International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Applying Machine Learning Algorithms for the Classification of Sleep Disorders

GOLI KARTHEEK
Student
Department of Computer Science
Sreenidhi Institute of Science and Technology
Ghatkesar, Hyderabad, Telanagana
22311a05k1@cse.sreenidhi.edu.in

MUTHYALA VINAY GOUD
Student
Department of Computer Science
Sreenidhi Institute of Science and Technology
Ghatkesar, Hyderabad, Telanagana
22311a05l8@cse.sreenidhi.edu.in

VADLA SAI KALYAN
Student
Department of Computer Science
Sreenidhi Institute of Science and Technology
Ghatkesar, Hyderabad, Telangana
22311a05m3@cse.sreenidhi.edu.in

DEEPA PANSE
Assistant Professor
Department of Computer Science
Sreenidhi Institute of Science and Technology
Ghatkesar, Hyderabad, Telangana
deepa.p@sreenidhi.edu.in

Abstract: Sleep issues, particularly sleep apnea, can cause people to get ill and that is why it is important to have the correct diagnosis. Instead, sleep experts have complex and time-intensive methods of sorting out the various sleep stages manually. This paper introduces a ML classification algorithm based on the publicly available Sleep Disorder Data, comprising 400 records and 13 characteristics. A number of profound and technique-based ML models are discussed, and their effectiveness in accurately detecting sleep disorders is evaluated. Some of the key aspects in the dataset that can aid in finding patterns are lifestyle factors and sleep health indicators. These tendencies may indicate that a person has a sleeping problem. The analysis of the models showed that bagged models (the Voting Classifier with RF and DT) had the best performance. Accuracy, precision, recall, and F1-score of the algorithm are all 0.973 which implies that the algorithm is useful in classifying the sleep disorders and is reliable. These results suggest that the suggested ML methods offer a chance to create sleep problem diagnoses that are smarter, quicker, and more accurate. This would enhance decision-making and health of the doctors and patients.

***"Index Terms -** Machine learning algorithms, deep learning, classification, sleep disorder, Voting algorithm."*

I. INTRODUCTION

Sleep is a necessity to both physical and mental health. It gives mental and memory assistance and physical energy. Quality of sleep affects thinking, especially in youngsters and elderly individuals, who are more prone to harm. Sleep deprivation may cause heart disease, diabetes and obesity. This complicates measuring sleep phases, thus making sleep issues hard to identify, yet very critical. In order to evaluate the stage of sleep, doctors, nurses, and sleeping professionals have to manually scan the polysomnography (PSG) findings. It is difficult to be messed up and time consuming [1].

Philips found 55% of 13,000 individuals in 13 nations were unhappy with their sleep in 2021. COVID-19, sleep apnea, and insomnia influenced the quality of sleep. Specifically, 37 percent of them had sleeplessness, 29 percent snoring, 22 percent shift-work sleep difficulty, and 12 percent sleep apnea as a result of the pandemic [2]. These statistics indicate that issues with sleep are widespread and they need a better method of diagnosis and categorization.

Doctors have categorized sleep into five stages: awareness, N1, N2, N3, and REM. A vigilant individual senses his environment and has irregular, rapid brain waves. The lightest sleep stage is N1. The slowing of brain waves and relaxation of the muscles. N2 is more powerful and N3 is more profound, it is difficult to awaken. REM sleep contains rapid eye movements and brain waves such as during wakefulness. All these steps play a crucial role in the repair of the body and thinking. These stages can be observed by physicians with the help of PSG that records EEG and ECG measurements to examine the health of the brain and its body during sleep [3, 4, and 5].

Automatic sleep phase classification and predictions were done by several researchers with the help of MLAs. We can classify them as DL and classic ML algorithms. Smaller datasets are handled by traditional MLAs such as support machines, and decision trees. Entropy and energy signals are used to identify sleep phases by extracting characteristics manually by MLA. In contrast, DL algorithms use neural networks to automatically learn complicated patterns from data. They are founded on brain functioning. Because of their capability to handle large and complex data, conventional MLAs might be unnecessary [6], [7]. The most popular sleep-stage categorization method uses EEG data to feed classical and deep learning models [8].

