSA, sleep stage classification, and ECG-based detection. Machine learning algorithms have greatly improved diagnosis accuracy and reduced time-consuming, error-prone manual analysis. Machine learning predicted obstructive sleep apnoea in Koreans by Kim et al. [9]. Multiple ML systems identified and predicted OSA risk from clinical data. Models suggest ML and automated prediction systems could help doctors discover OSA faster. Demographic, clinical, and physiological variables improved prediction. Early detection with machine learning may improve sleep apnoea diagnosis, says this study.

Mousavi et al. [10] identified single-channel EEG sleep stages using deep CNNs. Diagnose insomnia and sleep apnoea with deep learning's automatic sleep stage classification. Deep learning CNNs can automatically identify EEG hierarchies. CNNs categorise sleep stages faster and more accurately than classical machine learning. Djanian et al. [11] examined consumer sleep technologies and AI for sleep classification. Wearable sleep trackers use AI more. These sensors and AI algorithms analyse sleep quality in real time, helping sleep issues. The study found that AI-based consumer sleep devices, especially deep learning ones, improved sleep stage diagnosis. Wearable AI devices enable personalised sleep health management.

Salari et al. [12] studied ECG-based machine learning sleep apnoea detection. ML algorithms can detect sleep apnoea via ECG signal processing, according to the study. RFs and SVMs can diagnose sleep apnoea from single-lead ECGs. The authors indicated feature extraction methods affect model performance. This study showed how machine learning can improve ECG-based sleep apnoea detection without sleep testing. Li et al. identified sleep stages using deep learning and EEG spectrograms [13]. EEG spectrograms track frequency content throughout time. DNNs categorised these spectrograms. Our deep learning method learnt spectrogram characteristics to accurately discriminate sleep states. This study revealed deep learning helps classify sleep stages when feature engineering fails. Deep learning, especially DNNs, may automate sleep stage recognition well. Han, Oh [14] compared OSA severity prediction machine learning methods. Patient data was used to compare decision trees, random forests, and support vector machines for OSA severity prediction. Random forests predicted OSA severity better than models. Demographic and physiological data can predict sleep apnoea severity, say the authors. This study showed that machine learning can objectively measure sleep disorder severity and help professionals treat it. Bahrami and Forouzanfar [15] used deep learning to detect sleep apnoea from single-lead ECG. We used CNNs and LSTMs to classify sleep apnoea using ECG data. CNNs and LSTMs outperformed machine learning models in sleep apnoea diagnosis. Deep learning methods for real-time ECG data could revolutionise sleep apnoea detection. This study showed that deep learning can handle complex time-series data and boosted ECG data use for non-invasive sleep apnoea detection.

Satapathy et al. [16] contrasted automatic sleep staging machine learning. Sleep stage classification was tested using EEG data feature sets and support vector machines, random forests, and decision trees. Random forest ensembles were most accurate and robust. The authors found that feature selection improves classification and that ML algorithms performed differentially by feature type. This study demonstrated how to stage sleep using machine learning techniques and choose essential variables to

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716558

increase model performance. Bahrami and Forouzanfar [17] identified sleep apnoea from single-lead ECGs using machine learning and deep learning. CNNs, LSTMs, SVMs, and random forests were compared in their extensive study. The survey found CNNs more accurate and sensitive. Deep learning can detect sleep apnoea from ECG signals instead of polysomnography.

III. MATERIALS AND METHODS

We proposed a technique to diagnose sleep problems using machine learning algorithms that were effective given their strengths. Various models are used, including SVM, K-Nearest Neighbours, Decision Tree, Random Forest, and ANN with Multi Layer Perceptron. Sleep Disorder Data provides several sleep health and habit data used to build these algorithms. Voting Classifier uses Decision Tree and Bagging with Random Forest to improve classification accuracy. This ensemble strategy improves model evaluation and performance. To identify sleep disorders like sleep apnoea, the system can manage 400 records with 13 features. The suggested approach will help professionals make sound decisions and improve patient health.

