Sleep Disorder Identification Using Machine Learning and Deep Learning Models

C Rekha[1], Marthala Amruthavalli[2], Basineni Bharath Kumar[3], Gopidinne Harshitha[4], Voodi Hemanth[5]
[1]*Assistant Professor, Department of Artificial Intelligence and Machine Learning, Annamacharya Institute of Technology & Sciences, Tirupati, India*
[2,3,4,5]*Student, Department of Artificial Intelligence and Machine Learning, Annamacharya Institute of Technology & Sciences, Tirupati, India*

***Abstract*—sleep disorders, including sleep apnea, significantly impact human health and overall quality of life. Accurate classification of sleep disorders is essential for effective diagnosis and treatment; however, manual sleepstage classification by experts is often time consuming and prone to error. This study explores the application of machine learning algorithms (MLAs) for sleep disorder classification using the publicly available Sleep Health and Lifestyle Dataset, which comprises 400 records and 13 features related to sleep and daily activities. Conventional MLAs such as k-nearest neighbours, support vector machines, decision trees, and random forests were compared with a deep learning approach based on artificial neural networks (ANN). To enhance model performance, a genetic algorithm was employed to optimise the parameters of each model. Experimental results revealed notable performance variations among the evaluated algorithms. The ANN model demonstrated superior classification capability compared to conventional machine learning methods, achieving strong precision, recall, and F1-score. These findings highlight the potential of optimised deep learning models, particularly ANN, in enabling reliable and efficient sleep disorder classification, thereby contributing to improved healthcare outcomes.**

***Index Terms*—Machine learning algorithms, deep learning, classification, genetic algorithm**

I. INTRODUCTION

1.1 Background And Motivation

Sleep disorder identification has emerged as a vital area of research in the field of intelligent healthcare systems, driven by the increasing need for accurate, non-invasive, and automated diagnosis of sleep-related conditions. Traditional diagnostic techniques such as polysomnography limit accessibility due to their high cost, complexity, and requirement for clinical supervision, especially for continuous monitoring scenarios. The evolution of machine learning and deep learning techniques, particularly Artificial Neural Networks (ANNs), has significantly enhanced the ability to analyze physiological and behavioral data with high accuracy, making them highly suitable for sleep disorder detection tasks. This project is motivated by the potential of intelligent sleep monitoring systems to bridge the gap between patients and healthcare providers, enabling early diagnosis and continuous health assessment in real-world environments. The challenges in this domain such as variability in physiological signals, noise in sensor data, imbalanced datasets, and differences in individual sleep patterns necessitate robust and optimized computational solutions. Hence, this research focuses on designing a hybrid learning framework that integrates ANN with Genetic Algorithms (GA), where GA is used to optimize network parameters such as weights, biases, and architecture to improve performance and generalization. Additionally, preprocessing techniques such as normalization, feature extraction, noise filtering, and data balancing are employed to enhance model robustness. Sleep disorder identification systems have promising applications in wearable health devices, telemedicine platforms, and remote patient monitoring, supporting timely clinical intervention and improved patient care. The motivation further extends to the growing importance of AI-driven healthcare technologies, where automated sleep analysis offers a more efficient and scalable approach to diagnosis and treatment planning. By addressing the limitations of existing diagnostic methods and improving detection accuracy, this project aims to deliver a reliable and practical solution for real-world healthcare applications, ultimately enhancing accessibility, patient wellbeing, and quality of life.

1.2 Objectives

This study aims to achieve the following objectives:

1.2.1 Develop a Machine Learning and Deep Learning-Based System for Accurate Sleep Disorder Identification

The primary objective of this project is to design and implement a highly accurate sleep disorder identification system using advanced machine learning and deep learning techniques, particularly Artificial Neural Networks (ANNs) optimized with Genetic Algorithms (GA). The system will be trained on physiological and behavioral datasets such as heart rate, EEG signals, respiration patterns, and sleep cycles to ensure robustness and high accuracy in identifying disorders like insomnia, sleep apnea, and narcolepsy, overcoming the limitations of traditional diagnostic techniques.