II. RELATED WORK

Many research have used ML to categorize sleep disorders such OSA, sleep stages, and ECGs. ML algorithms have significantly boosted the accuracy of diagnoses and reduced manual analysis, which is tiresome and error-prone. Kim et al. [9] made predictions of Korean OSA using ML. OSA risk prediction was carried out with a set of ML systems based on clinical data. It has been simulated that ML and automated prediction systems can assist doctors to diagnose OSA in less time. Predictions were improved with demographic, clinical and physiological parameters. This article demonstrates that ML can be used to diagnose sleep apnea at an early age.

Mousavi et al. [10] detected single-channel EEG sleep phases using deep CNNs. EEG hierarchies may be used by DL CNNs to diagnose insomnia or sleep apnea. Compared to conventional ML, CNNs classify the sleep stages quicker. Djanian et al. [11] investigated the use of AI and sleep solutions by customers to get a better sleep. There is an increasing use of wearable sleep monitors based on AI. The actual sleep quality is monitored by the sensors and AI algorithms and can help with sleep disorders. The researchers concluded that AI-based consumer sleep devices and in particular deep learning devices improved sleep phase recognition. Sleep health is possible with wearable AI devices. Salari et al. [12] studied sleep apnea ML with the use of ECG. The article asserts that ECG readings can be used to detect sleep apnea using ML algorithms. Single-lead ECGs can be used to detect sleep apnea using RFs and SVMs. The authors indicated that the performance of the model can be influenced by feature extraction procedure. This experiment discovered that, with no sleep testing, ML can improve the detection of sleep apnea using ECG. Li et al. characterized sleep phases with the help of DL and EEG spectrogram [13]. EEG spectrograms exhibit alterations in frequency across time. DNNs group spectrograms. The spectrogram features were used to train our DL model to identify sleep phases. In this study, feature engineering was unsuccessful in the classification of sleep stages, so DL was used to aid this classification. DNNs and DL could be applied in identifying sleep stages. Han and Oh [14] compared OSA severity prediction ML algorithms. We used patient data to evaluate the DT, RF, and SVM to find which predicted OSA severity best. RF was more predictive of OSA than models. Through the authors, demographic and physiological factors indicate the severity of sleep apnea. This study reported that ML can effectively evaluate the severity of sleep problems and help the experts to treat them. Bahrami and Forouzanfar [15] determined sleep apnea in the single-lead ECG using DL. Sleep apnea was recognized by CNNs and LSTMs using ECG data. CNNs and LSTMs were superior to ML models in the diagnosis of sleep apnea. DL of real-time ECG data can alter the detection of sleep apnea. This article demonstrated that DL is capable of dealing with more complex time-series data and improving the use of ECG data to detect non-invasive sleep apnea.

Satapathy et al. [16] compared sleep stages ML. The feature sets of EEG data, support vectors, random forests, and decision trees were tested on the categorization of sleep stages. Random forest ensembles were most accurate and reliable. The authors identified that feature selection enhances classification and that the performance of the ML algorithms is influenced by the type of features. This study demonstrates how to stage sleep with the help of ML techniques and select the key factors to maximize model performance. Bahrami and Forouzanfar [17] discovered sleep apnea with the use of ML and DL single-lead ECGs. CNNs, LSTMs, SVMs, and random forests were compared in their extensive study. The survey showed that CNNs were more accurate compared to polysomnography in identifying sleep apnea in ECG signals.

III. MATERIALS AND METHODS

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Our ML algorithm method of sleep disorder detection was effective because of its advantages. Other models are SVM, KNN, DT, RF and ANN with Multi Layer Perceptron. Sleep Disorder Data provides a lot of sleep and health data for these algorithms. Voting Classifier employs DT and Bagging with Random Forest in order to enhance classification. Testing and improving models is easier using this group technique. With 13 criteria, the technology has the ability to detect sleep disorders such as sleep apnea using 400 records. The specified strategy can help practitioners make informed decisions and promote patient well-being.

