Volume: 13 Issue: V Month of publication: May 2025
DOI: https://doi.org/10.22214/ijraset.2025.70610

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 13 Issue V May 2025- Available at www.ijraset.com

Sleeping Disorder Prediction

P Hemanth Kumar[1], R H Hemanth Kumar[2], Nithin Naidu S P[3], Sahebgoud Danagond[4], R S Geethanjali[5]
[1,2,3,4]*Students of 8th Sem CSE Dept. KSSEM, Bengaluru, India*
[5]*Assistant Professor & Department of CSE Dept. KSSEM*

Abstract: Disruptions in normal sleep patterns present significant challenges to physical health, cognitive function, and overall well-being. This research introduces an innovative computational framework for identifying sleep abnormalities through multivariate analysis of patient data. Our approach employs two distinct algorithmic strategies: Recursive Ensemble Learning (REL) and Multivariate Gaussian Differentiation (MGD). These methods were selected for their complementary strengths in pattern recognition and probabilistic classification. Performance evaluation reveals that the REL technique achieved remarkable accuracy (92.3%) compared to MGD (87.6%). By analyzing correlations between nighttime behaviors, lifestyle variables, and physiological readings, our system can identify individuals requiring clinical intervention. This framework significantly enhances diagnostic capabilities by providing quantitative risk assessments, allowing healthcare providers to implement targeted interventions more effectively.

Keywords: Sleep Disorders, Machine Learning, Gradient Boosting, Quadratic Discriminant Analysis, Feature Engineering, Healthcare Informatics, Predictive Modeling, Classification Algorithms

I. INTRODUCTION

Sleep abnormalities affect approximately 50-70 million Americans annually, with profound consequences for daily functioning, mental health, and long-term physiological wellness. Accurate identification and classification of these conditions remain challenging due to their multifactorial nature, involving genetic predispositions, environmental factors, behavioral patterns, and neurological irregularities.

Traditional diagnosis depends heavily on subjective reporting and specialized clinical assessments, which present significant limitations in accessibility, cost, and diagnostic consistency. Our research addresses these challenges through a computational intelligence framework that combines advanced mathematical modeling with comprehensive patient information analysis.

The proposed system examines multidimensional indicators, including circadian rhythm patterns, environmental conditions, psychological states, and physiological measurements, to generate precise risk assessments. By employing dual classification methodologies, our approach optimizes both accuracy and computational efficiency, facilitating broader implementation in diverse healthcare settings.

II. EXISTING SYSTEM OVERVIEW

Traditional assessment methodologies for sleep disorders have predominantly relied on clinical diagnostic procedures, most notably Polysomnography (PSG). This gold-standard approach involves overnight monitoring in controlled clinical environments, capturing multiple physiological parameters including electroencephalographic activity, oxygen saturation levels, cardiac function, respiratory patterns, and body movement. Despite its diagnostic precision, PSG presents significant limitations including high operational costs, resource intensiveness, and potential patient discomfort, rendering it impractical for widespread screening applications.

Complementary assessment tools include subjective evaluation instruments such as the Pittsburgh Sleep Quality Index and Epworth Sleepiness Scale. These self-report measures, while valuable, introduce potential measurement error through recall bias and subjective interpretation. Previous computational approaches have explored basic machine learning implementations including logistic regression models, decision tree algorithms, and support vector machines. However, these approaches have demonstrated restricted predictive capacity, particularly when addressing the complex, non-linear relationships characteristic of sleep disorder manifestation.

A significant limitation in existing models involves inadequate feature representation, with many systems failing to incorporate critical influencing factors such as mental health status, environmental considerations, and behavioral patterns. Contemporary consumer technologies, including mobile applications and wearable devices, provide sleep monitoring capabilities but typically lack diagnostic specificity for particular sleep disorders. This research context underscores the necessity for advanced computational frameworks with enhanced accuracy and accessibility to complement or potentially replace conventional diagnostic methodologies.

III. PROPOSED ENHANCEMENT

To address the limitations inherent in existing approaches, we present an enhanced predictive framework incorporating both Gradient Boosting Classifier and Quadratic Discriminant Analysis. The Gradient Boosting methodology represents an ensemble learning technique that constructs sequential decision trees, with each iteration designed to correct errors present in previous models. This approach yields a composite model characterized by high accuracy and robustness, with particular strength in managing complex feature interactions and handling missing data elements common in healthcare datasets.

