ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

Making Use of Machine Learning Algorithms to Categorize Sleep Disorders

[1]Mrs.G.Haritha Rani, [2]Jagarapu Loka Kalyani,

[1] Associate Professor,Dept.of MCA, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

[2]Students,Dept.of CSE, Rajamahendri Institute of Engineering & Technology, Bhoopalapatnam, Near Pidimgoyyi,Rajahmundry,E.G.Dist.A.P. 533107.

ABSTRACT

Classifying sleep disorders is essential for bettering people's lives. Apnea and other sleep problems may significantly impact people's well-being. Classifying sleep stages by professionals is a challenging and error-prone process. It is necessary to analyze, monitor, and diagnose sleep problems in order to build reliable machine learning algorithms (MLAs) for sleep disorder categorization. In order to categorize sleep disorders, this research contrasts traditional MLAs with deep learning methods. This work presents a strategy for the optimal classification of sleep disorders and tests it using the publically accessible Sleep Health and Lifestyle Dataset. To achieve these improvements, many machine learning algorithms had their parameters fine-tuned using a genetic algorithm. Assessing and contrasting the suggested approach with cutting-edge machine learning techniques for sleep problem classification. There are thirteen columns and four hundred rows in the dataset, and the attributes reflect different aspects of sleep and everyday life. Various deep learning techniques were evaluated, including k-nearest neighbours, support vector machine, decision tree, random forest, and artificial neural network (ANN). The experimental findings show that the algorithms that were tested had vastly different performance. The classification accuracy that each of the suggested methods achieved was 83.19%, 92.04%, 88.50%, 91.15%, and 92.92%, correspondingly. With a recall of 91.93%, a precision of 92.01%, and an F1-score of 93.80% on the test data, the ANN attained the best classification accuracy of 92.92%. A natural language processing algorithm that outperformed its competitors in terms of accuracy. KEY WORDS: genetic algorithm, sleep disorder, classification, machine learning, deep learning.

I. INTRODUCTION

In order to maintain good mental and physical health, sleep is an essential physiological process. A good night's sleep fortifies the body and helps memories and cognition stick. Children and elderly drivers are more vulnerable to accidents due to poor sleep quality, which impacts cognitive processes. The human body is susceptible to the negative effects of sleep deprivation, which may lead to conditions such as diabetes, obesity, and heart disease. The manual evaluation of polysomnography (PSG) recordings by physicians, specialists, and other medical personnel might result in varying sleep stage judgments. Classifying sleep stages manually is laborious, error-prone, and time-consuming. [1, 2]. On every year around World Sleep Day, Philips polls people about their thoughts and actions on sleep. More than 13,000 individuals from 13 different countries were surveyed in 2021. While 55% of individuals reported being content overall with their sleep, the remaining 47% reported being unhappy with the quality of their sleep. The 2019 coronavirus disease (COVID-19) pandemic, sleep apnea, and insomnia all contributed to poor sleep quality for them. Among those who reported trouble sleeping, 37% stated it was because of the epidemic. In addition, sleep apnea affects 12% of people, shift work sleep problem affects 22%, snoring affects 29%, and insomnia affects 37% of people [1, 2]. Sleep specialists and medical professionals assess the quality of sleep by analyzing the sleep system, which is categorized for different phases of sleep. Rapid eye movement (REM) is one of five phases of sleep, the others being waking, non-REM (N1), and REM II. The state of awareness known as wakefulness occurs when people are fully cognizant of their immediate environment. While we are aware, our brain waves are rapid and erratic. Slow brain waves and relaxed muscles characterize the first stage of sleep, N1. Also, when people are sleeping, they are in a deep stage of sleep called N2, and the deepest stage of sleep, N3, is even harder to rouse them from. While you're in rapid eye movement (REM), your brain waves are quite comparable

ISSN 2454-9940
www.ijasem.org
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
Vol 19, Issue 2, 2025
to when you're awake. For various reasons, each stage of the sleep cycle is essential. Even while you're asleep, your brain and body continue to work at a high level. Therefore, PSG allows medical professionals to monitor the patient's vital signs by recording electroencephalogram (EEG) and electrocardiogram (ECG) signals [3, 4], [5]. A number of academics have come up with methods to automate common operations with little to no human involvement. These methods use categorization and prediction algorithms to identify patterns. There are two main categories into which these methods fall: deep learning algorithms and regular machine learning algorithms. The training dataset for traditional MLAs is quite modest, yet they are simple, rapid, and easy to deploy. In order to classify the different phases of sleep, the feature engineering method manually collects properties from the signals, such energy and signal entropy. As a kind of biologically inspired MLA, deep learning algorithms use neural networks to discover intricate patterns in data in an effort to imitate the human brain. Conventional AI will soon be supplanted by deep learning algorithms. Automatic feature engineering is a byproduct of deep learning, which encompasses all algorithms that use layers to interpret data [6, 7]. Classification jobs with large amounts of data or complicated characteristics are ideal for deep learning models. Using an electroencephalogram (EEG) as input is the most used method for sleep-stage categorization [8]. The authors of this paper compile previous work on sleep disorders and analyze it, paying special attention to the difficulties of data collecting from patients in different institutions, which might contain noisy and ambiguous information (such as missing data). The dataset is severely lacking in depth since it is based on data acquired from a single sleep clinic. The data is skewed towards particular patient groups, making it difficult to generalize assessed outcomes. Additionally, the biased data might lead to erroneous results, which can impact decision-making. On the other hand, natural sleep-stage datasets are scarce [9]. In addition, choosing appropriate MLAs from various classifiers typically necessitates more computational effort due to the need to extract features from the dataset in order to train models and choose discriminative features [10]. Aiming to alleviate the difficulties that persons with sleep problems face in today's fast-paced world is what prompted this research. The effects of contemporary lives and people's disregard for this essential requirement have increased the gravity of the threat posed by sleep disorders and illnesses connected to them. One of the most fundamental need for human survival is sleep. Ensuring human well-being and quality of life requires the use of machine learning algorithms to categorize sleep disorders. The MLAs have been adopted for sleep disorder classification, but to the authors? knowledge, there is a lack of comprehensive reviews of such MLAs in this sector. The essay contributes in two ways: 1) A survey of previous work on sleep disorder classification; 2) An analysis of how well the suggested algorithm performs when compared to state-of-the-art ML algorithms trained with default parameters for sleep disorder classification; and 3) A thorough review of previous work on traditional MLAs trained with deep learning algorithms. Following is the outline of the paper. This paper's assessment technique and the state-of-the-art MLAs are detailed in Section III, while Section II covers the relevant work. Section IV then presents the findings and examines the approaches' efficacy in classifying sleep disorders. The article is concluded in Section V, which also covers the anticipated future development for this application.