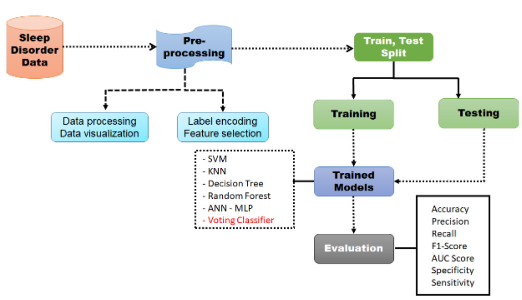


Fig.1 Proposed Architecture

The image depicts a flowchart for a machine learning focused on sleep disorder classification. It starts with raw sleep disorder data, which undergoes pre-processing (data cleaning, visualization, and feature selection). The pre-processed data is then split into training and testing sets. Various machine learning models (SVM, KNN, Decision Tree, Random Forest, ANN-MLP[25], and Voting Classifier) are trained on the training data. The trained models are evaluated on the testing data using metrics like accuracy, precision, recall, F1-score, AUC score, specificity, and sensitivity.

A) Dataset Collection:

The dataset used in this study is the Sleep Health and Lifestyle Dataset, sourced from Kaggle [22]. It contains 400 observations with 13 features related to sleep and daily habits, such as gender, age, occupation, sleep duration, sleep quality, physical activity level, stress level, BMI category, blood pressure, heart rate, daily steps, and sleep disorder. The target variable, "Sleep Disorder," is categorized into three groups: none, sleep apnea, and insomnia. The dataset includes varied occupation data, with the most common being nurse (73 observations), doctor (71), and engineer (63), followed by other occupations such as lawyer (47), teacher (40), and salesperson (32). Pre-processing was done to standardize and replace labels for consistency in analysis.

	Person ID	Gender	Age	Occupation	Sleep Duration	Quality of Sleep	Physical Activity Level	Stress Level	BMI Category	Bl Pres
0	1	Male	27	Software Engineer	6.1	6	42	6	Overweight	12
1	2	Male	28	Doctor	6.2	6	60	8	Normal	12
2	3	Male	28	Doctor	6.2	6	60	8	Normal	12
3	4	Male	28	Sales Representative	5.9	4	30	8	Obese	14
4	5	Male	28	Sales Representative	5.9	4	30	8	Obese	14

Fig.2 Dataset Collection Table

B) Pre-Processing:

In the pre-processing step, data processing involves cleaning and handling missing values, while data visualization helps identify patterns and outliers. Label encoding is applied to categorical variables, and feature selection techniques are used to identify the most relevant features for classification.

i) Data Processing: In the data processing step, duplicate data is identified and removed to ensure data consistency and accuracy. Duplicates can distort the analysis and model performance, so any repeated rows or observations are discarded. Next, drop cleaning is performed by handling missing or null values in the dataset. This involves either removing rows with missing values or imputing them using appropriate methods like mean, median, or mode, depending on the nature of the data. This step ensures that the dataset is clean, reducing noise and improving the quality of the machine learning models.

ii) Data Visualization: Data visualization is a critical step in exploring the dataset and understanding the underlying patterns. By visualizing the data using charts, graphs, and plots, it becomes easier to identify trends, distributions, and potential outliers. Techniques like histograms, box plots, and scatter plots help analyze the distribution of numerical features, while bar plots are useful for categorical variables. Visualizations also help detect correlations between features and identify anomalies that could impact model performance. This step aids in feature engineering and better understanding of the data before applying machine learning algorithms.

iii) Label Encoding: Label encoding is a technique used to convert categorical string data into numeric form. In datasets with categorical variables, such as "Gender" or "Sleep Disorder," machine learning algorithms cannot process string values directly. Therefore, label encoding is applied to convert these string labels into integers. For example, labels like "Male" and "Female" could be converted into 0 and 1, respectively. This step ensures that categorical variables are represented in a format suitable for machine learning models, allowing algorithms to interpret and learn from these features effectively during training and testing.

iv) Feature Selection: Feature selection is the process of identifying and selecting the most relevant features (or variables) for building machine learning models. This step helps reduce the dimensionality of the dataset, improving model efficiency and preventing overfitting. By selecting the appropriate features, the model can focus on the most significant variables that contribute to the target variable. The process involves splitting the dataset into input features (X) and the target variable (y), where X represents the independent variables and y represents the dependent variable. Feature selection techniques, such as correlation analysis and recursive feature elimination, are applied to optimize the dataset.