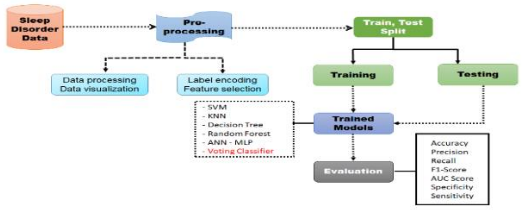


Fig.1 Proposed Architecture

The picture shows a flowchart for a MLprogram that focuses on classifying sleep disorders. It starts with raw data on sleep disorders which is cleaned, visualized and features selected before being used. After being pre-processed, the data is divided into sets for training and testing. It trains various ML models on the training data, including SVM, KNN, DT, RF, ANN-MLP[25] and Voting Classifier. Accuracy, precision, recall, F1-score, AUC score, specificity, and sensitivity are some of the measures we apply to test our trained models using the testing data.

A) Dataset Collection:

The Sleep Health and Lifestyle Dataset of Kaggle [22] was used in this research. There are 400 observations and 13 aspects on sleep and everyday habits including gender, age, job, length of time you sleep, how well you sleep, how active you are, how stressed you are, BMI category, blood pressure, heart rate, daily steps and your sleep problem. "Sleep Disorder," the target variable, is divided into three groups: none, sleep apnea, and insomnia. The data contains data on numerous occupations. The most prevalent ones are nurse (73 observations), doctor (71), and engineer (63). There are also other occupations such as lawyer (47), teacher (40), and salesperson (32). The standardization and updating of labels before the analysis was required to have uniform results.

	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:

Pre-processing includes cleaning and dealing with missing statistics, and uses data visualization to identify patterns and outliers. Label encoded is categorical variables. Using feature selection, the most significant categorization criteria are found.

i) Data Processing: To ensure accuracy and duplicates, data processing eliminates duplicates. Duplicates can impact analysis and model performance, therefore, duplicate rows and observations are eliminated. Second, drop cleaning eliminates the nulls or missing values in the data. Removal of rows with missing values or the filling of the rows with the mean, median, or mode can be done depending on the type of data. This stage cleans data, decreasing noise and enhancing ML models.

ii) Data Visualization: Visualization of data assists to determine trends in the data set. Data charts, graphs and plots are used to identify trends, distributions and outliers. The numerical distribution is depicted in histograms, box plots, and scatter plots. Numbers are revised using bar graphs. The visualizations can also demonstrate the correlation and outliers of the features that could cause poor performance of the model. The process enhances the understanding of data and feature engineering prior to utilizing the ML algorithms.

iii) Label Encoding: Label encoding transforms text data that is categorical into an integer. String values in categorical datasets, such as Gender or Sleep Disorder, cannot be directly used in ML algorithms. Label encoding is a process of converting text labels into numbers. You can change "Male" and Female to 0 and 1. This step ensures categorical variables are encoded in a manner that ML models can read and learn about such variables during training and testing.

iv) Feature Selection: The process of selecting the most important features (or variables) to use in ML models is referred to as feature selection. This helps to minimize the size of datasets, enhancing the performance of the models and avoiding overfitting. The choice of the appropriate features enables the model to concentrate on the most significant variables that influence the target variable. The input features (X) and the target variables (y) are divided into the data sets, where X is a set of independent variables and y is a dependent variable. Correlation analysis and recursive feature reduction are used to select the best features to improve the dataset.

C) Training & Testing:

To determine the performance of the model, the data is divided into two; training and testing. In most cases, 80% of the data is used to train the model, while the other 20% is utilized to test it. This split helps make sure that the model gets trained on a lot of the data and tested on data it hasn't seen before to determine how well it generalizes. The division is usually performed using such tools as train-test-split of scikit-learn that ensures that the dataset is divided into training and testing in a random and unbiased manner.

D) Algorithms:

Random Forest [20] is a combination of numerous decision trees that make the categorization more accurate and reliable. It reduces overfitting as well as enhances generalization by employing random groups of characteristics and data. This team approach does well in discovering crucial variables in diagnosing sleep disorders and effective predictions.

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

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

Support Vector Machine [18] The sorting of sleep disorders is done with the help of Support Vector Machine [18] which is used to determine the most suitable hyperplane to subdivide the various classes in the data set. It can be applied to high dimensional data, making it an effective means of discovering complex trends in sleep patterns and activities, which will aid in proper diagnosis.

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

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] classifier ranks sleep disorders based on their proximity to other data points in the feature space. KNN finds patterns and similarities by looking at the nearest neighbors of a particular instance. This provides us with a simple method of categorizing the sleep disorders in relation to the historical information.

A model predicting the categories of sleep disorders is built through the Decision Tree [19] technique, by making a series of binary decisions, based on input features. Its simple structure allows one to comprehend it easily thus making it suitable in searching some important things that can affect sleep health and diagnoses more precise.

ANNs[21] are employed in modeling complex interactions in the dataset and training how to classify sleep disorders with the input features. The multilayer perceptron architecture allows the system to detect details in complex patterns of sleep data, which enhances its performance in terms of making predictions and overall diagnostic outcomes.

The Voting Classifier combines the predictions of multiple models, including the Random Forest, decision trees and bagging classifiers. This combination method allows the classification to be more precise, utilising the advantages of each algorithm and reducing errors, as well as providing a more reliable and robust prediction of 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 the ability of a test to distinguish between sick and healthy individuals. To gain an idea of the accuracy of a test we should figure out the percentage of true positive and true negative cases of all the examples we considered. Mathematically, this can be stated as:

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

Precision: Precision is used to determine the number of the samples or instances labeled as positive that were really correct. The equation to determine the accuracy is, therefore,:

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

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Recall: In machine learning, recall is a measure of how well a model can find all the examples of a certain class that are important. It is a proportion of the positives that are predicted correctly to the total true positives. This will give you some notion on how good a model is at capturing the examples of a particular class.

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

F1-Score: F1 score is a metric that can be used to measure the accuracy of a ML model. It adds the accuracy and recall scores of the model. The statistic of accuracy is the number of times that a model made a valid guess in the entire dataset.

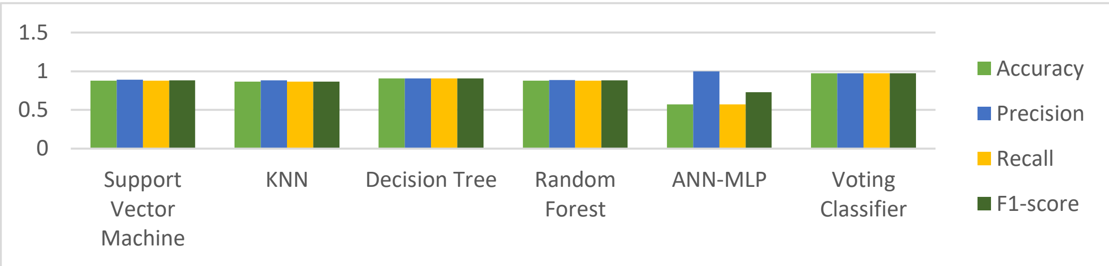
$$F1\ Score = 2 * \frac{Recall\ \mathrm{X}\ Precision}{Recall + Precision} * 100 (7)$$

The results in Table 1 indicate the accuracy, precision, recall and F1-score of each method. The Voting Classifier has the highest scores with all measures at 0.973. We also demonstrate the statistics of other algorithms to be able to compare them.

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



Graph 1 depicts the accuracy, precision, recall and F1-score in light green, blue, light yellow and green respectively. In all criteria, the Voting Classifier does better than the other methods, with the highest values compared to the other models. The above graph presents these facts in an easy to understand manner.

V. CONCLUSION

This paper has shown that ML algorithms can be used to successfully detect sleep disorders based on publicly accessible data on the Sleep Disorder Data set. The Voting Classifier, which is based on bagging with RF and DT, has the best accuracy when compared to a number of DL and traditional ML approaches. The results were good in all the assessment matrices with accuracy of 97.3, precision of 97.3, recall of 97.3 and F1-score of 97.3. These findings also suggest that the Voting Classifier created in this paper is an extremely solid and strong tool to detect sleep disorders. The findings reveal the overall improvement of almost all the measures that are assessed, implying that the model can be used to effectively and timely identify sleep issues, thus benefiting the patient and enhancing clinical judgment. The Voting Classifier may be suggested as a useful tool to automate the process of diagnosing sleep disorders since it has a high accuracy rate. This will result in better diagnoses and a better outcome to those with sleeping disorders.

Future directions of the research involve exploring additional advanced ML algorithms, such as DL systems, such as CNNs and recurrent networks, to improve the accuracy of sleep disorder classification. The system could be improved with real-time information provided by wearable devices, which would help the system to make more predictions. Also, it would be better to expand the dataset to include diverse demographics and sleeping issues to better generalize the model.

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.