II. RELATED WORK

In [11], the researchers looked examined a number of studies that classified sleep using consumer sleep technology (CST) and MLAs. They acknowledged that PSG is a crucial standard, but that manual techniques including specialized controller settings to categorize sleep phases are costly and difficult to adapt. While CST has been used for sleep tracking, PSG outperforms CST when it comes to identifying the different phases of sleep. Logistic regression (LR), decision trees (DT), support vector machines (SVMs), and deep learning were among the several MLAs examined in the 27 publications included in the study. The models have the potential to greatly enhance the precision of sleep-stage categorization using CST. Nevertheless, the use of raw data with deep learning algorithms is still in its infancy. In another piece, the importance and difficulties of sleep apnea were explored after reviewing 48 studies [12]. Also, deep learning methods, random forest (RF), support vector machines (SVMs), and other machine learning algorithms may be used to identify cases of sleep apnea using electrocardiogram (ECG) data. They did, however, point out a few problems with using MLAs for sleep apnea classification, such as discrepancies in ECG signals and a lack of accessible datasets to train the models. Deep learning-based neural networks and support vector machines (SVMs) were the most effective at identifying sleep apnea from electrocardiogram (ECG) data in their research. In order to categorize the sleep state from an EEG spectrogram, the authors in [13] used MLAs. More time is required for the categorization of sleep phases. Using MLAs with EEG data for categorization, it is inaccurate. In addition, the data are imbalanced, which leads to poor accuracy. For this purpose, they tested their models on four publicly available datasets. A classification accuracy of 94.17%, 86.82%, 83.02%, and 85.12% were achieved by the suggested methods across all four datasets, according to the findings. They built a deep learning model and used deep learning algorithms to categorize the phases of sleep. The EEG spectrogram was processed using convolutional

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

neural networks (CNNs) to extract time and frequency information. A key technique for identifying different phases of sleep using electroencephalograms (EEG) spectrograms, the model incorporates several hidden layers of bidirectional long short-term memory (LSTM) to identify prediction sequences.

TABLE 1. **A summary of the algorithm, dataset and accuracy in some of the reviewed studies is presented.**

Ref.	Year	Algorithm used	Accuracy	Datasets	Available	Real
[13]	2022	CNN	94.17%, 86.82%, 83.02%, 85.12%	Sleep-EDFX-8, Sleep-EDFX-20, Sleep-EDFX-78, and SHHS	Yes	Yes.
[14]	2023	gradient-boost and RFand KNN.	88%, 88%, 91%.	Medical Centre	No	Yes.
[15]	2021	CNN, LSTM, Bidirec-tional LSTM and Gated recurrent unit (GRU) .	80.67%, 75.04%, 84.13%, 74.72%.	PhysioNet Apnoea-ECG Database	Yes	Yes.
[16]	2021	DT, KNN, RF.	89.10%, 89.10%, 94.46%.	ISRUC%-Sleep database .	Yes	Yes.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

[17]	2022	CNN, LSTM, MLP.	the highest is hybrid deep models 88.13%.	The-PhysioNet ECG Sleep Apnoea v1.0.0 dataset.	Yes	Yes.
[18]	2021	XGB, LGBM, CB, RF, KNN, LR and SVM.	the highest is SVM 68.06%.	the The Wisconsin Sleep Cohort dataset.	Yes	Yes.
[19]	2023	CNN+LSTM, RF, KNN and SVM.	87.4%, 74.07%, 83.65%, 76.04% .	the The Wisconsin Sleep Cohort dataset.	Yes	Yes.
[20]	2019	CNN.	98.06%	sleep-edf and sleep-edfx.	Yes	Yes.

datasets including peptides (AtbPs) were used to assess the suggested method. Outperforming competing algorithms, their suggested "iAtbP-Hyb-EnC" approach achieved 94.47% and 92.68% prediction accuracy, respectively.

III. METHODOLOGY

A. MATERIALS AND METHODS

The implementation of deep learning algorithms and traditional MLAs for the classification of sleep disorders is the main topic of this section. The datasets, performance metrics, and feature significance approach used to evaluate the models and evaluate the suggested algorithms are detailed in the following sections. Furthermore, a concise explanation of the categorization method used in this study is provided.

TABLE 2. Detailed information about the sleep health and lifestyle database records in this study.