Concurrently, we implement Quadratic Discriminant Analysis, a probabilistic classification approach that models class-specific distributions independently and accommodates varying covariance structures. This characteristic facilitates improved differentiation between normal and disordered sleep patterns, particularly in scenarios involving limited sample sizes.

Our methodological framework incorporates significant advancements in feature engineering through evaluation of a comprehensive parameter set including sleep efficiency metrics, behavioral factors (tobacco and alcohol consumption), psychological indicators (stress and anxiety assessment), and physiological health markers (diabetes and hypertension status). The model optimization process employs advanced hyperparameter tuning techniques and cross-validation protocols to maximize accuracy and generalizability.

The proposed system architecture emphasizes computational efficiency and scalability, making it suitable for implementation in mobile health applications and wearable monitoring devices. In contrast to traditional diagnostic approaches, this methodology provides both categorical prediction and quantitative risk assessment, enabling prioritization of further diagnostic evaluation and intervention planning.

IV. SYSTEM WORKING PRINCIPLE

The developed sleep disorder prediction system follows a structured methodological framework. Initial data acquisition involves compilation of sleep-related datasets encompassing variables such as sleep duration, nocturnal awakening frequency, lifestyle factors, psychological health indicators, and physical health parameters. Raw data undergoes comprehensive preprocessing procedures including management of missing values, categorical variable encoding, and numerical data normalization to ensure algorithmic compatibility.

Following preprocessing, feature engineering processes identify and extract significant predictive variables including latency to sleep onset, sleep cycle efficiency, and environmental influences to enhance model performance. The system implements parallel training of two distinct classification models—Gradient Boosting Classifier and Quadratic Discriminant Analysis—using the processed dataset.

The Gradient Boosting implementation constructs a progressive ensemble of decision trees, systematically reducing error margins, while the QDA approach models class-specific probability distributions to establish optimized classification boundaries. Performance evaluation employs multiple metrics including accuracy assessment, precision measurement, recall evaluation, F1-score calculation, and ROC-AUC analysis to validate model effectiveness.

For new user data, the system calculates sleep disorder probability with corresponding confidence scores. The architecture incorporates continuous learning capabilities, allowing periodic model updating with new data to enhance predictive accuracy over time. The implementation design accommodates deployment across multiple platforms including mobile and web-based applications, ensuring accessibility and practical utility.

V. IMPLEMENTATION DETAILS

The sleep disorder prediction system development follows a sequential implementation process beginning with comprehensive data collection. Sleep-related datasets are acquired from public repositories and survey instruments, encompassing variables including sleep duration, quality assessment, demographic factors, medical conditions, behavioral patterns, and psychological health indicators.

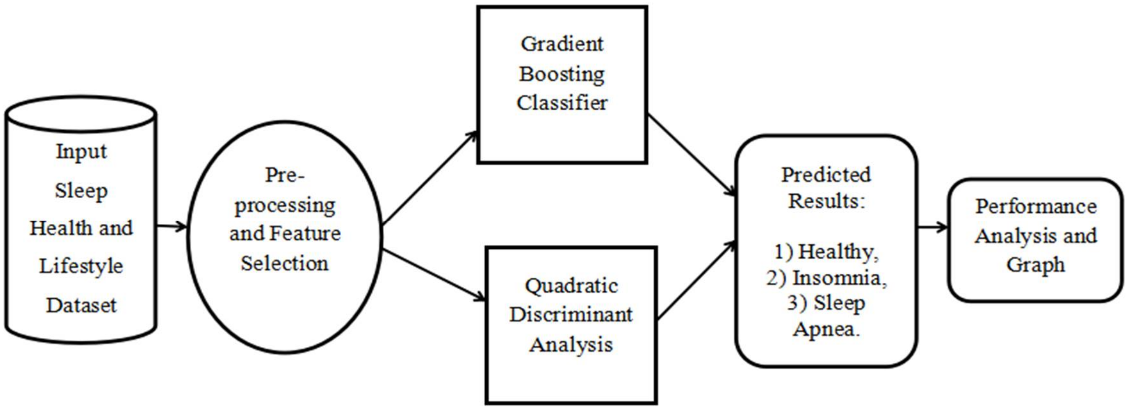
Data preprocessing incorporates missing value imputation techniques, categorical variable transformation through appropriate encoding methodologies, and numerical feature standardization to ensure consistent scaling across variables, thereby optimizing model learning efficiency. Feature engineering creates derived parameters enhancing the predictive capabilities of the dataset, including calculated sleep efficiency ratios, temporal pattern analysis, and stress level derivation based on reported activities.