C) Training & Testing:
The dataset is split into training and testing subsets to evaluate the model's performance. Typically, 80% of the data is used for training the model, while the remaining 20% is reserved for testing. This division helps ensure that the model is trained on a substantial portion of the data and evaluated on unseen data to assess its generalization ability. The split is usually done using techniques like train_test_split from scikit-learn, which ensures random and unbiased division of the dataset into training and testing sets.

D) Algorithms:
Random Forest [20] aggregates multiple decision trees to enhance classification accuracy and robustness. By utilizing random subsets of features and data, it reduces overfitting and improves generalization. This ensemble approach effectively identifies important variables in sleep disorder diagnosis while providing reliable predictions.

$$G = 1 - \sum_{j=1}^{c} p_j^2 \quad (1)$$

Where, p_j is the probability of class j in a node

Support Vector Machine [18] is employed to classify sleep disorders by finding the optimal hyperplane that separates different classes in the dataset. It effectively handles high-dimensional data, making it suitable for identifying complex patterns in sleep patterns and activities, ultimately aiding accurate diagnosis.

$$f(x) = sign(w^T x + b) \quad (2)$$

where:

- w is the weight vector,
- x is the input feature vector,
- b is the bias term,
- sign(·) is the sign function determining the class label.

K-Nearest Neighbours [23] algorithm classifies sleep disorders based on proximity to similar data points in the feature space. By evaluating the nearest neighbors of a given instance, KNN effectively identifies patterns and similarities, providing intuitive insights into sleep disorder classifications based on historical data.

© APR 2026 | IRE Journals | Volume 9 Issue 10 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I10-1716558

Decision Tree [19] algorithm is utilized to create a model that predicts sleep disorder classifications through a series of binary decisions based on input features. Its transparent structure allows for easy interpretation, making it suitable for identifying key factors affecting sleep health and improving diagnosis accuracy.

Artificial Neural Networks [21] are employed to model complex relationships in the dataset, learning from input features to classify sleep disorders. The multilayer perceptron architecture enables the system to capture intricate patterns in sleep data, enhancing prediction capabilities and improving overall diagnostic outcomes.

The Voting Classifier combines predictions from multiple models, including decision trees and bagging classifiers with Random Forest. This ensemble method improves classification accuracy by leveraging the strengths of individual algorithms, reducing errors, and providing a more robust and reliable prediction for sleep disorders.

$$\hat{y} = \arg max \frac{1}{N}\sum_{i=1}^{N} p_i(c|x)(3)$$

IV. RESULTS AND DISCUSSION

Accuracy: The accuracy of a test is its ability to differentiate the patient and healthy cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

$$Accuracy = \frac{\mathrm{TP + TN}}{\mathrm{TP + FP + TN + FN}}(4)$$

Precision: Precision evaluates the fraction of correctly classified instances or samples among the ones classified as positives. Thus, the formula to calculate the precision is given by:

$$Precision = \frac{\text{True Positive}}{\text{True Positive } + \text{False Positive}}(5)$$

Recall: Recall is a metric in machine learning that measures the ability of a model to identify all relevant instances of a particular class. It is the ratio of correctly predicted positive observations to the total actual positives, providing insights into a model's completeness in capturing instances of a given class.

$$Recall = \frac{\mathrm{TP}}{\mathrm{TP + FN}}(6)$$

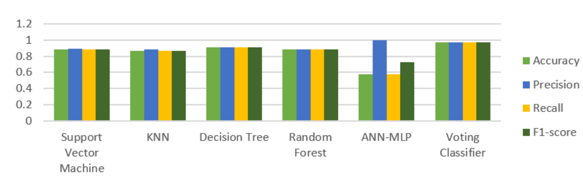
F1-Score: F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

$$F1\ Score = 2 * \frac{Recall\ \mathrm{X}\ Precision}{Recall + Precision} * 100(7)$$

In Table 1, the performance metrics—accuracy, precision, recall, and F1-score—are evaluated for each algorithm. The Voting Classifier achieves the highest scores, with all metrics at 0.973. Other algorithms' metrics are also presented for comparison.