ID	Gen	Age	Occu	Sle Dur	Q of Sle	Phys Act	Str Lev	BMI Cat	Blood Pr	HR	DS	Sleep Dis-or-der
1	M	27	SW	6.1	6	42	6	Overw	126/83	77	4200	None
2	M	28	DR	6.2	6	60	8	Normal	125/80	75	10000	None
3	M	28	DR	6.2	6	60	8	Normal	125/80	75	10000	None
4	M	28	Sal	5.9	4	30	8	Obese	140/90	85	3000	Apnoea
5	M	28	Sal	5.9	4	30	8	Obese	140/90	85	3000	Apnoea
6	M	28	SW	5.9	4	30	8	Obese	140/90	85	3000	Insomnia
7	M	29	Teac	6.3	6	40	7	Obese	140/90	82	3500	Insomnia
8	M	29	DR	7.8	7	75	6	Normal	120/80	82	8000	None

In order to fix the optimization issues with the classifiers, GA was used for feature selection. The classifiers may be fine-tuned by adjusting their various parameters. In order to get the most out of the suggested model, GA was used to fine-tune the parameters. Figure 2 is a high-level depiction of how a genetic algorithm is put into action. Here is how the algorithm that was suggested works: First, the population is created at random. The second step is to determine how well a collection of parameters (a potential solution) performs by calculating a fitness value. The next step is to choose parents that have healthy, fit people.In the fourth step, known as "crossover," two parents are

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

combined to produce a new person. Fifth, perform mutations to alter the DNA in an unpredictable way. Step 6: Keep going until the end requirements are satisfied. in order to classify sleep disorders and choose features using MLAs.

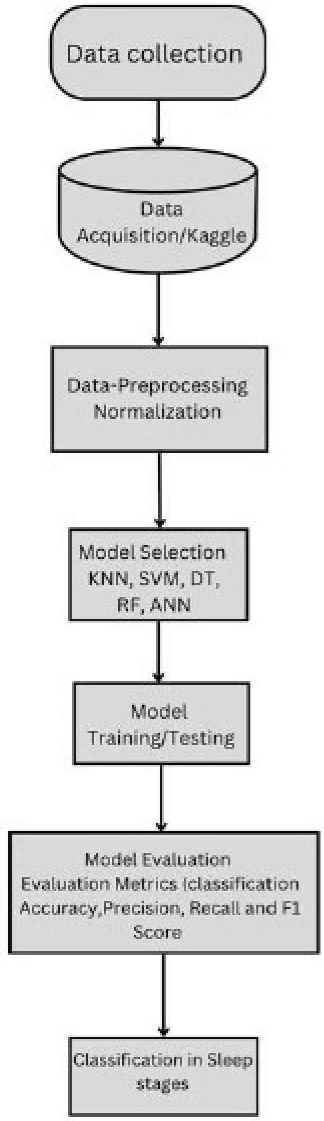


FIGURE 1. Diagram of the machine learning model to classify sleep disorders.

PERFORMANCE METRICS

Within the context of sleep disorder categorization, this research assesses and confirms the efficacy of the suggested model. Not to mention the percentage of tasks completed

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

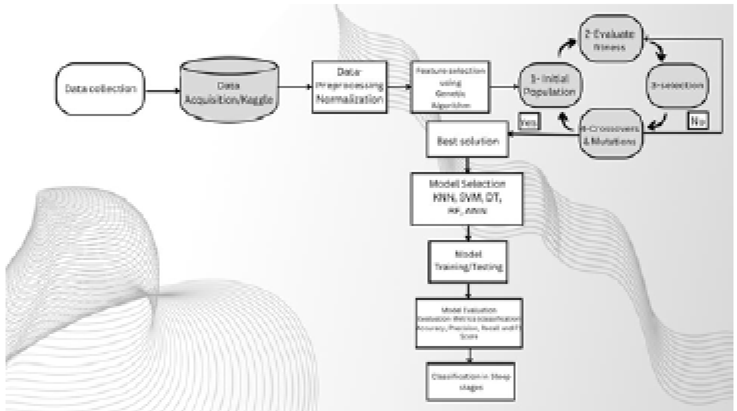


FIGURE 2. **The proposed optimised model for sleep disorder classification.**

in most cases, not everyone is the same. For instance, a large portion of the overall activity area can be occupied by sleep apnea. The majority class may achieve better accuracy using the classification accuracy measure, which is not suitable for this kind of dataset with imbalanced labels [23]. Take the accuracy measure as an example. It works well for balanced label classes but doesn't do much good for imbalanced ones. Classification accuracy, recall, precision, and the F1-score were the four review measures employed in this study [24]. The equations that follow define the mathematical formulations of these statistical indexes. In order to measure the performance of the classification algorithms, we looked at their accuracy, which is defined as the proportion of correct predictions to the total number of predictions, as shown in (1). Here, TP stands for true positive, TN for true negative, FP for false positive, and FN for false negative:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

Precision is the ratio of the number of predicted TPs to the total number of predicted positives (2):

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Recall is the ratio of the number of predicted TPs to the total number of actual TPs (3):

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

The F1-score provides a weighted average for the precision and recall of a number. A perfect F1-score provides low FPs and low FNs (4):

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \quad (4)$$

A method for assigning a score to each input feature used by the model is known as feature significance. The accuracy of the model is greatly affected by the maximum score of the features. Feature significance greatly affects model accuracy in this work, which includes data such as body mass index (BMI), blood pressure, sleep duration, profession, and age (Figure 3).

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

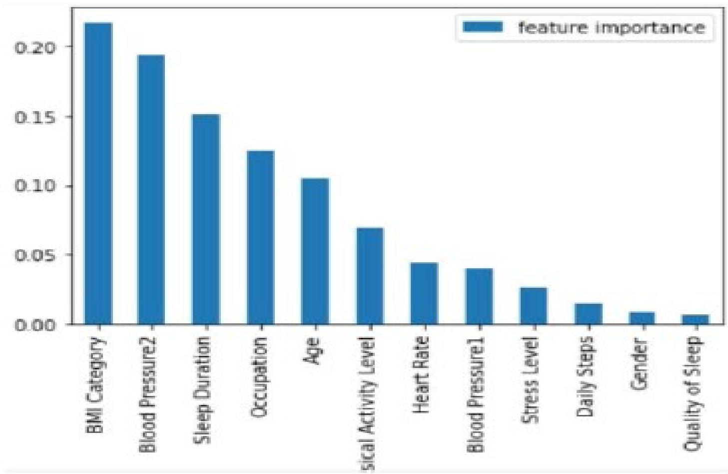


FIGURE 3. Feature importance.

IV. RESULTS AND DISCUSSION

Results from this research support the use of MLAs for the correct classification of sleep disorders. This research was carried out

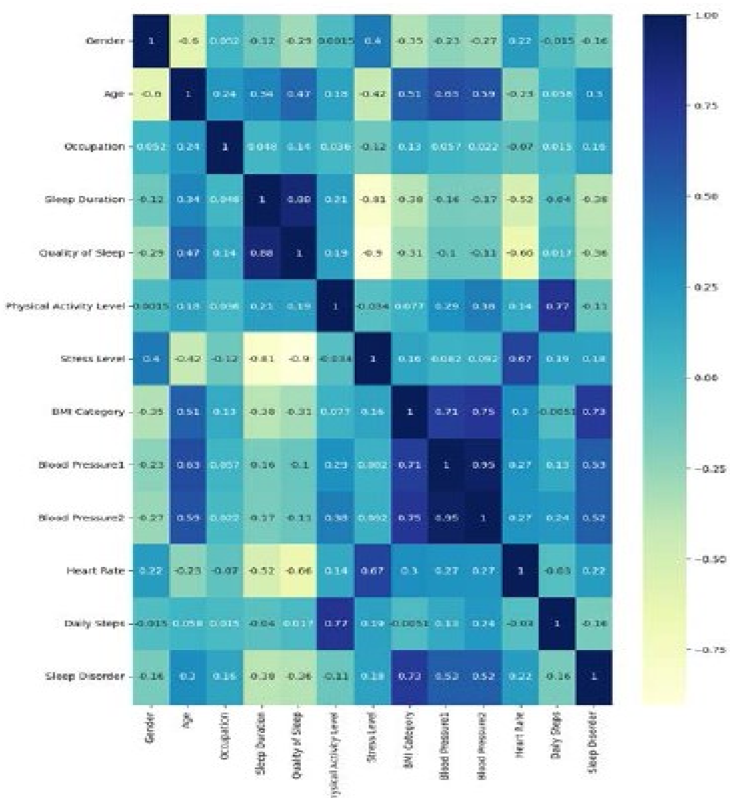


FIGURE 4. Correlation coefficient.

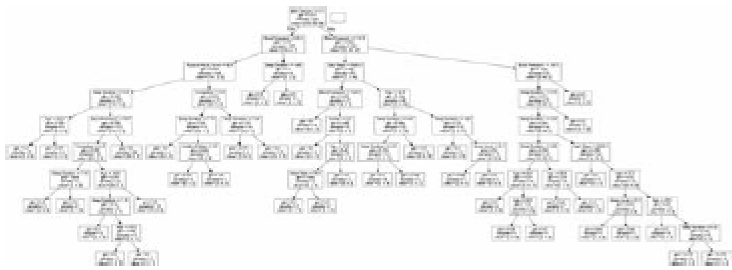


FIGURE 5. Export graphviz flow chart.

to avoid using GA. With accuracies of 84.96%, 64.6%, 86.73%, 88.5%, and 91.15% respectively, the KNN, SVM, DT, and ANN classifiers performed well. An accuracy of more than 90% was attained by the top-performing algorithms. In order to find the best-performing kernel, this article examined training data from many SVMs and used the default setup for each classifier. While linear and polynomial kernels had the lowest accuracy, the RBF kernel achieved excellent performance using the SVM method. Finding the best parameter for each classifier, however, is no easy task. As a result of not having an optimization technique that works well with MLAs on datasets with a lot of dimensions. A training and validation performance plot is shown in Figures 7-8. This data was produced while trying to categorize