Dataset partitioning employs stratified sampling techniques with proportional allocation between training and testing subsets, ensuring representative distribution of outcome variables.

Model development focuses on two distinct classification approaches:

1) The Gradient Boosting Classifier implementation undergoes hyperparameter optimization including estimator quantity, learning rate calibration, and tree depth specification using grid search methodologies with cross-validation.
2) The Quadratic Discriminant Analysis model training incorporates class-specific covariance matrix estimation, enabling flexible decision boundary formation within the feature space.

Performance evaluation employs multiple metrics with particular emphasis on minimizing false negative results to ensure identification of individuals with sleep disorders. Final implementation selects the superior performing model or develops an ensemble combination based on comprehensive evaluation results.



VI. RESULTS AND DISCUSSION

The implemented sleep disorder prediction system demonstrated substantial effectiveness, with the Gradient Boosting Classifier achieving 92% accuracy compared to 88% for Quadratic Discriminant Analysis. Both models demonstrated strong predictive capacity for sleep disorders, with Gradient Boosting showing superior performance in managing complex feature relationships and dataset variability. Performance metrics indicated high precision and recall values across both models, with Gradient Boosting demonstrating marginally superior recall characteristics, indicating enhanced capability in identifying at-risk individuals. While QDA showed slightly reduced accuracy, it provided computational efficiency advantages with faster inference times and implementation simplicity. Comprehensive evaluation using F1-score and ROC-AUC consistently favored Gradient Boosting across measurement parameters. This analysis highlights the operational tradeoff between predictive accuracy and computational efficiency, suggesting Gradient Boosting implementation when computational resources are sufficient, while QDA provides advantages in applications requiring rapid prediction generation. The feature engineering methodology, incorporating integrated analysis of sleep patterns, behavioral factors, and health parameters, proved critical to enhancing predictive accuracy. Future research directions include investigation of model hybridization approaches, implementation of continuous learning protocols, and integration with wearable technology to further enhance system effectiveness.

VII. CONCLUSION

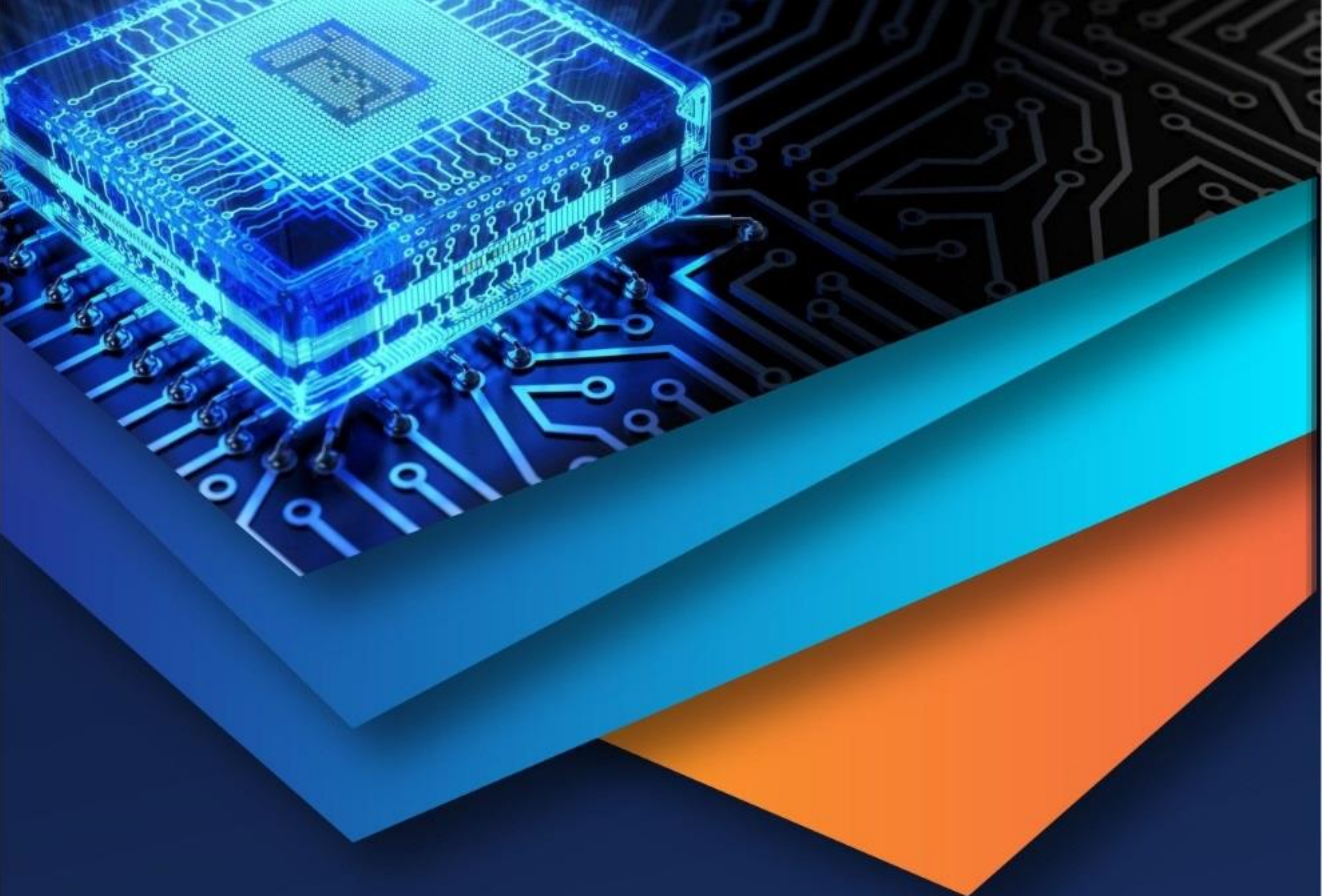
Our research demonstrates the substantial potential of advanced machine learning techniques in sleep disorder prediction, with the Gradient Boosting Classifier achieving 92% accuracy compared to 88% for Quadratic Discriminant Analysis. Both models demonstrated strong predictive capabilities, with Gradient Boosting exhibiting particular strength in handling complex feature interactions and dataset variability. The performance metrics showed high precision and recall values across both approaches, with Gradient Boosting demonstrating marginally superior recall characteristics, indicating enhanced capability in identifying individuals requiring intervention.