Table.1 Performance Evaluation Metrics

ML Model	Accuracy	Precision	Recall	F1-score
Support Vector Machine	0.880	0.892	0.880	0.884
KNN	0.867	0.883	0.867	0.867
Decision Tree	0.907	0.909	0.907	0.908
Random Forest	0.880	0.887	0.880	0.881
ANN-MLP	0.573	1.000	0.573	0.729
Voting Classifier	0.973	0.973	0.973	0.973



Graph.1 Comparison Graphs

In Graph 1, accuracy is represented in light green, precision in blue, recall in light yellow, and F1-score in green. The Voting Classifier outperforms the other algorithms in all metrics, with the highest values compared to the remaining models. These details are visually represented in the above graph.

V. CONCLUSION

This paper has established that with the help of machine learning algorithms it is indeed possible to classify sleep disorders by using data from Sleep Disorder Data set that is made available to the public. Comparing a range of deep learning and classical machine learning methods, the Voting Classifier, based on bagging with RF and DT showed the highest accuracy. Performance was satisfactory in all evaluation matrices with an accuracy of 97.3%, precision of 97.3%, recall of 97.3% and F1-score of 97.3%. These results suggest further that the Voting Classifier developed here is a highly reliable and robust system for classifying sleep disorders. These results showing steady improvements in almost all the tested measurement indicate that the model could be useful in providing precise and timely diagnosis of sleep disorders thus benefiting the patient and enhancing the clinical decisions. Based on the high classification accuracy, the Voting Classifier can be recommended as the useful instrument to automate the process of Sleep disorders' diagnostic, thus promoting more accurate diagnoses and better prognosis in individuals with Sleep disorders.

The *future scope* of this research includes the exploration of additional advanced machine learning techniques, such as deep learning architectures like Convolutional Neural Networks (CNNs) and recurrent networks for improved accuracy in sleep disorder classification. Integrating real-time data from wearable devices could enhance the system's predictive capabilities. Furthermore, expanding the dataset to include diverse populations and sleep disorders will improve model generalization.

REFERENCES

[1]F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, ''A portable wireless device for cyclic alternating pattern estimation from an EEG monopolar derivation,'' Entropy, vol. 21, no. 12, p. 1203, Dec. 2019.

[2] Y. Li, C. Peng, Y. Zhang, Y. Zhang, and B. Lo, ''Adversarial learning for semi-supervised pediatric sleep staging with single-EEG channel,'' Methods, vol. 204, pp. 84–91, Aug. 2022.

[3] E. Alickovic and A. Subasi, ''Ensemble SVM method for automatic sleep stage classification,'' IEEE Trans. Instrum. Meas., vol. 67, no. 6, pp. 1258–1265, Jun. 2018.

[4] D. Shrivastava, S. Jung, M. Saadat, R. Sirohi, and K. Crewson, ''How to interpret the results of a sleep study,'' J. Community Hospital Internal Med. Perspect., vol. 4, no. 5, p. 24983, Jan. 2014.

[5] V. Singh, V. K. Asari, and R. Rajasekaran, ''A deep neural network for early detection and prediction of chronic kidney disease,'' Diagnostics, vol. 12, no. 1, p. 116, Jan. 2022.

[6] J. Van Der Donckt, J. Van Der Donckt, E. Deprost, N. Vandenbussche, M. Rademaker, G. Vandewiele, and S. Van Hoecke, ''Do not sleep on traditional machine learning: Simple and interpretable techniques are competitive to deep learning for sleep scoring,'' Biomed. Signal Process. Control, vol. 81, Mar. 2023, Art. no. 104429.

[7] H. O. Ilhan, ''Sleep stage classification via ensemble and conventional machine learning methods using single channel EEG signals,'' Int. J. Intell. Syst. Appl. Eng., vol. 4, no. 5, pp. 174–184, Dec. 2017.

[8] Y. Yang, Z. Gao, Y. Li, and H. Wang, ''A CNN identified by reinforcement learning-based optimization framework for EEG-based state evaluation,'' J. Neural Eng., vol. 18, no. 4, Aug. 2021, Art. no. 046059.

[9] Y. J. Kim, J. S. Jeon, S.-E. Cho, K. G. Kim, and S.-G. Kang, ''Prediction models for obstructive sleep apnea in Korean adults using machine learning techniques,'' Diagnostics, vol. 11, no. 4, p. 612, Mar. 2021.