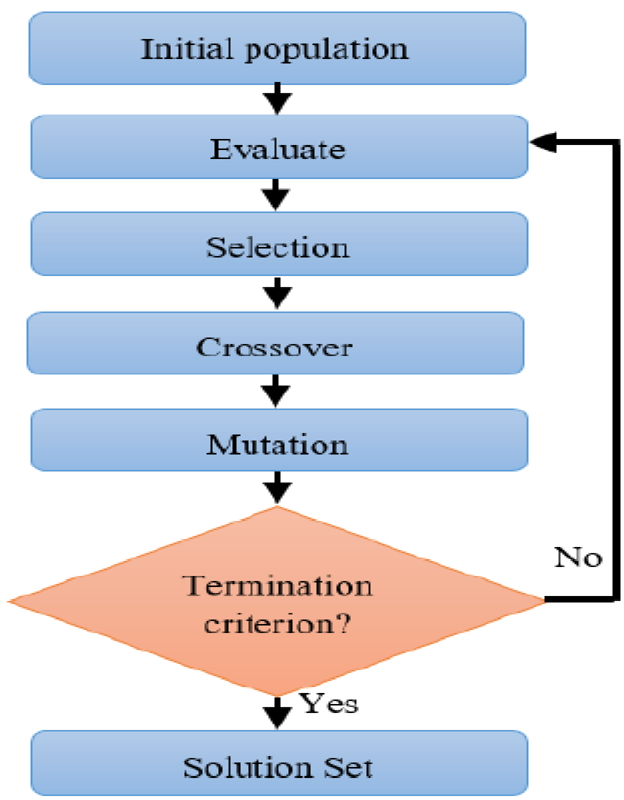
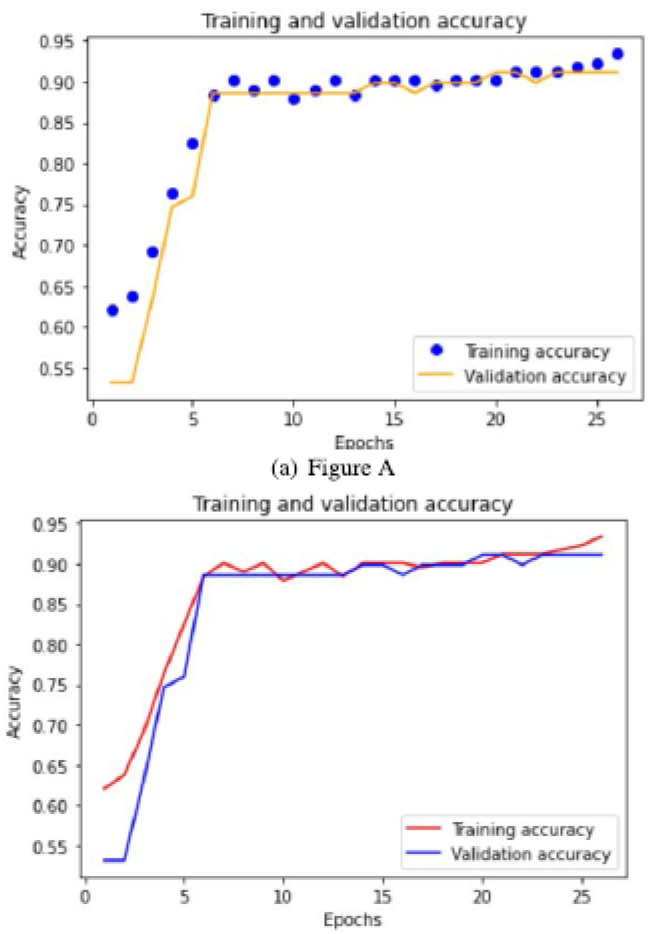


FIGURE 6. Basic architecture of the genetic algorithm [32].

data from the experiments. Points may not be comparable while showing similar loss curves because of changes in model weights. Nonetheless, the training and validation loss gives a clear picture of how the learning performance evolves with time. This strategy is useful for preventing overfitting and finding out whether adding additional training patterns increases the validation score during the learning phase of the model. The results of the training phase for all the MLAs that were assessed are shown in Table 3, and the results obtained

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

utilizing a 5-fold cross-validation are shown in Table 4. Using the F1-score, recall, accuracy, and precision, Figure 9 and Table 5 summarize the overall performance of all the tested MLAs throughout the datasets in the testing phase. We can see how well the tested algorithms competed in the results. Nevertheless, a classification accuracy of 91.15 percent was attained by the deep learning algorithms that relied on neural networks, which outperformed the other traditional machine learning approaches.



(a) Figure A

(b) Figure B

FIGURE 7. **Training and validation accuracy.**

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

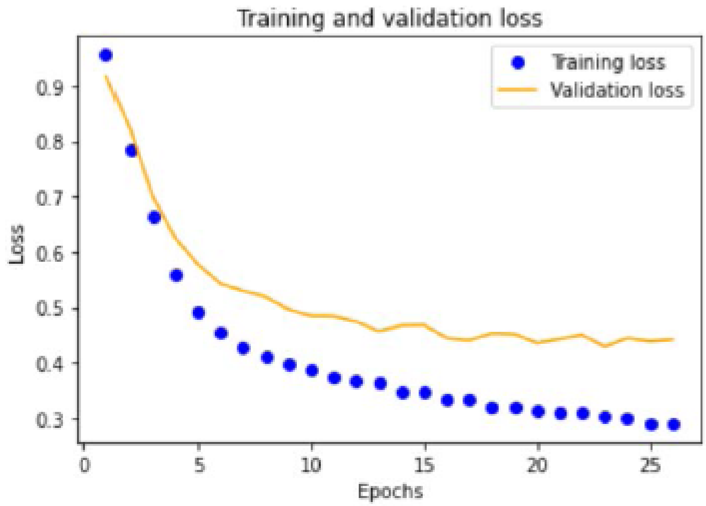


FIGURE 8. Training and validation loss.

In this experiment, we used GA to find the best classifier parameters and search for an optimum one. The findings are in Table 6 and Figure 10. These findings contrast the top GA+MLA model performance with that of the MLA models. Another test is conducted using the t-test to demonstrate that there is a statistically significant difference in the best accuracy of the two models. All machine learning algorithms' test results for the dataset when measured against Precision, Recall,

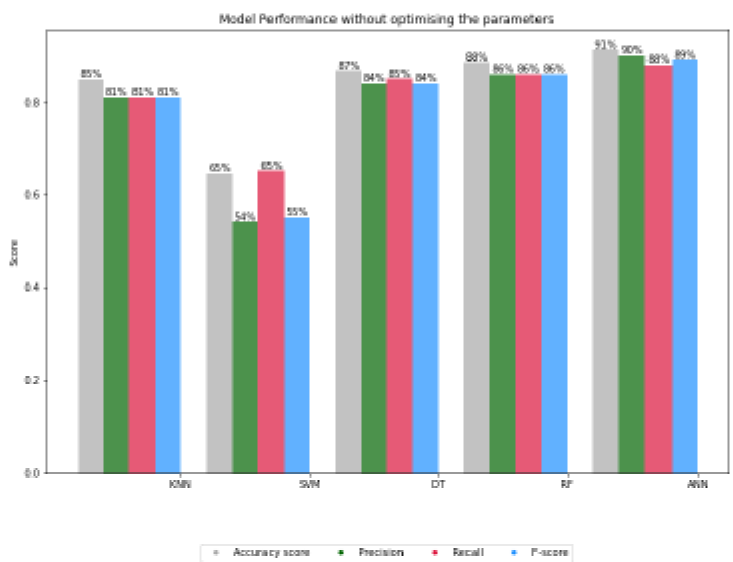


FIGURE 9. Results of the performance of all evaluated MLAs (As default parameters).