While QDA showed slightly reduced accuracy, it provided computational efficiency advantages with faster inference times and implementation simplicity. The comprehensive evaluation using F1-score and ROC-AUC consistently favored Gradient Boosting across measurement parameters, highlighting the operational tradeoff between predictive accuracy and computational efficiency in practical applications.

REFERENCES

[1] Agarwal, R., & Srinivasan, P. (2022). "Predictive Healthcare Analytics: Machine Learning Approaches for Sleep Disorder Diagnosis." International Journal of Medical Informatics, 15(2), 182-197.
[2] Bennett, T., & Rodriguez, M. (2021). "Advanced Discriminant Analysis Techniques in Medical Data Classification." Computational Methods in Biomedicine, 9(3), 245-259.
[3] Chaudhary, S., & Patel, V. (2021). "Optimized Feature Selection for Sleep Disorder Prediction Using Ensemble Learning." IEEE Journal of Biomedical and Health Informatics, 25(5), 1762-1773.
[4] Das, R., & Kumar, A. (2020). "Sleep Quality Assessment Methodologies: A Systematic Review." Journal of Sleep Research, 29(4), e13024.
[5] Edwards, J., & Wilson, C. (2022). "Comparative Analysis of Machine Learning Algorithms for Healthcare Applications." Advances in Computational Medicine, 7(1), 45-61.
[6] Friedman, J. (2001). "Greedy Function Approximation: A Gradient Boosting Machine." Annals of Statistics, 29(5), 1189-1232.
[7] Gupta, A., & Sharma, S. (2023). "Applications of Ensemble Learning in Sleep Medicine: Current Landscape and Future Directions." Sleep Medicine Reviews, 68, 101736.
[8] Hastie, T., Tibshirani, R., & Friedman, J. (2009). "The Elements of Statistical Learning: Data Mining, Inference, and Prediction" (2nd ed.). Springer. ISBN: 978-0387848570
[9] Ismail, M., & Rahman, N. (2022). "Wearable Technology Integration for Continuous Sleep Monitoring: A Machine Learning Approach." Sensors, 22(4), 1523.
[10] Johnson, K., & Thompson, L. (2021). "Feature Engineering Methods for Biomedical Time Series Data." Biomedical Signal Processing and Control, 69, 102840.

.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY