IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 13, Issue 7, July 2025

DOI: 10.17148/IJIREEICE.2025.13703

SWEET DREAMS POWERED BY MACHINES: MACHINE LEARNING IN SLEEP DISORDER ANALYSIS

SANDI SUNANDA[1], RAYALA SURESH BABU[2], BANDARI CHAITANYA KRISHNA[3]

ASST.PROFESSOR, CSE DEPARTMENT, GNITC, HYDERABAD[1]

ASST.PROFESSOR, EEE DEPARTMENT, GNITC, HYDERABAD[2]

B.TECH IV YEAR, CSE DEPARTMENT, GNITC, HYDERABAD[3]

Abstract: Sleep disorders, such as Insomnia, Sleep Apnea, and other conditions, significantly impact individuals' health and well-being. Accurate and efficient classification of these disorders can aid in early diagnosis and effective treatment, enhancing the quality of life for affected individuals. The existing systems predominantly rely on Artificial Neural Networks (ANN) for classification, which, while effective, can be computationally intensive and less interpretable. This study proposes a Random Forest-based approach for classifying sleep disorders, utilizing a dataset consisting of 400 samples with 13 relevant features. Random Forest model was selected for its robustness, interpretability, and superior ability to handle complex, non-linear relationships within the data. By employing this algorithm, the study aims to classify sleep disorders into three classes: Insomnia, None, and Sleep Apnea, demonstrating improved performance compared to traditional ANN-based systems. The evaluation of the Random Forest model is conducted using standard performance metrics, including accuracy, precision, recall, and F1-score, which show that the proposed approach outperforms existing models, offering enhanced accuracy and reliability in the classification of sleep disorders.

I. INTRODUCTION

Sleep is a vital physiological function necessary for physical and mental health. Sleep helps strengthen the body and consolidate the brain and memories. Sleep quality affects cognitive functions, particularly in children and older drivers at increased risk of accidents. Sleep deprivation can affect the human body and cause health problems like heart disease, diabetes and obesity. Physicians, doctors, medical professionals and experts must manually evaluate polysomnography (PSG) records, which can lead to different assessments of sleep stages. Manual classification is prone to human error and is time-consuming for sleep-stage classification. These techniques can be split into conventional (traditional) machine learning algorithms Random Forest algorithm. Sleep disorders, such as Insomnia, Sleep Apnea, and other related conditions, significantly affect the quality of life, health, and daily functioning of individuals. These disorders can lead to severe health problems, including cardiovascular diseases, diabetes, and mental health issues, highlighting the importance of early and accurate diagnosis. Traditional methods of diagnosing sleep disorders, such as polysomnography, are often time-consuming, expensive, and require specialized medical facilities. As a result, there is an increasing need for automated, efficient, and accurate classification methods that can assist healthcare professionals in identifying and managing sleep disorders. Machine learning algorithms have gained significant attention in the medical field due to their ability to analyze complex patterns in data, providing valuable insights for disease diagnosis and prediction. Existing approaches for classifying sleep disorders often rely on Artificial Neural Networks (ANN), which, despite their effectiveness, have limitations such as high computational costs, a lack of interpretability, and a tendency to overfit, especially when dealing with smaller datasets. The proposed system utilizes a dataset consisting of 400 samples and 13 features, including demographic data, sleep quality indicators, physical activity levels, and vital health statistics. The goal is to classify these samples into three classes: Insomnia, None (no disorder), and Sleep Apnea. The Random Forest algorithm is employed due to its ability to manage various types of data, handle missing values, and generate feature importance scores, which offer valuable insights into the factors contributing to each sleep disorder classification.

II. RELATED WORK

Deep learning model to automatically classify sleep stages using raw PSG signals. The model extracts feature from a one-dimensional CNN. To evaluate the proposed model, they used databases (sleep-edf and sleep-edfx) that are publicly available online. The proposed model obtained high accuracy for two to six sleep classes at 88.22% and 88.00%. The authors suggested that deep learning is a promising approach for automated sleep-stage classification that can replace the job of classical methods to avoid manual experts. This approach is prone to human error.

 This work is licensed under a Creative Commons Attribution 4.0 International License

A summary of the algorithm, dataset and accuracy in some of the reviewed studies is presented in have developed an efficient method that integrated a heterogeneous feature representation and a genetic algorithm-based ensemble learning model to predict ant tubercular peptides to help in the search for a new treatment to strive tuberculosis.

An artificial neural network (ANN) is a supervised learning algorithm that mimics the human brain. It is a combination of interconnected nodes called artificial neurons. The ANN comprises multiple hidden layers that are between the input and output layers. Every entry contains a neural weight. Each input is fed to each neuron of the first layer, and every layer is completely linked to the next layer and is assigned a weight. A weighted sum is sent via a threshold function to an activation function. The output of the activation function determines whether a neuron is activated, and the activated neuron is passed to the output.

DRAWBACKS: -

- This can lead to long training times and high demands on hardware resources like GPUs.
- Ann overfitting, especially when the dataset is small or imbalanced.
- Hyperparameter Tuning Complexity. Requires Extensive Data Preprocessing.

III. LITERATURE SURVEY

Title: Classification of sleep disorders using random forest on sleep health and lifestyle dataset. Author: X. Wu, D. Hong, Z. Huang, and I. A. Hidayat Year: 2023 Description: This study aims to classify sleep disorders using the Random Forest method on the Sleep Health and Lifestyle dataset. This dataset contains information about sleep, lifestyle, and relevant health factors. In this study, the dataset was processed and divided into training and testing subsets. The Random Forest model was trained using the training subset with sleep and health related features. The quality of the split in each decision tree was measured using the Gini Index. The model was evaluated using the testing subset to measure its accuracy and classification performance. The evaluation results showed that the Random Forest model was able to predict sleep disorders with good accuracy. Analysis of class distributions, correlation relationships between features, and visualization by gender provided insights into the factors that influence sleep disorders. This research has the potential to contribute to the field of health and medicine, especially in the recognition and diagnosis of sleep disorders.

Title: Genetic-algorithm-based neural network for fault detection and diagnosis: Application to grid-connected photovoltaic systems. Author: A. Hichri, M. Hajji, M. Mansouri, K. Abodayeh, K. Bouzrara, H. Nounou. Year: 2022. Description: Modern photovoltaic (PV) systems have received significant attention regarding fault detection and diagnosis (FDD) for enhancing their operation by boosting their dependability, availability, and necessary safety. As a result, the problem of FDD in grid connected PV (GCPV) systems is discussed in this work. Tools for feature extraction and selection and fault classification are applied in the developed FDD approach to monitor the GCPV system under various operating conditions. This is addressed such that the genetic algorithm (GA) technique is used for selecting the best features and the artificial neural network (ANN) classifier is applied for fault diagnosis. Only the most important features are selected to be supplied to the ANN classifier. The classification performance is determined via different metrics for various GA-based ANN classifiers using data extracted from the healthy and faulty data of the GCPV system. A thorough analysis of 16 faults applied on the module is performed. In general terms, the faults observed in the system are classified under three categories: simple, multiple, and mixed. The obtained results confirm the feasibility and effectiveness with a low computation time of the proposed approach for fault diagnosis.

Title: Ensemble computational intelligent for insomnia sleep stage detection via the sleep ECG signal. Author: P. Tripathi, M. A. Ansari, T. K. Gandhi, R. Mehrotra, M. B. B. Heyat. Year: 2022. Description: Insomnia is a common sleep disorder in which patients cannot sleep properly. Accurate detection of insomnia disorder is a crucial step for mental disease analysis in the early stages. The disruption in getting quality sleep is one of the big sources of cardiovascular syndromes such as blood pressure and stroke. The traditional insomnia detection methods are time-consuming, cumbersome, and more expensive because they demand a long time from a trained neurophysiologist, and they are prone to human error, hence, the accuracy of diagnosis gets compromised. Therefore, the automatic insomnia diagnosis from the electrocardiogram (ECG) records is vital for timely detection and cure. In this paper, a novel hybrid artificial intelligence (AI) approach is proposed based on the power spectral density (PSD) of the heart rate variability (HRV) to detect insomnia in three classification scenarios: (1) subject-based classification scenario (normal Vs. insomnia), (2) sleep stage-based classification (REM Vs. W. stage), and (3) the combined classification scenario using both subject-based and sleep stage-based deep features. The ensemble learning of random forest (RF) and decision tree (DT) classifiers are used to perform the first and second classification scenarios, while the linear discriminant analysis (LDA) classifier is used to perform the third combined scenario. The proposed framework includes data collection, investigation of the ECG signals, extraction of the signal HRV, estimation of the PSD, and AI-based classification via hybrid machine learning classifiers.

DOI: 10.17148/IJIREEICE.2025.13703

Title: Performance analysis of machine learning algorithms on automated sleep staging feature sets. Author: S. Satapathy, D. Loganathan, H. K. Kondaveeti, and R. Rath, Year: 2021 Description: Abstract With the speeding up of social activities, rapid changes in lifestyles, and an increase in the pressure in professional fields, people are suffering from several types of sleep-related disorders. It is a very tedious task for clinicians to monitor the entire sleep durations of the subjects and analyze the sleep staging in traditional and manual laboratory environmental methods. For the purpose of accurate diagnosis of different sleep disorders, we have considered the automated analysis of sleep epochs, which were collected from the subjects during sleep time. The complete process of an automated approach of sleep stages' classification is majorly executed through four steps: pre-processing the raw signals, feature extraction, feature selection, and classification. In this study, we have extracted 12 statistical properties from input signals. The proposed models are tested in three different combinations of features sets. In the first experiment, the feature set contained all the 12 features. The second and third experiments were conducted with the nine and five best features. The patient records come from the ISRUC-Sleep database. The highest classification accuracy was achieved for sleep staging through combinations with the five-feature set. From the categories of the subjects, the reported accuracy results were found to exceed above 90%. As per the outcome from the proposed system the random forest classification techniques achieved best accuracy incomparable to that of the other two classifiers.

Title: Adversarial learning for semi-supervised pediatric sleep staging with single-EEG channel. Author: Y. Li, C. Peng, Y. Zhang, Y. Zhang, and B. Lo. Year: 2022. Description: Despite the progress recently made towards automatic sleep staging for adults, children have complicated sleep structures that require attention to the pediatric sleep staging. Semi-supervised learning, i.e., training networks with both labeled and unlabeled data, greatly reduces the burden of epoch-by-epoch annotation for physicians. However, the inherent class imbalance problem in sleep staging task undermines the effectiveness of semi-supervised methods such as pseudo-labeling. In this paper, we propose a Bi-Stream Adversarial Learning network (BiSALnet) to generate pseudo-labels with higher confidence for network optimization. Adversarial learning strategy is adopted in Student and Teacher branches of the two-stream networks. The similarity measurement function minimizes the divergence between the outputs of the Student and Teacher branches, and the discriminator continuously enhances its discriminative ability. In addition, we employ a powerful symmetric positive definite (SPD) manifold structure in the student branch to capture the desired feature distribution properties. The joint discriminative power of convolutional features and nonlinear complex information aggregated by SPD matrices is combined by the attention feature fusion module to improve the sleep stage classification performance.

IV. PROPOSED WORK

In this study, the authors review research in the field of sleep disorders, focusing on such challenges as data collection, which includes data that are often noisy and uncertain (e.g. missing data) from various hospitals from patients during sleep. The dataset has many limitations due to the data being collected from only one sleep clinic. It is challenging to generalize evaluated results due to the bias of the data towards certain groups of patients, and the biased data can lead to inaccurate results that can influence decision making. However, there is a lack of natural sleep-stage datasets. Moreover, feature extraction from the dataset is required to train models and select discriminative features, which usually requires more computational effort to select well-suited MLAs from different classifiers. This study is motivated toward the requirement to handle the challenges caused through sleep disorders in the modern lifestyle, especially for people suffering from sleep disorders.

An RF classifier is an ensemble learning algorithm that creates multiple random DTs to combine the predictions, improve model predictive accuracy and manage overfitting. The model can use two random processes: Bootstrapping and random selection of features.

Bootstrapping guarantees that the model does not utilize the like data for each tree, so that the model is minimally sensitive to conversions in the training data data. Random feature selection reduces the correlation between the trees and aggregates. The proposed method contributes to the field of medical diagnostics by providing a reliable and interpretable tool that enhances the early detection and management of sleep disorders.

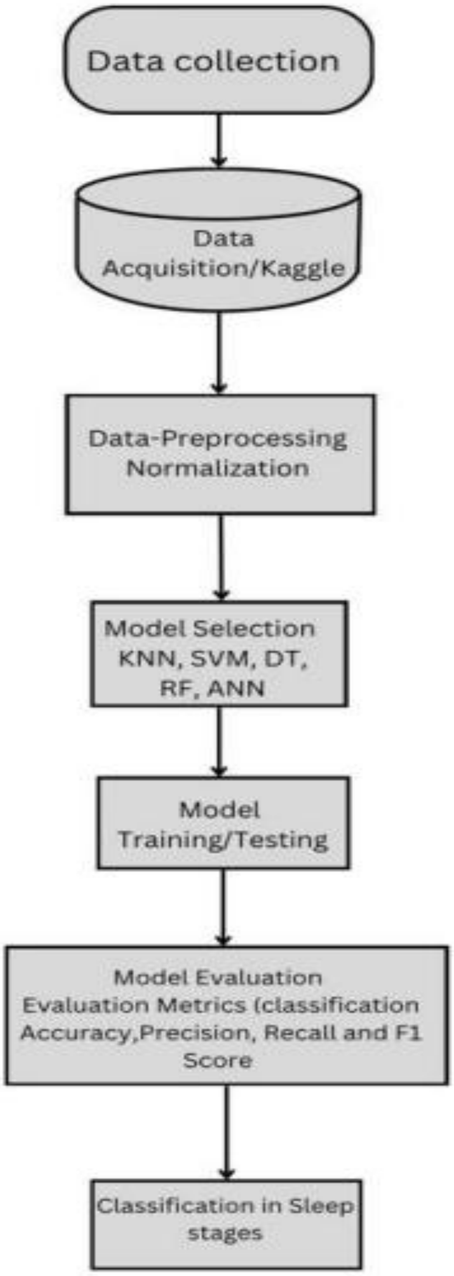
ADVANTAGES: -

- it reduces the variance of the model, leading to improved predictive performance.
- Interpretability and Feature Importance.
- RF inherently reduces overfitting through the use of multiple trees and random feature selection.
- Handles Missing Data Effectively.Flexibility in Handling Large Datasets.

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IJIREEICE.2025.13703

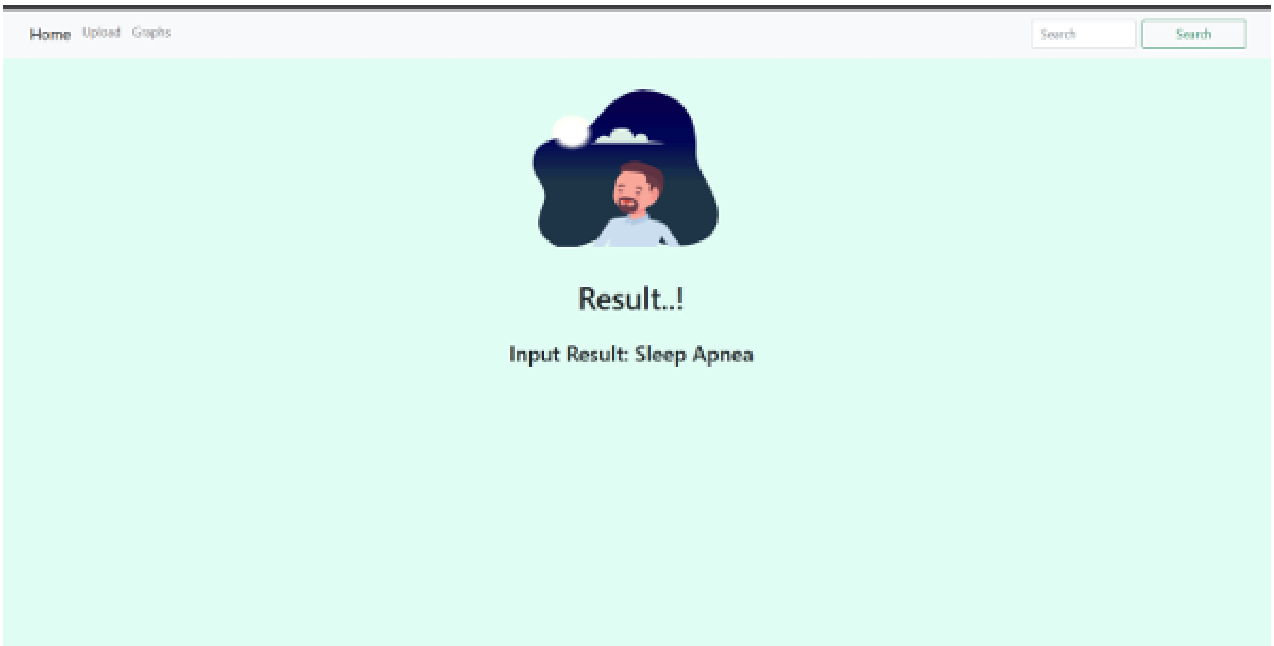
System Architecture:

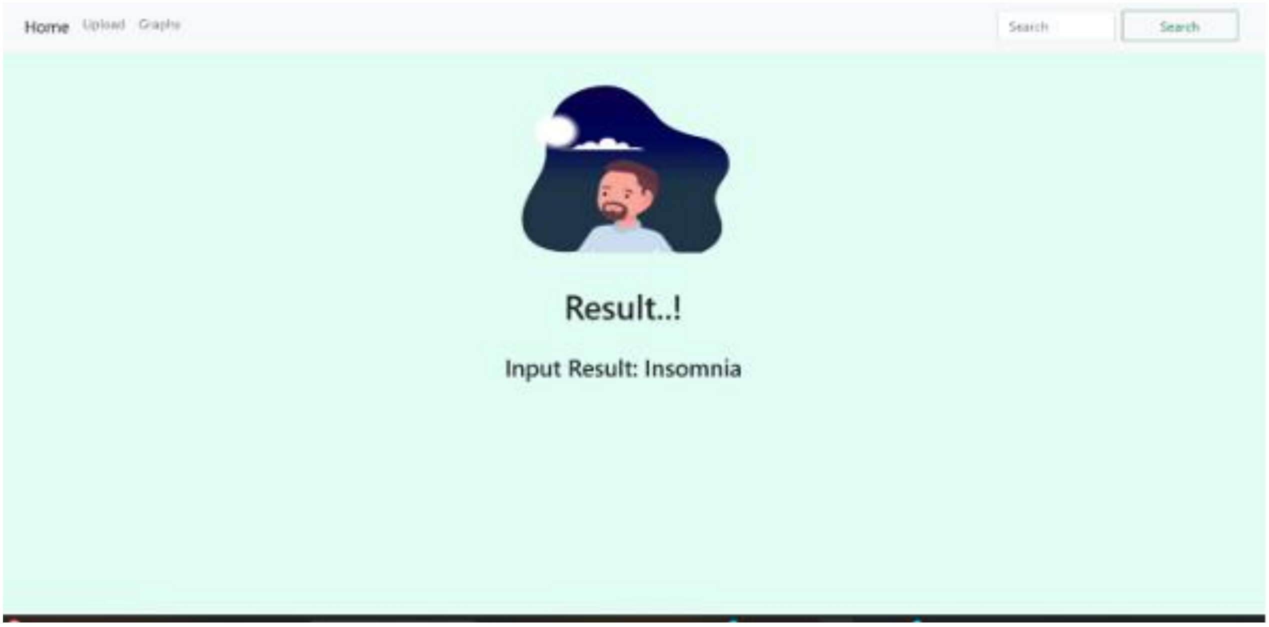


ADVANTAGES:

➢ It reduces the variance of the model, leading to improved predictive performance.
➢ Interpretability and Feature Importance.
➢ RF inherently reduces overfitting through the use of multiple trees and random feature selection.
➢ Handles Missing Data Effectively.

V. RESULTS





VI. CONCLUSION

The optimized model for sleep disorder classification was proposed using machine learning algorithms. The study originally implemented Random Forest Algorithm and demonstrating that MLAs can effectively classify sleep disorders by learning from high-dimensional data without relying on expert-defined features. Among these models, the optimized ANN with GA achieved less accuracy and satisfactory precision, recall, and F1-score values. To implemented the Random Forest algorithm, which outperformed the existing models by achieving an accuracy of 95%. The Random Forest model demonstrated superior performance due to its ability to handle complex data structures, reduce overfitting, and provide interpretability, which makes it highly suitable for real-world applications in sleep disorder classification. Despite the limitations of a relatively small dataset, the Random Forest model has proven to be a robust alternative, showcasing its effectiveness in accurately classifying sleep disorders.

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering
Impact Factor 8.414 ∺ Peer-reviewed & Refereed journal ∺ Vol. 13, Issue 7, July 2025
DOI: 10.17148/IJIREEICE.2025.13703

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. Department of Computer Science and Engineering, GNITC 54 Sweet Dreams Powered by Machines: Machine Learning in Sleep Disorder Analysis
[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 singlelead 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. 36120 VOLUME 12, 2024 T. S. Alshammari: Applying MLAs for the Classification of Sleep Disorders
[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. Department of Computer Science and Engineering, GNITC 55 Sweet Dreams Powered by Machines: Machine Learning in Sleep Disorder Analysis
[22] (2023). Sleep Health and Lifestyle Dataset. [Online]. http://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyledataset Available:
[23] F. Ordóñez and D. Roggen, ‘‘Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition,’’ Sensors, vol. 16, no. 1, p. 115, Jan. 2016.

 This work is licensed under a Creative Commons Attribution 4.0 International License

IJIREEICE ISSN (O) 2321-2004, ISSN (P) 2321-5526

International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering

Impact Factor 8.414 ⩳ Peer-reviewed & Refereed journal ⩳ Vol. 13, Issue 7, July 2025

DOI: 10.17148/IJIREEICE.2025.13703

[24] D. M. W. Powers, ''Evaluation: From precision, recall and Fmeasure to ROC, informedness, markedness and correlation,'' 2020, arXiv:2010.16061.

[25] F. Pedregosa, ''Scikit-learn: Machine learning in Python,'' J. Mach. Learn. Res., vol. 12, pp. 2825–2830, Nov. 2011.

[26] M. Bansal, A. Goyal, and A. Choudhary, ''A comparative analysis of Knearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning,'' Decis. Anal. J., vol. 3, Jun. 2022, Art. no. 100071.

[27] M. Q. Hatem, ''Skin lesion classification system using a K-nearest neighbor algorithm,'' Vis. Comput. Ind., Biomed., Art, vol. 5, no. 1, pp. 1–10, Dec. 2022.

[28] V. G. Costa and C. E. Pedreira, ''Recent advances in decision trees: An updated survey,'' Artif. Intell. Rev., vol. 56, no. 5, pp. 4765–4800, May 2023.

[29] 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.

[30] 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.

[31] S. Kuanar, V. Athitsos, N. Pradhan, A. Mishra, and K. R. Rao, ''Cognitive analysis of working memory load from eeg, by a deep recurrent neural network,'' in Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP), Apr. 2018, pp. 2576–2580.

[32] A. Hichri, M. Hajji, M. Mansouri, K. Abodayeh, K. Bouzrara, H. Nounou, and M. Nounou, ''Genetic-algorithm-based neural network for fault detection and diagnosis: Application to grid-connected photovoltaic systems,'' Sustainability, vol. 14, no. 17, p. 10518, Aug. 2022.

[33] 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. Department of Computer Science and Engineering, GNITC 5