1.2.2 Enable Real-Time Monitoring and Early Detection of Sleep Disorders

Beyond classification, the system aims to support near real-time monitoring and analysis of sleep data collected from wearable devices and remote sensors. This objective enables early detection of abnormalities in sleep patterns, allowing timely medical intervention and continuous health tracking. The system can be integrated into telemedicine platforms and smart healthcare environments to provide alerts, diagnostic insights, and personalized sleep health recommendations.

1.2.3 Make the System Accessible for Healthcare, Home Monitoring, and Assistive Technologies

A significant focus of the project is to ensure the sleep disorder identification system is accessible and usable in home-based monitoring environments, wearable health devices, and clinical decision support systems. By ensuring compatibility with low-cost sensors and user-friendly interfaces, the system can be adopted by individuals, clinicians, and caregivers, promoting inclusive, preventive, and patient-centric healthcare solutions.

1.3 Scope

This research focuses on the following major areas:

1.3.1 Focus on Sleep Disorder Identification Using ANN Optimized with Genetic Algorithm

The study emphasizes the use of Artificial Neural Networks enhanced with Genetic Algorithm optimization to accurately classify different sleep disorders. It considers variations in physiological signals across age groups, lifestyles, and health conditions, ensuring adaptability and reliability in diverse real-world conditions.

1.3.2 Development of an Intelligent Sleep Monitoring and Alert System

The research includes the implementation of a system capable of converting analyzed sleep data into meaningful health insights and alerts. These outputs are intended to support early diagnosis, clinical decision-making, and personalized treatment planning. Real-time or near real-time performance is emphasized to ensure prompt detection of sleep abnormalities.

1.3.3 Design with Accessibility and User-Centric Healthcare Functionality

The system is designed for use by a wide range of users including patients, healthcare professionals, and researchers. A simple and intuitive interface ensures ease of use for individuals with minimal technical knowledge, making the system suitable for both clinical and personal health monitoring applications.

1.3.4 Potential for Future Integration and Expansion

While the current scope focuses on identifying common sleep disorders, the system architecture supports future scalability. It can be extended to include additional physiological parameters, integration with IoT-based wearable devices, cloud-based health monitoring systems, and AI-driven personalized treatment recommendations, paving the way for advanced intelligent sleep healthcare solutions.

II. LITERATURE SURVEY

2.1 Traditional Methods of Sleep Disorder Identification

Historically, sleep disorder identification has been carried out using clinical observation, questionnairebased assessments, and polysomnography (PSG), along with conventional signal processing techniques. These approaches relied on manual feature extraction from physiological signals such as EEG, ECG, respiratory patterns, and oxygen saturation levels. While these traditional methods laid the foundation for sleep diagnostics, they also presented several limitations:

2.1.1 Reliance on Specialized Clinical Equipment

Many traditional sleep disorder diagnostic systems require specialized equipment such as polysomnography devices,
EEG monitors, and hospital-based sleep labs to collect and analyze physiological data. These systems are expensive, require trained medical personnel, and are not suitable for continuous or home-based monitoring, limiting their accessibility for large populations.

2.1.2 Sensitivity to Environmental and Patient Conditions

Conventional diagnostic approaches are often affected by environmental conditions such as clinical settings, sensor
placement, and patient discomfort during monitoring. Variations in patient physiology, movement during sleep, and external disturbances can lead to inconsistent data collection and reduced diagnostic accuracy.

2.1.3 Manual Feature Extraction and Expert Dependency

Earlier sleep disorder detection systems relied heavily on handcrafted feature extraction techniques such as frequency band analysis, waveform characteristics, and threshold-based classification. This process required significant domain expertise from clinicians and signal processing specialists, making it time-consuming and prone to human error when applied across diverse patient populations.

2.1.4 Limited Scalability and Adaptability

Most traditional methods are not easily scalable for large datasets or continuous monitoring applications. Adding new parameters or adapting the system for different types
of sleep disorders often requires redesigning the diagnostic framework. These systems also struggle to handle complex and nonlinear relationships present in physiological data, limiting their effectiveness in real-world, large-scale healthcare applications.

2.2 Advances In Machine Learning and Deep Learning for Sleep Disorder Identification

Recent advancements in machine learning and deep learning techniques have significantly improved the accuracy and efficiency of sleep disorder identification systems. Unlike traditional diagnostic approaches, modern models can automatically learn complex patterns and relationships from physiological and behavioural sleep data such as EEG, ECG, respiration signals, and heart rate variability.

Emergence of ANN and Hybrid Optimization Techniques: ANN-based models have shown superior performance in sleep disorder classification tasks. Optimization using Genetic Algorithms (GA) helps in automatically tuning network parameters, reducing training time and improving performance on smaller and complex biomedical datasets.

Optimized Sleep Pattern Analysis Using ANN and Genetic Algorithm: Artificial Neural Networks (ANNs) are used to model complex, nonlinear relationships in physiological sleep data. Genetic Algorithms (GA) optimize ANN parameters such as weights and architecture for improved performance. This hybrid approach enhances accuracy and generalization in detecting sleep disorders like insomnia and sleep apnea.

Comparative Performance: Compared to traditional models like SVM, KNN, and Decision Trees, ANN combined with Genetic Algorithm provides higher accuracy and adaptability. The hybrid ANN-GA model handles noisy signals and patient variability more effectively. This makes it more robust and suitable for real-world sleep disorder identification applications.

2.3 Applications And Challenges In Sleep Disorder Identification

Applications Across Healthcare Domains: Sleep disorder identification systems are used in clinical diagnosis support, remote monitoring, wearable health devices, and telemedicine platforms. They enable continuous tracking, early detection, and personalized healthcare for improved patient management.

Challenges in Implementation: Despite advancements, several challenges persist in sleep disorder detection systems. These include inter-patient variability in physiological signals, noise in sensor data, missing or imbalanced datasets, and variations in sleep environments. Additionally, ensuring real-time analysis and maintaining patient privacy and data security remain critical concerns in practical deployments.

Need for Robust and Scalable Frameworks: Modern systems use preprocessing techniques like normalization, noise filtering, feature extraction, and data balancing to improve reliability. Integrating ANN with Genetic Algorithm optimization enhances accuracy and scalability for real-world healthcare applications.

III. METHODOLOGY

3.1 Dataset Preparation

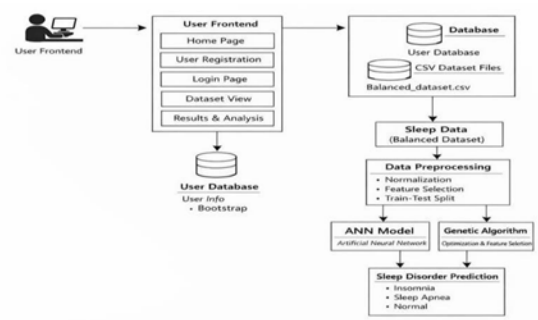
The dataset plays a crucial role in building an accurate and reliable sleep disorder identification system. It consists of structured physiological and behavioral data representing sleep conditions such as normal sleep, insomnia, and sleep apnea, including signals like heart rate, EEG, respiration, oxygen levels, and sleep duration collected under varying conditions to improve generalization. To enhance diversity and prevent overfitting, preprocessing techniques such as normalization, noise filtering, feature extraction, and data balancing are applied, ensuring accurate performance across real-world scenarios.

3.2 System Architecture

The Sleep Disorder Classification System using Machine Learning and Deep Learning models is designed with multiple structured layers including user interface, database management, data preprocessing, model training, optimization, prediction, and result generation. The system starts with a user-friendly frontend interface that includes the Home Page, User Registration, Login Page, Dataset View, and the Results & Analysis section.Users can securely log in to access the balanced sleep dataset, while user details and CSV dataset files are stored in the database. The system uses a balanced sleep dataset containing various health and lifestyle parameters such as sleep duration, stress level, BMI, heart rate, and other medical indicators.

After retrieving the dataset, the system moves to the data preprocessing stage, where the data is cleaned, normalized, and important features are selected. The dataset is then divided into training and testing sets, typically in an 80:20 ratio. The processed data is passed to the Artificial Neural Network (ANN) model, which acts as the deep learning classifier. The ANN consists of an input layer, hidden layers with activation functions, and an output layer for multi-class classification. To further improve model performance, a Genetic Algorithm is applied for optimization and feature selection, enhancing accuracy and reducing overfitting.

Once the model is trained and optimized, the system performs sleep disorder prediction. When a user inputs sleep-related parameters through the interface, the data undergoes preprocessing and is fed into the trained ANN model. The system then classifies the condition into categories such as Insomnia, Sleep Apnea, or Normal sleep. Finally, the predicted result along with performance analysis is displayed on the user dashboard, providing an efficient and intelligent solution for sleep disorder classification.



