



International Journal of Advanced Research in Education and Technology (IJARETY)

Volume 13, Issue 3, May - June 2026

Impact Factor: 9.971



Improving Sleeping Disorder Diagnosis through Optimized Machine Learning Approaches

S. Niranjani¹, M. Sagarika², K. Karthik³, M. Manideepak⁴

Assistant Professor, Department of CSE, Guru Nanak Institute of Technology, Hyderabad, Telangana, India

Students, Department of CSE, Guru Nanak Institute of Technology, Hyderabad, Telangana, India

ABSTRACT: Classifying sleep disorders such as obstructive sleep apnea and insomnia is critical for improving human health and quality of life. Traditional expert-based classification methods, particularly visual inspection, are often error-prone and time-consuming. To address these limitations, machine learning algorithms (MLAs) are increasingly being used to analyze, monitor, and diagnose sleep disorders. This study compares the effectiveness of different MLAs for sleep disorder classification using the Sleep Health and Lifestyle Dataset. Two experiments were conducted: in the first, five key features were selected using the Gradient Boosting Regressor with the Mean Decrease Impurity (MDI) technique, while in the second, two key features were chosen using the same method. The existing system employs the Random Forest classifier, which uses an ensemble of decision trees to categorize sleep stages. However, this approach suffers from computational inefficiency and limited generalization. The proposed system introduces XG Boost, a powerful gradient boosting algorithm that improves classification accuracy and efficiency through optimized tree pruning and regularization techniques. XG Boost demonstrates superior performance by effectively capturing feature interactions and reducing overfitting. This study highlights the potential of XG Boost in advancing sleep disorder classification, providing a more reliable and scalable solution for clinical applications.

I. INTRODUCTION

Sleep disorders, such as obstructive sleep apnea and insomnia, pose significant challenges to human health and well-being, affecting millions of individuals worldwide. Accurate classification of these disorders is essential for timely diagnosis and effective treatment, ultimately improving the quality of life. Traditionally, sleep disorder classification has relied on expert-based methods, including visual inspection of sleep patterns through polysomnography (PSG) recordings. However, these conventional approaches are often time-consuming, prone to human error, and require significant expertise. To address these limitations, machine learning algorithms (MLAs) have emerged as a powerful tool for analyzing sleep patterns, offering automated, data-driven solutions for diagnosing sleep disorders. MLAs can process vast amounts of sleep-related data efficiently, identifying complex patterns that may be difficult to detect through manual inspection.

1.1 SCOPE OF THE PROJECT

The scope of the project “Improved Sleeping Disorder Diagnosis Through Optimized Machine Learning Approaches” is to develop an intelligent and efficient machine learning-based system for the early detection and accurate diagnosis of sleeping disorders. Sleep disorders have become a common health problem affecting people of different age groups, leading to issues such as fatigue, stress, poor concentration, and serious health complications. This project focuses on using machine learning techniques to analyze sleep health and lifestyle data to identify disorders such as insomnia, sleep apnea, and other sleep-related conditions in an automated manner.

II. LITERATURE SURVEY

Detection and Prediction of Sleep Disorders by Covert Bed-Integrated RF Sensors Zijing Zhang, Thomas B. Conroy, Ana C Krieger, Edwin C. Kan 2024

Respiratory disturbances during sleep are a prevalent health condition that affects a large adult population. The gold standard to evaluate sleep disorders including apnea is overnight polysomnography, which requires a trained technician for live monitoring and post-processing scoring.

A Hierarchical Approach for the Diagnosis of Sleep Disorders Using Convolutional RNN

A. Wadichar, S. Murarka, D. Shah, A. Bhurane, M. Sharma, H. S. Mir, and U. R. Acharya 2024.

Sleep disorders significantly impact human health, necessitating accurate and efficient diagnostic methods. Traditional manual approaches to sleep disorder classification are time-consuming and prone to human error. This study proposes a

hierarchical approach for sleep disorder diagnosis using a **Convolutional Recurrent Neural Network (CRNN)**, which combines the feature extraction capability of CNNs with the sequential learning ability of RNNs.

Enhancing Accuracy of Sleep Disorder with Logistic Regression Model

Tanishq Soni, Deepali Gupta, Mudita Uppal 2023.

Now-a-days, sleeping disorder are a topic of concern among the youth and person with different health problems. It is important to note that sleep problems are not a recent occurrence and are not something to be sought after or embraced as a trend, even if it is true that sleep-related concerns have gained greater attention and discussion in recent years as a result of increasing knowledge and study. This disorder can create various problems like anxiety, depression and burden on person's mental health. In few past year's artificial intelligence and machine learning has become a powerful and useful tool in market.

Daily River flow Simulation Using Ensemble Disjoint Aggregating M5-Prime Model

Khabat Khosravi, Nasrin Fathollahzadeh Attar, Sayed Bateni, Changhyun Jun 2023

Accurate prediction of daily river flow (Q_t) is a challenging task in hydrological modeling, particularly vital for flood mitigation and water resource management. This study introduces an advanced M5 Prime (M5P) predictive model designed to estimate Q_t and one- and two-day-ahead river flow forecasts (i.e. Q_{t+1} and Q_{t+2}).

III. EXISTING SYSTEM

The existing system for sleep disorder classification primarily relies on the Random Forest classifier, an ensemble learning method that constructs multiple decision trees and aggregates their outputs to improve prediction accuracy. Random Forest is known for its robustness in handling noisy data and reducing overfitting through bagging techniques. It has been widely used for medical data classification due to its ability to handle large feature spaces and provide feature importance rankings. In this approach, multiple trees are trained on different subsets of the dataset, and the final classification is determined based on majority voting.

Disadvantages of existing system:

- High Computational Cost
- Slow Training and Prediction
- Lack of Interpretability
- Overfitting Risk
- Limited Feature Interactions

IV. PROPOSED SYSTEM

The proposed system aims to automate and enhance the loan approval process using machine learning, specifically a Decision Tree classifier with an accuracy of 80%. By integrating AI-driven predictions, the system reduces manual effort and improves decision-making efficiency. A web-based interface built with Flask allows users to input key financial and demographic details, which are then processed by the trained model to determine loan approval status. To ensure accessibility and reliability, the application is deployed on a cloud-based platform (Render) with an automated reloading mechanism for continuous availability. The system also incorporates data validation techniques to prevent errors and ensure accurate predictions. By minimizing human bias and optimizing processing time, the proposed system provides a faster, fairer, and more transparent loan evaluation framework, benefiting both financial institutions and borrowers.

Advantages of proposed system:

1. Higher Accuracy
2. Faster Computation
3. Better Generalization
4. Feature Importance Insights
5. Handles Missing Data Efficiently

V. SYSTEM ARCHITECTURE

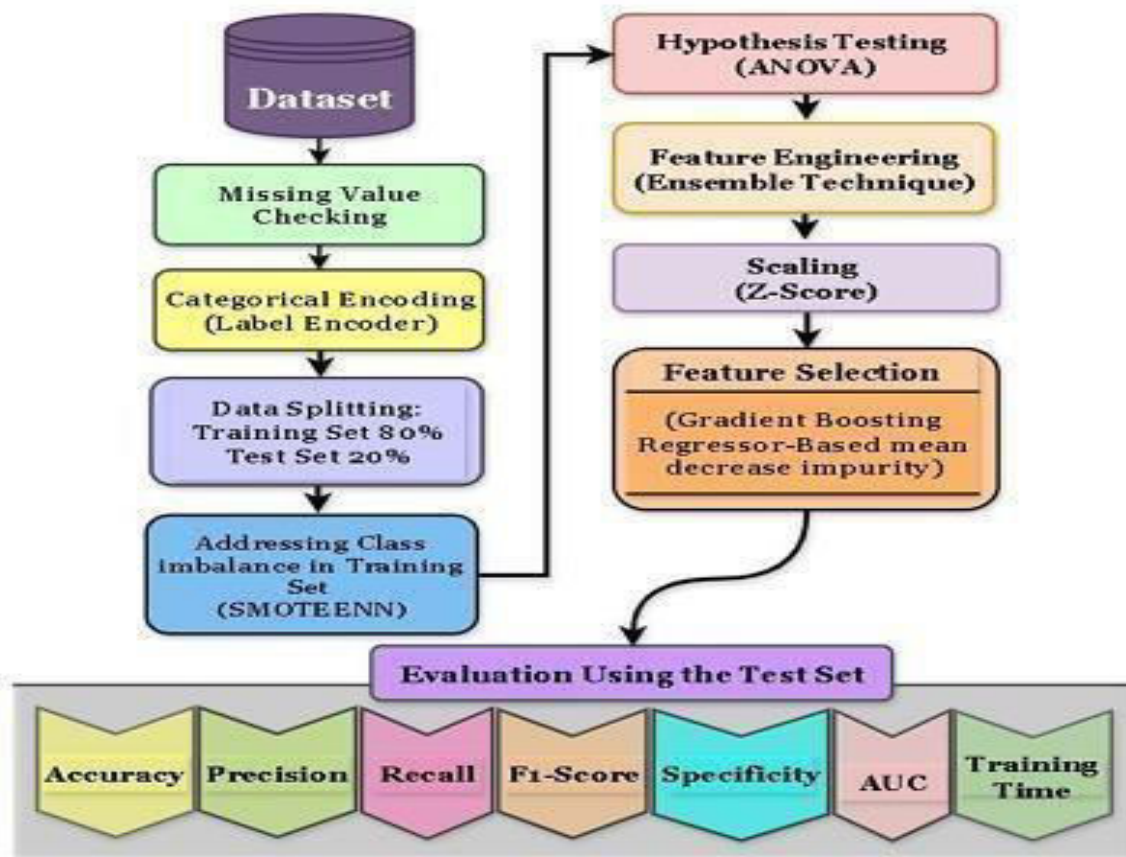


Fig 1.1 System Architecture

The system architecture for the Improved Sleeping Disorder Diagnosis Through Optimized Machine Learning Approaches project begins with collecting the sleep disorder dataset containing patient health and sleep-related information. The data is first preprocessed by checking for missing values and converting categorical data into numerical form using label encoding. preprocessing, the dataset is divided into training (80%) and testing (20%) sets, and class imbalance in the training data is handled using the SMOTEENN technique.

VI. METHODOLOGY

Modules Name:

1. Compiling data
2. Deciphering data
3. Standardizing dataset
4. Processing data
5. Configuring algorithm

Modules Explanation:**1) Compiling Data:**

This phase involves gathering relevant sleep-related datasets, such as the Sleep Health and Lifestyle Dataset. The data is sourced from medical records, wearable devices, and sleep monitoring studies to ensure a comprehensive and diverse dataset.

2) Deciphering Data:

In this step, the collected data is analyzed to identify key patterns, missing values, and inconsistencies. Exploratory Data Analysis (EDA) techniques are used to understand the distribution of features and their correlation with sleep disorders

3) Standardizing Datasets:

To enhance model performance, the dataset undergoes preprocessing, including normalization, handling missing values, and feature selection. The Mean Decrease Impurity (MDI) technique is used to retain only the most impactful features for classification.

4) Processing Data:

The cleaned and standardized dataset is fed into the machine learning pipeline. Various preprocessing steps, such as data augmentation and transformation, are applied to enhance model generalization and accuracy.

5) Configuring the Algorithm:

Machine learning models, including Random Forest and XGBoost, are trained on the dataset. Hyperparameter tuning techniques such as Grid Search and Bayesian Optimization are employed to optimize performance.

VII. ANALYSIS AND DISCUSSION

XG Boost is an optimized gradient boosting algorithm designed for high-performance machine learning tasks. Unlike Random Forest, which builds trees independently, XG Boost constructs trees sequentially, where each new tree corrects the errors made by previous ones. It uses a gradient descent approach to minimize loss functions, making it highly efficient in capturing complex relationships within the data.

1.Usability

The system is designed with completely automated process hence there is no or less user intervention.

2.Reliability

The system is more reliable because of the qualities that are inherited from the chosen platform python. The code built by using python is more reliable.

3.Performance

This system is developing in the high level languages and using the advanced back-end technologies it will give response to the end user on client system with in very less time.

4.Supportability

The system is designed to be the cross platform supportable. The system is supported on a wide range of hardware and any software platform, which is built into the system.

5.Implementation

The system is implemented in web environment using Jupyter notebook software. The server is used as the intelligence server and windows 10 professional is used as the platform. Interface the user interface is based on Jupyter notebook provides server system.

VIII. CONCLUSION

Sleep disorder classification plays a crucial role in diagnosing and managing conditions such as obstructive sleep apnea and insomnia. Traditional methods, such as manual visual inspection, are prone to errors and require significant time and expertise. This study has demonstrated the potential of machine learning algorithms, particularly XG Boost, in enhancing the accuracy and efficiency of sleep disorder classification. By leveraging advanced feature selection techniques and optimized classification algorithms, the proposed system outperforms traditional approaches, offering a more reliable and scalable solution for clinical applications. The findings of this study highlight the advantages of XG Boost in capturing feature interactions, reducing overfitting, and improving computational efficiency.

IX. FUTURE ENHANCEMENT

The future scope of the sleep disorder detection system can be significantly expanded by integrating real-time sleep monitoring using wearable devices such as smartwatches and fitness trackers. These devices can continuously collect sleep-related data, enabling accurate and real-time analysis of sleep patterns. The project can also be improved by implementing advanced machine learning and deep learning techniques to enhance prediction accuracy and identify complex sleep disorder patterns more effectively. In addition, a user-friendly mobile or web application can be developed so that both patients and doctors can conveniently access the system from anywhere, improving usability and healthcare accessibility.

The system can further be enhanced by providing personalized health recommendations, including sleep improvement tips, health alerts, and treatment suggestions based on individual patient data. Future versions of the project may also support the detection of additional sleep disorders by training the model on larger and more diverse datasets, which would improve the overall performance and reliability of the system. Integration with hospital management systems and cloud platforms can make data storage, sharing, and healthcare implementation more efficient and practical in real-world environments.

REFERENCES

- [1] L. Tharmalingam. (2024). Sleep Health and Lifestyle Dataset. Accessed: Jun. 26, 2024. [Online]. Available: <https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset>
- [2] T. S. Alshammari, "Applying machine learning algorithms for the classification of sleep disorders," IEEE Access, vol. 12, pp. 36110–36121, 2024.
- [3] X. Xiao, Y. Rui, Y. Jin, and M. Chen, "Relationship of sleep disorder with neurodegenerative and psychiatric diseases: An updated review," Neurochemical Res., vol. 49, no. 3, pp. 568–582, Mar. 2024.
- [4] M. Braun, S. Dietz-Terjung, U. Sommer, C. Schoebel, and C. Heiser, "Stated patient preferences for overnight at-home diagnostic assessment of sleep disorders," Sleep Breathing, vol. 28, no. 5, pp. 1939–1949, Oct. 2024.
- [5] 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.
- [6] S. K. Satapathy and D. Loganathan, "Automated classification of multi class sleep stages classification using polysomnography signals: A nine layer 1D-convolution neural network approach," Multimedia Tools Appl., vol. 82, no. 6, pp. 8049–8091, Mar. 2023.
- [7] G. Kong, C. Li, H. Peng, Z. Han, and H. Qiao, "EEG-based sleep stage classification via neural architecture search," IEEE Trans. Neural Syst. Rehabil. Eng., vol. 31, pp. 1075–1085, 2023.
- [8] Y. Shao, B. Huang, L. Du, P. Wang, Z. Li, Z. Liu, L. Zhou, Y. Song, X. Chen, and Z. Fang, "Reliable automatic sleep stage classification based on hybrid intelligence," Comput. Biol. Med., vol. 173, May 2024, Art. no. 108314.
- [9] X. Ji, Y. Li, P. Wen, P. Barua, and U. R. Acharya, "MixSleepNet: A multi type convolution combined sleep stage classification model," Comput. Methods Programs Biomed., vol. 244, Feb. 2024, Art. no. 107992.
- [10] C.-H. Tai, T.-Y. Liao, S.-P. Chen, and M.-H. Chung, "Sleep stage classification using light gradient boost machine: Exploring feature impact in depressive and healthy participants," Biomed. Signal Process. Control, vol. 88, Feb. 2024, Art. no. 105647.
- [11] 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.
- [12] 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.
- [13] A. Şenol, T. Talan, and C. Aktürk, "A new hybrid feature reduction method by using MCMSTClustering algorithm with various feature projection methods: A case study on sleep disorder diagnosis," Signal, Image Video Process., vol. 18, no. 5, pp. 4589–4603, Jul. 2024.
- [14] 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.
- [15] W. Z. T. Tareq, "Sleep disorders detection and classification using random forests algorithm," in Decision Making in Healthcare Systems. Cham, Switzerland: Springer, 2024, pp. 257–266.
- [16] M. Warunlawan, P. Homsud, P. Sapphaphab, O. Rinthon, and S. Pechprasarn, "Identification of crucial factors in sleep quality using machine learning models and MRMR feature selection technique," in Proc. 15th Biomed. Eng. Int. Conf. (BMEiCON), Oct. 2023, pp. 1–5.
- [17] H. Zhu, Y. Wu, Y. Guo, C. Fu, F. Shu, H. Yu, W. Chen, and C. Chen, "Towards real-time sleep stage prediction and online calibration based on architecturally switchable deep learning models," IEEE J. Biomed. Health Informat., vol. 28, no. 1, pp. 470–481, Jan. 2024. 21001 VOLUME 13, 2025 M. A. Rahman et al.: Improving Sleep Disorder Diagnosis Through Optimized Machine Learning Approaches
- [18] Z. Zhang, T. B. Conroy, A. C. Krieger, and E. C. Kan, "Detection and prediction of sleep disorders by covert bed-integrated RF sensors," IEEE Trans. Biomed. Eng., vol. 70, no. 4, pp. 1208–1218, Apr. 2023.
- [19] A. Wadichar, S. Murarka, D. Shah, A. Bhurane, M. Sharma, H. S. Mir, and U. R. Acharya, "A hierarchical approach for the diagnosis of sleep disorders using convolutional recurrent neural network," IEEE Access, vol. 11, pp. 125244–125255, 2023.



International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 9.971