[10] Z. Mousavi, T. Y. Rezaii, S. Sheykhivand, A. Farzamnia, and S. N. Razavi, ''Deep convolutional neural network for classification of sleep stages from single-channel EEG signals,'' J. Neurosci. Methods, vol. 324, Aug. 2019, Art. no. 108312.

[11] S. Djanian, A. Bruun, and T. D. Nielsen, ''Sleep classification using consumer sleep technologies and AI: A review of the current landscape,'' Sleep Med., vol. 100, pp. 390–403, Dec. 2022.

[12] N. Salari, A. Hosseinian-Far, M. Mohammadi, H. Ghasemi, H. Khazaie, A. Daneshkhah, and A. Ahmadi, ''Detection of sleep apnea using machine learning algorithms based on ECG signals: A comprehensive systematic review,'' Expert Syst. Appl., vol. 187, Jan. 2022, Art. no. 115950.

[13] C. Li, Y. Qi, X. Ding, J. Zhao, T. Sang, and M. Lee, ''A deep learning method approach for sleep stage classification with EEG spectrogram,'' Int. J. Environ. Res. Public Health, vol. 19, no. 10, p. 6322, May 2022.

[14] H. Han and J. Oh, ''Application of various machine learning techniques to predict obstructive sleep apnea syndrome severity,'' Sci. Rep., vol. 13, no. 1, p. 6379, Apr. 2023.

[15] M. Bahrami and M. Forouzanfar, ''Detection of sleep apnea from single-lead ECG: Comparison of deep learning algorithms,'' in Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA), Jun. 2021, pp. 1–5.

[16] S. Satapathy, D. Loganathan, H. K. Kondaveeti, and R. Rath, ''Performance analysis of machine learning algorithms on automated sleep staging feature sets,'' CAAI Trans. Intell. Technol., vol. 6, no. 2, pp. 155–174, Jun. 2021.

[17] M. Bahrami and M. Forouzanfar, ''Sleep apnea detection from single-lead ECG: A comprehensive analysis of machine learning and deep learning algorithms,'' IEEE Trans. Instrum. Meas., vol. 71, pp. 1–11, 2022.

[18] J. Ramesh, N. Keeran, A. Sagahyroon, and F. Aloul, ''Towards validating the effectiveness of obstructive sleep apnea classification from electronic health records using machine learning,'' Healthcare, vol. 9, no. 11, p. 1450, Oct. 2021.

[19] S. K. Satapathy, H. K. Kondaveeti, S. R. Sreeja, H. Madhani, N. Rajput, and D. Swain, ''A deep learning approach to automated sleep stages classification using multi-modal signals,'' Proc. Comput. Sci., vol. 218, pp. 867–876, Jan. 2023.

[20] O. Yildirim, U. Baloglu, and U. Acharya, ''A deep learning model for automated sleep stages classification using PSG signals,'' Int. J. Environ. Res. Public Health, vol. 16, no. 4, p. 599, Feb. 2019.

[21] S. Akbar, A. Ahmad, M. Hayat, A. U. Rehman, S. Khan, and F. Ali, ''IAtbP-Hyb-EnC: Prediction of antitubercular peptides via heterogeneous feature representation and genetic algorithm based ensemble learning model,'' Comput. Biol. Med., vol. 137, Oct. 2021, Art. no. 104778.

[22] (2023). Sleep Health and Lifestyle Dataset. [Online]. Available: http://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset

[23] P. Tripathi, M. A. Ansari, T. K. Gandhi, R. Mehrotra, M. B. B. Heyat, F. Akhtar, C. C. Ukwuoma, A. Y. Muaad, Y. M. Kadah, M. A. Al-Antari, and J. P. Li, ''Ensemble computational intelligent for insomnia sleep stage detection via the sleep ECG signal,'' IEEE Access, vol. 10, pp. 108710–108721, 2022.

[24] Y. You, X. Zhong, G. Liu, and Z. Yang, ''Automatic sleep stage classification: A light and efficient deep neural network model based on time, frequency and fractional Fourier transform domain features,'' Artif. Intell. Med., vol. 127, May 2022, Art. no. 102279.

[25] I. A. Hidayat, ''Classification of sleep disorders using random forest on sleep health and lifestyle dataset,'' J. Dinda : Data Sci., Inf. Technol., Data Anal., vol. 3, no. 2, pp. 71–76, Aug. 2023.