3.3 Deep Learning Model

The Artificial Neural Network (ANN) serves as the core deep learning component of the sleep disorder identification system.

3.3.1 Model Architecture

An Artificial Neural Network (ANN) was implemented with multiple dense layers optimized for structured healthcare data classification. The input layer consists of neurons corresponding to the number of selected features in the dataset. Hidden layers use the ReLU (Rectified Linear Unit) activation function to capture complex and non-linear relationships between physiological parameters and sleep disorders. Dropout layers are included to prevent overfitting and improve generalization capability. The final output layer uses a Softmax activation function to perform multi-class classification, producing probability scores for the classes: Insomnia, Sleep Apnea, and Normal Sleep. The model effectively learns patterns from structured health data without requiring manual feature engineering.

3.3.2 Compilation and Training

The model was built with the categorical cross-entropy loss function and trained using the Adam optimizer with an appropriate learning rate. Accuracy was used as the primary evaluation metric. The training process was carried out for multiple epochs with a defined batch size to ensure stable convergence. The integration of Genetic Algorithm further improved model performance by optimizing feature subsets and hyperparameters.

3.4 Training And Validation

The dataset was divided into 80% training and 20% testing sets to evaluate the performance of the proposed system. During training, the ANN model learned complex associations between sleep-related features and disorder categories. The validation set was used to monitor model accuracy and loss to prevent overfitting. Performance evaluation was conducted using metrics such as accuracy, precision, recall, F1-score, and confusion matrix. Comparative analysis was also performed between ANN and traditional machine learning models (SVM, Random Forest, KNN, Decision Tree). The optimized ANN-GA model demonstrated superior classification accuracy and better generalization performance compared to conventional approaches.

3.5 User Interface

The user interface is designed to be simple, intuitive, and suitable for both healthcare professionals and patients. The system allows users to input sleep-related parameters such as sleep duration, stress level, heart rate, and physical activity level through a structured form interface. Once the input data is submitted, the system preprocesses the information and feeds it into the trained ANN-GA model for prediction. The predicted sleep disorder category is displayed clearly on the screen along with relevant insights. The interface includes validation mechanisms to handle incorrect or incomplete data entries. The system is responsive and can be deployed as a web-based or desktop-based healthcare support tool, making it suitable for real-world applications.

IV. IMPLEMENTATION

4.1 Tools and Technologies

To develop the Sleep Disorder Identification system using ML and DL models, a combination of efficient tools and technologies was utilized. Python was used as the primary programming language due to its simplicity and strong support for data science libraries. Data preprocessing and manipulation were performed using Pandas and NumPy. Machine Learning models such as SVM, KNN, Decision Tree, and Random Forest were implemented using the Scikit-learn library. The Artificial Neural Network (ANN) was developed using TensorFlow and Keras frameworks to enable efficient deep learning model construction and training. Matplotlib and Seaborn were used for visualization of performance metrics, including confusion matrix and accuracy graphs. The Genetic Algorithm was implemented to optimize feature selection and model hyperparameters, improving classification accuracy and reducing overfitting. The system can be integrated into a web application using Flask or Django for deployment in healthcare environments.

4.2 Code Overview

The implementation of the Sleep Disorder Identification system is divided into three main components:

4.2.1 Loading Data and Preprocessing

The dataset is loaded using Pandas and analyzed for missing values and inconsistencies. The categorical data is converted into a machine-readable format, and the numerical features are scaled to a consistent range. After that, the dataset is divided into two parts: one used for training the model and the other reserved for testing its performance.

4.2.2 Constructing and training the CNN Model

Machine Learning models and the ANN model are constructed and trained using the processed dataset. The Genetic Algorithm is applied to optimize feature subsets and hyperparameters. The best-performing model is saved for future predictions.

4.2.3 Prediction and Classification

New patient data is input into the system, pre-processed, and passed to the trained model. The system outputs the

Predicted sleep disorder category along with probability scores and evaluation metrics.