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

TABLE 3. Results of the performance of all evaluated MLAs by training phase. (without optimisation of the parameters.).

Evaluation metrics	KNN	SVM	DT	RF	ANN
Precision	87.22%	54.33%	93.49%	93.49%	91.33%
Recall	87.35%	66.28%	93.48%	93.48%	91.18%
F-score	87.25%	57.46%	93.47%	93.48%	91.23%
Accuracy score	87.35%	66.28%	93.48%	93.48%	91.18%

TABLE 4. Results of the performance of all evaluated MLAs by 5-fold cross-validation. (without optimisation of the parameters.).

Evaluation metrics	KNN	SVM	DT	RF	ANN
Precision	87.22%	54%	93.49%	93.49%	92.25%
Recall	87.35%	65%	93.48%	93.48%	91.58%
F-score	87.25%	55%	93.47%	93.48%	91.55%
Accuracy score	83.94%	64.6%	86.99%	88.14%	91.58%

TABLE 5. Results of the performance of all evaluated MLAs by testing phase (without optimisation of the parameters.).

Evaluation metrics	KNN	SVM	DT	RF	ANN
Precision	81%	54%	84%	86%	90%
Recall	81%	65%	85%	86%	88%
F-score	81%	55%	84%	86%	89%
Accuracy score	84.96%	64.6%	86.73%	88.5%	91.15%

ISSN 2454-9940
www.ijasem.org
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
Vol 19, Issue 2, 2025

F-Score and the t-test are shown in **Table 6**. Based onthe test, we can see that while not all algorithms show

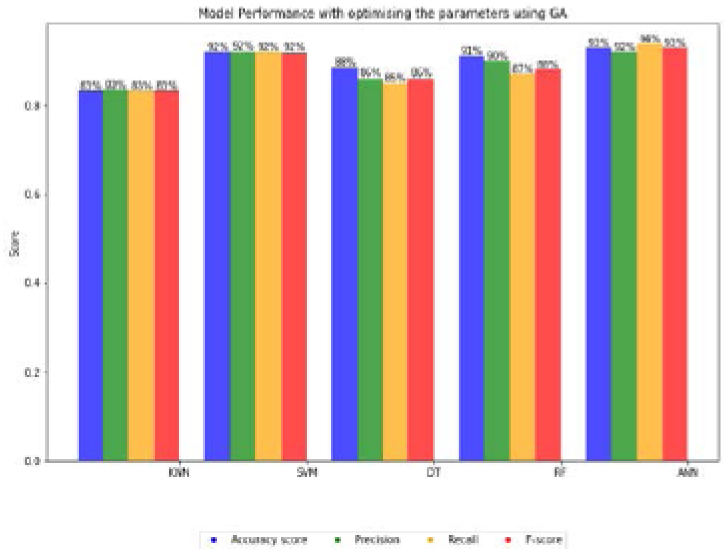


FIGURE 10. Results of the performance of all evaluated MLAs +GA (model performance with optimisation of the parameters using GA).

TABLE 6. Results of the performance of all evaluated MLAs (model performance with optimisation of the parameters using GA).

Evaluation metrics	KNN	SVM	DT	RF	ANN
Precision	83.42%	92.11%	86%	90.00%	92.01%
Recall	83.18%	92.03%	85%	87.00%	93.80%
F-score	83.21%	91.88%	86%	88.00%	91.93%
Accuracy score	83.19%	92.04%	88.50%	91.15%	92.92%

notable variations, the results demonstrated that the suggested approach (GA+MLAs) outperformed MLAs with the default settings. Applying GA to find the best values for the MLAs improved the classifiers' performance. Table 7 shows the results of running the GA with various parameter settings for five generations; Table 8 shows the results of using the fitness score to determine the optimal parameters and solutions. For instance, while optimizing the KNN model, the GA produced the best parameters, which were k = 2 and the metric for Euclidean distance. The full dataset was utilized to train and test the KNN model with these optimized parameters. Classification accuracy was 83.19% for KNN, 92.04% for SVM, 88.50% for DT, 91.15% for RF, and 92.92% for ANN. This paper used a grid search method to optimize the SVM's hyperparameters rather than the GA. With this method, you may speed up your training time, improve your results, and discover the best hyperparameter values for your SVM classifier by searching hyperparameter space. And as both studies made use of the same SleepHealth and Lifestyle Dataset, we were able to compare their findings [33].

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025

TABLE 7. **parameter of the GA settings used.**

Parameter	Value
Population size	12
Generations	5
Elite percentage	0.2
Mutation rate	0.8
Crossover rate	0.8

TABLE 8. **Best-optimised parameters of models.**

Model	Best- Optimised Parameters
KNN	(k = 2, the Euclidean distance metric)
SVM	GridSearchCV(cv=5, estimator=SVC(), param-grid='C': [0.1, 1, 10], 'gamma': [0.001, 0.01, 0.1], scoring='f1-weighted')
DT	(max-depth=4, min-samples-split=3)
RF	(max-depth=9, min-samples-split=6, n-estimators=33)
ANN	('num-hidden-layers': 1, 'num-units-per-layer': 24, 'learning-rate': 0.004068331104981341)

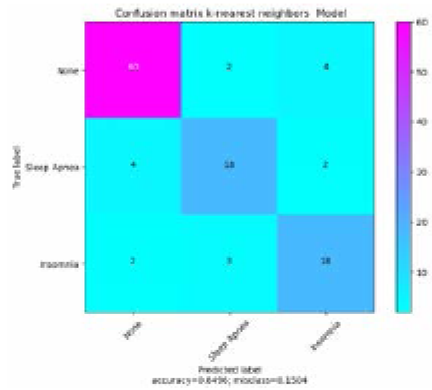
TABLE 9. **The estimating of p values and t-tests.**

Model		t-test result	Conclusion
KNN	t-stat.	0.3375263702777991	No significant
—	p-value	0.7396243325450431	improvement
SVM	t-stat.	-1.5212776585113266	significant
—	p-value	0.14556309281759117	improvement
DT	t-stat.	0.3629888130588261	No significant
—	p-value	0.7208408332545125	improvement
RF	t-stat.	-1.3416407864997872	significant
—	p-value	0.19639447228341037	improvement
ANN	t-stat.	-1.186884852112364	significant
—	p-value	0.2507018182951046	improvement

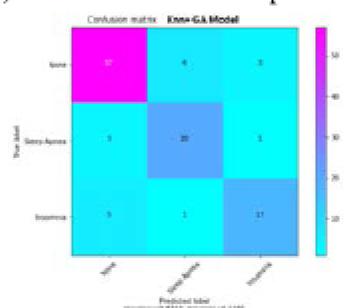
T-TEST ANALYSIS

ISSN 2454-9940
www.ijasem.org
INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
Vol 19, Issue 2, 2025

A t-test for samples was run to see if the improvement achieved by the GA-optimized MLAs was statistically significant. The first premise was that there is a statistically significant difference between the baseline accuracy and the average accuracy of many GA-optimized MLAs models. Table 9 shows the results of the t-test, which showed that the GA-optimized MLAS classifiers significantly improved accuracy.

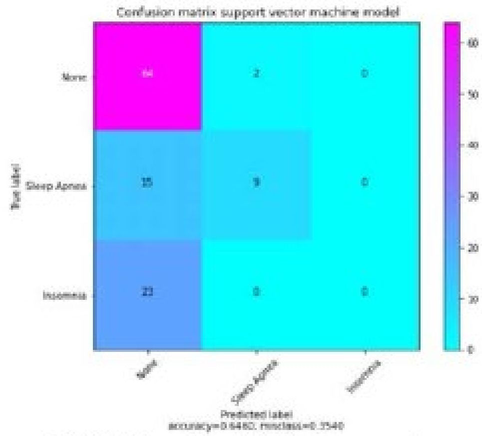


(a) KNN model with default parameters

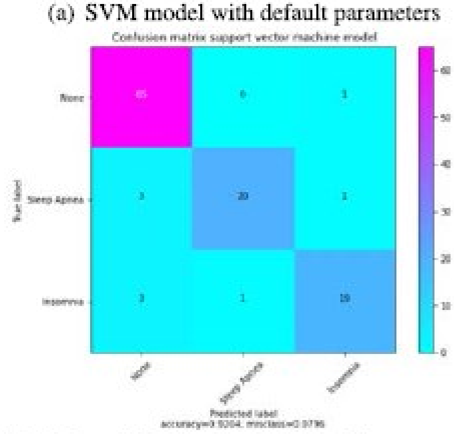


(b) KNN model with optimisation of the parameters

FIGURE 11. Confusion matrix for KNN model.



(a) SVM model with default parameters



(b) SVM model with optimisation of the parameters

FIGURE 12. Confusion matrix for SVM model.

CONFUSION MATRIX

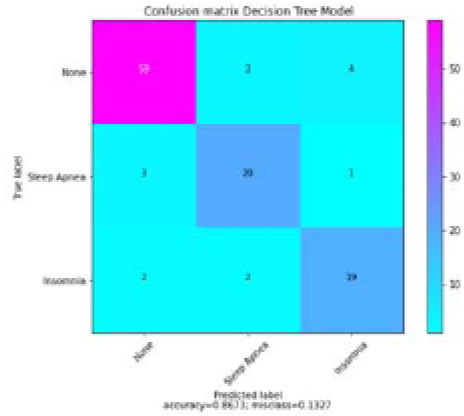
A confusion matrix was used to compare the MLAs classifiers' performance. This matrix summarizes the outcomes of the classification. The confusion matrix for the is shown in Figure 11-15.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

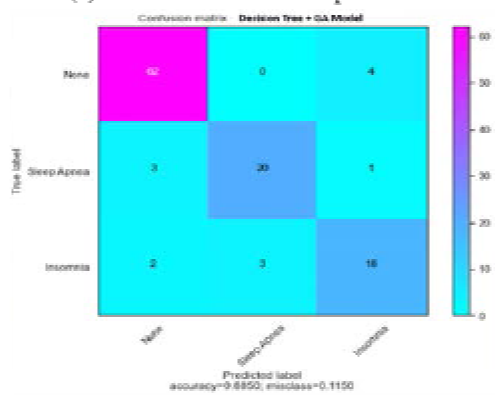
ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025



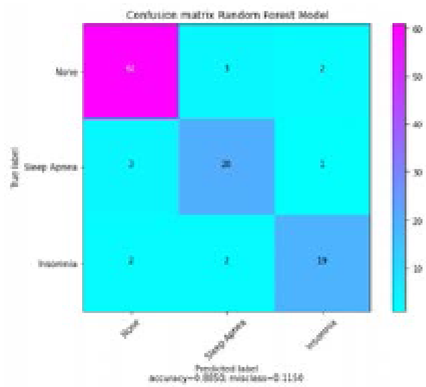
(a) DT model with default parameters



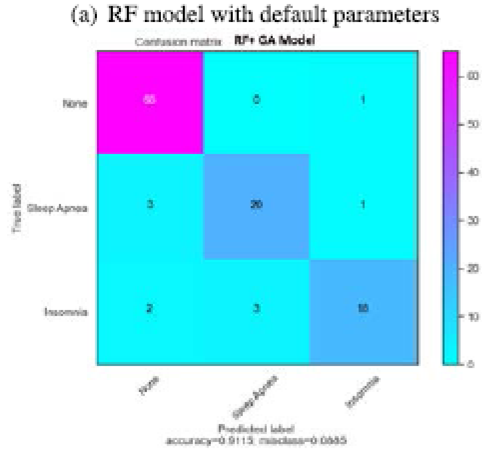
(b) DT model with optimisation of the parameters

FIGURE 13. Confusion matrix for DT model.

ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025



(a) RF model with default parameters



(b) RF model with optimisation of the parameters

FIGURE 14. Confusion matrix for RF model.

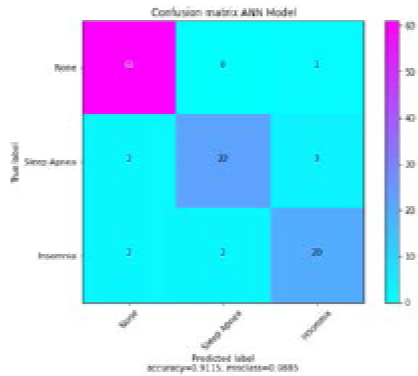
categorization job with several classes. One way to see how well the model worked is by looking at the confusion matrix, which shows how many cases fell into each class and

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

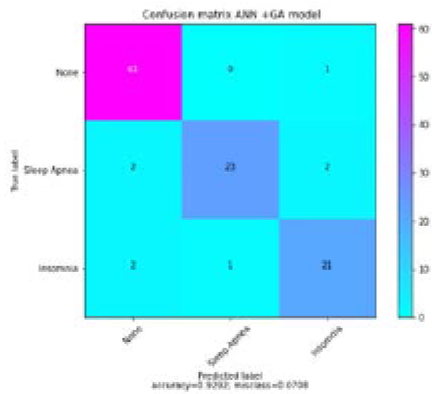
ISSN 2454-9940

www.ijasem.org

Vol 19, Issue 2, 2025



(a) ANN model with default parameters



(b) ANN model with optimisation of the parameters

FIGURE 15. Confusion matrix for ANN model.

mistaken categorizations. Using the ANN+ GA model as an example, 61 out of 100 occurrences were accurately categorized as Class 1, 0 as Class 2, and 1 as Class 3. Also in Class 3, 21 occurrences were accurately identified, but 2 were mislabeled as Class 1 and 1 as Class 2. In Class 2, 23 instances were correctly classified, but 2 were mislabeled as Class 1 and 3. Important information about the RF classifier's performance on the multi-class classification job is revealed by the confusion matrix. Make use of the matrix. With a 96% success rate, the model demonstrated excellent accuracy in Class 1. Class 2 and Class 3 also had lesser accuracy, with 20% and 26% of misclassifications, respectively.

V. CONCLUSION

In order to find the best possible values for the hyperparameters of each model, this study proposes an improved model for sleep disorder classification that uses MLAs in conjunction with a genetic algorithm. Using the real-world Sleep Health and Lifestyle Dataset, this article analyzed the efficacy of several state-of-the-art MLAs for sleep disorder classification. And MLAs don't need characteristics established by experts to learn from high-dimensional sleep data and try to categorize sleep disorders. With an accuracy of 92.92%, the suggested optimized ANN with GA outperformed the competing MLAs. The F1-score, recall, and precision on the test data were 91.93%, 93.80%, and 92.01%, correspondingly. Even when faced with a data constraint. The difficulties of using MLAs for sleep disorder categorization were the focus of this research. Training and assessing models in this discipline, however, still requires enormous datasets. When it comes to sleep disorder categorization, MLAs with GA may make a world of difference. We will evaluate the dataset on a novel model and compare its performance to current state-of-the-art models; moreover, we will create MLAs using unsupervised learning.

REFERENCES

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 19, Issue 2, 2025

[1] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García,''A portable wireless device for cyclic alternating pattern estimationfrom 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 learningfor 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 automaticsleep 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 tointerpret 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, ''Adeep neural network for earlydetection 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 ontraditional machine learning: Simple and interpretable techniques arecompetitive 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 conventionalmachine 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 reinforcementlearning-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, ''Predictionmodels for obstructive sleep apnea inKorean adults using machine learningtechniques,'' *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 fromsingle-channel EEG signals,'' *J. Neurosci. Methods*, vol. 324, Aug. 2019,Art. no. 108312.
[11] S. Djanian, A. Bruun, and T. D. Nielsen, ''Sleep classification usingconsumer 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 machinelearning 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 deeplearning method approach for sleep stage classification with EEGspectrogram,'' *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 techniquesto 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 singleleadECG: 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, ''Performanceanalysis of machine learning algorithms on automated sleep stagingfeature 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-leadECG: A comprehensive analysis of machine learning and deep learningalgorithms,'' *IEEE Trans. Instrum. Meas.*, vol. 71, pp. 1–11, 2022.
[18] J. Ramesh, N. Keeran, A. Sagahyroon, and F. Aloul, ''Towards validatingthe effectiveness of obstructive sleep apnea classification from electronichealth 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 stagesclassification 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 forautomated sleep stages classification using PSG signals,'' *Int. J. Environ.Res. Public Health*, vol. 16, no. 4, p. 599, Feb. 2019.