V. RESULT AND DISCUSSION

5.1 Model Performance

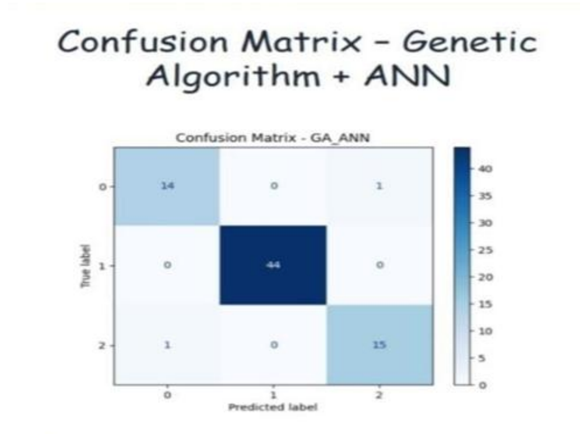
During the testing phase, the proposed Sleep Disorder Classification using ML and DL Models demonstrated strong performance. Initially, the Artificial Neural Network (ANN) model achieved an accuracy of 93.33% on the test dataset. The classification report showed good precision, recall, and F1-score values across all three classes (Insomnia, Sleep Apnea, and Normal), indicating that the ANN was able to effectively learn patterns from the sleep health parameters. To further enhance performance, a Genetic Algorithm (GA) was integrated with the ANN model for feature selection and parameter optimization. After applying the Genetic Algorithm, the optimized ANN model achieved a significantly improved accuracy of 97.33%. The precision, recall, and F1-score values also increased, showing better class-wise prediction and reduced misclassification. The improvement in performance demonstrates that the Genetic Algorithm successfully selected the most relevant features and optimized the ANN parameters, thereby enhancing generalization capability and reducing overfitting. The macro and weighted averages of precision and recall further confirm the balanced performance of the proposed hybrid model.Overall, the results indicate that the GA-optimized ANN model outperforms the standalone ANN model, making it a reliable and efficient approach for sleep disorder classification.

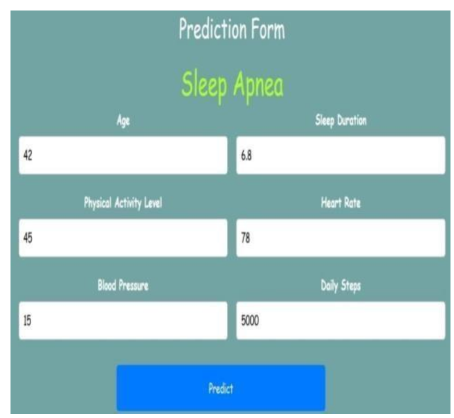
ANN Accuracy : 93.33

Genetic Algorithm + ANN Accuracy : 97.33

	precision	recall	f1-score	support
0	0.87	0.87	0.87	15
1	1.00	0.95	0.98	44
2	0.83	0.94	0.88	16
accuracy			0.93	75
macro avg	0.90	0.92	0.91	75
weighted avg	0.94	0.93	0.93	75

	precision	recall	f1-score	support
0	0.93	0.93	0.93	15
1	1.00	1.00	1.00	44
2	0.94	0.94	0.94	16
accuracy			0.97	75
macro avg	0.96	0.96	0.96	75
weighted avg	0.97	0.97	0.97	75





The confusion matrix for the Genetic Algorithm + ANN model shows strong classification performance across all three classes. It correctly predicted 14 instances of Class 0, 44 of Class 1, and 15 of Class 2, indicating high accuracy. Class 1 achieved perfect classification with no misclassifications, showing excellent precision and recall. Minor errors occurred between Class 0 and Class 2, with one instance from each class misclassified as the other. This suggests slight similarity in feature patterns between these two classes. Overall, the model maintains high precision, recall, and F1-score, and can be improved further with more data and better feature optimization.

5.2 System Usability

During the testing phase, the sleep disorder classification system built using a Genetic Algorithm optimized Artificial Neural Network (GA-ANN) achieved a validation accuracy of 97.33%, indicating strong generalization capability. The training and validation loss curves showed stable convergence, confirming that the model effectively learned sleep pattern features without overfitting. To enhance robustness, preprocessing and feature optimization techniques were applied, enabling the model to accurately interpret variations in physiological and lifestyle attributes such as sleep duration, BMI, stress levels, and heart rate. The system was deployed through a user-friendly web interface, allowing users to input their health and sleep parameters and receive instant predictions of sleep disorder categories. The lightweight ANN architecture ensured fast processing time, making it suitable for real-time clinical and personal health monitoring applications. The confusion matrix showed high precision and recall across most classes, though minor misclassifications occurred between closely related conditions such as insomnia and sleep apnea. Despite this, the system demonstrated high usability by delivering consistent and reliable predictions. Future enhancements may include integrating wearable sensor data and temporal sleep pattern analysis to further improve accuracy and user experience.

5.3 Comparison With Traditional Methods

Traditional sleep disorder classification methods often relied on manual analysis of clinical records, statistical techniques, and rule-based systems, which required handcrafted feature selection and domain expertise. These approaches were time-consuming, less scalable, and often sensitive to incomplete or noisy health data, leading to lower accuracy and limited generalization. In contrast, machine learning and deep learning models, particularly the proposed Genetic Algorithm optimized Artificial Neural Network (GA-ANN), automatically learn complex relationships among features such as sleep duration, BMI, stress level, and physiological parameters. Unlike traditional models, the GA component optimizes feature selection and network parameters, improving accuracy and reducing overfitting. The GA-ANN model demonstrated superior performance with an accuracy of 97.33%, outperforming conventional ML approaches in both precision and reliability. Furthermore, the end-to-end learning capability of ANN reduces manual intervention and enhances scalability for real-world healthcare applications. This shift toward intelligent ML and DL techniques significantly improves the effectiveness and practical usability of sleep disorder classification systems.

5.4 Future Work

Future enhancements for the sleep disorder classification system can focus on several important directions. Incorporating temporal sleep pattern analysis and real-time monitoring using wearable devices will allow the system to capture continuous physiological signals such as heart rate, oxygen levels, and sleep cycles for improved prediction accuracy. Integrating multimodal data sources, including polysomnography (PSG), actigraphy, and lifestyle factors, can provide deeper insights into complex sleep disorders. Developing lightweight and efficient models will enable deployment on mobile and embedded healthcare devices, increasing accessibility for remote monitoring. Expanding the dataset with more diverse patient populations and clinical conditions will improve robustness and generalization. Finally, ensuring data privacy, secure storage, and compliance with healthcare regulations will be essential for safe real-world adoption.

VI. CONCLUSION

The sleep disorder classification system, built using a Genetic Algorithm–optimized Artificial Neural Network (GAANN), reached an impressive validation accuracy of 97.33%, highlighting its strong capability to correctly identify various sleep disorders. By automating feature selection and optimizing network parameters, the GA-ANN approach showed improved robustness and precision compared to traditional machine learning methods. The use of preprocessing and feature optimization techniques further enhanced the model's generalization capability, enabling reliable performance across diverse patient data and health conditions. The system was implemented through a user-friendly interface that allows users to input sleep and health parameters and obtain instant predictions, supporting practical healthcare applications. Although minor misclassifications were observed between closely related conditions such as insomnia and sleep apnea, the system maintains a high level of reliability and usability. Future improvements may include integrating real-time wearable data and expanding the dataset for better performance. Overall, this project demonstrates the significant potential of hybrid ML-DL approaches in advancing intelligent sleep disorder diagnosis systems.

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, doi: 10.3390/e21121203.

[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, doi: 10.1016/j.ymeth.2022.05.006.

[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, doi: 10.1109/TIM.2018.2799053.

[4] D. Shrivastava, S. Jung, M. Saadat, R. Sirohi, and K. Crewson, "How to interpret the results of a sleep study," J. Community Hosp. Intern. Med. Perspect., vol. 4, no. 5, p. 24983, Jan. 2014, doi: 10.3402/jchimp. v4.24983.

[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, doi: 10.3390/diagnostics12010116.

[6] J. Van Der Donckt et al., "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, Art. no. 104429, Mar. 2023, doi: 10.1016/j.bspc.2022.104429.

[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, doi: 10.18201/ijisae.2017.262.

[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, Art. no. 046059, Aug. 2021, doi: 10.1088/1741-2552/abfa71.

[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, doi: 10.3390/diagnostics11040612.

[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, Art. no. 108312, Aug. 2019, doi: 10.1016/j.jneumeth.2019.108312.