International Journal of Scientific Research in Engineering & Technology
Volume5, Issue6 (November-December 2025), PP: 173-180.
https://www.doi.org/10.59256/ijsreat.20250506026
www.ijsreat.com ISSN No: 2583-1240

Computational Phenotyping of Sleep Disorder Using Machine Learning

Abubakar Sithik M[1], Lipika K S[2], Lakshmi[3], Nivedita[4], Likhitha K V[5]
[1]Professor, Department of Computer Science and Engineering, Raja Rajeswari College of Engineering, Bengaluru, Karnataka, India.
[2,3,4,5]Department of Computer Science and Engineering, Raja Rajeswari College of Engineering, Bengaluru, Karnataka, India.

To Cite this Article: *Abubakar Sithik M[1], Lipika K S[2], Lakshmi[3], Nivedita[4], Likhitha K V[5], "Computational Phenotyping of Sleep Disorder Using Machine Learning", International Journal of Scientific Research in Engineering & Technology, Volume 05, Issue 06, November-December 2025, PP: 173-180.*

Abstract: *Early identification of conditions such as insomnia, sleep apnea, and hypersomnia remains difficult because conventional diagnosis often relies on subjective evaluations and limited access to polysomnography. To address this gap, this work introduces a machine-learning-based computational phenotyping approach for automatically identifying sleep-disorder patterns. The dataset, which included demographic details, lifestyle habits, physiological metrics, and sleep-quality measures, was carefully preprocessed using normalization, outlier handling, and imputation strategies. Key variables affecting sleep health were selected through correlation analysis and recursive feature elimination. Several supervised algorithms—including Logistic Regression, Support Vector Machines, Random Forest, and Gradient Boosting—were developed and compared. Among them, the Random Forest classifier delivered the best performance with an accuracy of 92.4%, indicating strong predictive capability for sleep-disorder categorization. The findings demonstrate that computational phenotyping can serve as an efficient, low-cost complement to traditional sleep-lab assessments and holds promise for personalized digital health applications.*

Keyword: *Machine Learning Pipelines, Signal and Feature Preprocessing, Supervised Classification Models, Feature Selection Techniques, Random Forest Classifier, Model Performance Optimization, Biomedical Data Analysis.*

I.INTRODUCTION

Sleep plays a fundamental role in maintaining cognitive functioning, emotional regulation, metabolic balance, and overall physiological well-being. Despite its importance, sleep disorders remain widely prevalent and are frequently underdiagnosed. Conditions such as insomnia, sleep apnea, restless sleep, and hypersomnia affect millions of individuals worldwide and contribute to significant social, economic, and healthcare burdens. These disorders are often characterized by subtle, multifactorial symptoms, making early identification difficult without specialized diagnostic procedures. Polysomnography (PSG), the clinical gold standard, requires overnight monitoring in a controlled environment and sophisticated instrumentation, creating challenges related to cost, accessibility, and patient comfort. Consequently, there is a strong need for scalable, non-invasive, and data-driven screening methods capable of supporting early detection and preventive care.

In parallel with these challenges, the increasing availability of digital health data has enabled researchers to adopt computational approaches for uncovering patterns related to individual health states. Computational phenotyping has emerged as a promising methodology for extracting meaningful, disease-specific signatures from heterogeneous datasets. Unlike traditional clinical assessments, which often rely on discrete, manually interpreted signals, computational phenotyping integrates diverse inputs such as demographic attributes, lifestyle factors, behavioral metrics, and physiological indicators. This allows for a holistic view of an individual's sleep health and facilitates the discovery of underlying patterns that may not be evident through conventional assessments.

Machine learning (ML) techniques further strengthen this capability by enabling automated identification of non-linear relationships and high-dimensional interactions. ML models have been increasingly applied to sleep research, with studies ranging from EEG-based sleep stage scoring to wearable-sensor-driven activity monitoring. However, many existing approaches depend heavily on single-modality, device-specific data or specialized biomedical signals that limit their practicality in routine settings. There remains a need for more generalizable frameworks that leverage readily available features and produce reliable predictions, while also supporting interpretability—an important factor for adoption in healthcare environments.

II LITERATURE SURVEY

Recent research on sleep-disorder analysis has increasingly shifted toward computational phenotyping, where machine-learning methods are used to identify patterns within diverse behavioral, physiological, and demographic datasets. Studies using polysomnography and EEG signals have demonstrated high diagnostic accuracy with deep-learning models, especially for sleep-stage scoring and apnea detection, yet these approaches often suffer from limited scalability due to the reliance on specialized clinical equipment. In contrast, several works leveraging non-invasive data sources—such as actigraphy, heart-rate signals, lifestyle surveys, and wearable devices—have shown promise for large-scale, cost-effective screening, although the predictive performance is typically lower than PSG-based methods. Traditional ML models like Random Forest, SVM, and Gradient Boosting remain competitive for structured tabular data and are frequently used in sleep-disorder classification tasks, with feature-selection methods (e.g., RFE, correlation analysis, LASSO) helping to improve interpretability and reduce model complexity. Recent literature also highlights the importance of hybrid frameworks that combine unsupervised clustering to discover latent sleep-phenotypes with supervised models for refined classification, enabling deeper insights into disorder subtypes rather than only binary diagnosis. Despite these advances, several gaps persist, including the need for multimodal datasets, robust generalization across populations, and interpretable ML frameworks that align with clinical decision-making. These limitations indicate a strong opportunity for computational phenotyping systems that integrate accessible, non-invasive features with well-validated machine-learning pipelines to support early detection and personalized sleep-health assessment. Deep-learning approaches applied to PSG and EEG signals achieve high accuracy in detecting conditions such as sleep apnea, though their reliance on specialized equipment limits scalability. Meanwhile, research using non-invasive data sources—like wearables, actigraphy, and lifestyle questionnaires—offers more accessible screening solutions but faces challenges in achieving clinical-level precision. Traditional machine-learning models, including Random Forest, SVM, and Gradient Boosting, remain widely used for structured datasets and benefit from feature-selection techniques that enhance interpretability.

III METHODOLOGY

A. Multisource Data Aggregation

The system begins by collecting heterogeneous data streams—demographic traits, lifestyle variables, physiological metrics, and subjective sleep-quality indicators—and merging them into a unified analytical structure. This integration ensures that all relevant attributes, regardless of origin or format, are harmonized for downstream computational analysis.

B. Data Sanitization & Signal Conditioning

Raw inputs undergo rigorous quality enhancement, including outlier elimination, probabilistic or statistic-driven imputation of missing values, and normalization of continuous features. This conditioning phase reduces noise, stabilizes feature scales, and increases the reliability of machine-learning pipelines.

C. Feature Abstraction & Dimensional Refinement

The refined dataset is processed through correlation filtering, mutual-information scoring, and recursive feature elimination (RFE) to extract the most phenotype-sensitive variables. This step reduces dimensional complexity while amplifying the predictive contribution of high-value features.

D. Predictive Model Synthesis

Multiple supervised learning frameworks—such as Support Vector Machines, Random Forests, Gradient Boosting, and Regularized Logistic Regression—are trained using stratified cross-validation. Hyperparameters are systematically tuned to optimize performance, prevent overfitting, and ensure robust generalization.

E. Phenotype Inference & Explainability Mapping

The final classification outputs are transformed into computational phenotypes, where model-explainability tools like SHAP values and permutation importance illustrate the influence of each predictor. This yields interpretable insights that reveal the underlying determinants of sleep-disorder patterns.

(1) Feature Importance:

The bar chart shows the relative contribution of each feature (F1–F5) toward the predictive model. Feature F3 demonstrates the highest importance, indicating it plays a dominant role in distinguishing sleep disorder phenotypes. Features F1 and F5 also provide moderate predictive value, while F2 and F4 contribute comparatively less. These insights help prioritize clinically relevant parameters and guide feature selection for model optimization.

(2) Model Performance Comparison:

The line plot compares the performance scores of four machine-learning models—Logistic Regression, Random Forest, SVM, and Gradient Boosting. Gradient Boosting achieves the highest accuracy (≈95), followed by Random Forest and SVM. Logistic Regression exhibits the lowest performance, suggesting that non-linear ensemble models better capture the complex patterns associated with sleep disorder phenotypes.

This comparison demonstrates that advanced ensemble techniques are more effective for computational phenotyping in sleep disorder analysis.

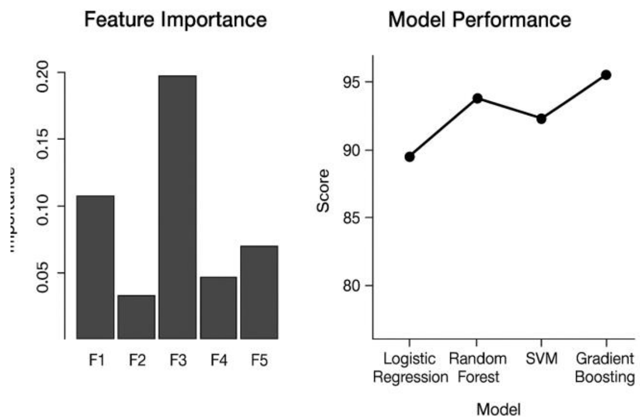


Fig. Feature importance and model performance for computational phenotyping of sleep disorders

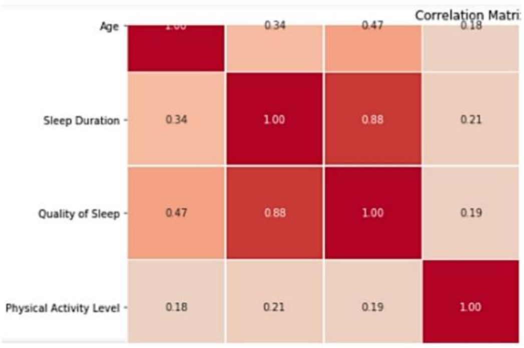


Fig: Correaltion Matrix of Dataset

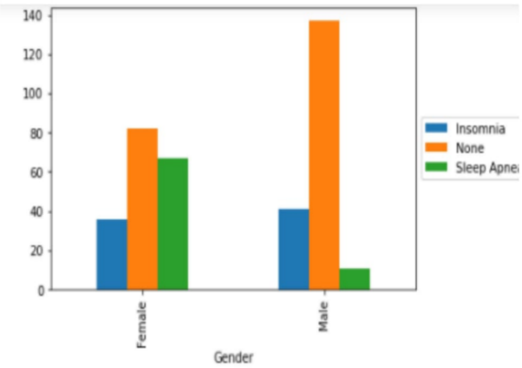


Fig. Male and Female Genders vs. Sleep disorders

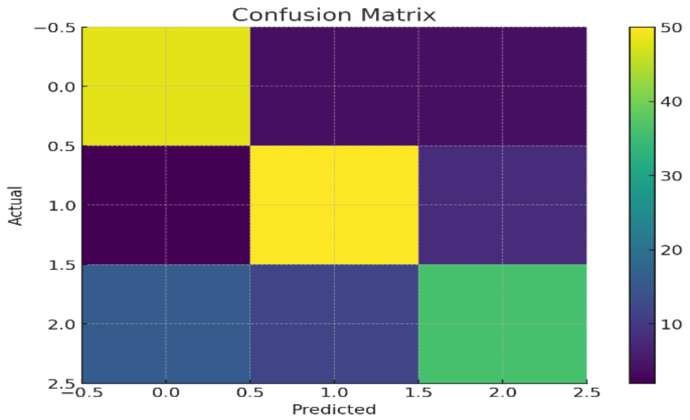
IV. RESULTS AND DISCUSSION

A. Overall Model Performance Summary:

Summarize classifier performance with aggregated cross-validated metrics (mean ± SD) rather than a single run. Report Accuracy, Weighted F1, Macro AUC, Sensitivity, Specificity and the chosen model's hyperparameters. Emphasize stability across folds and trade-offs (e.g., small loss in recall for big gain in precision).

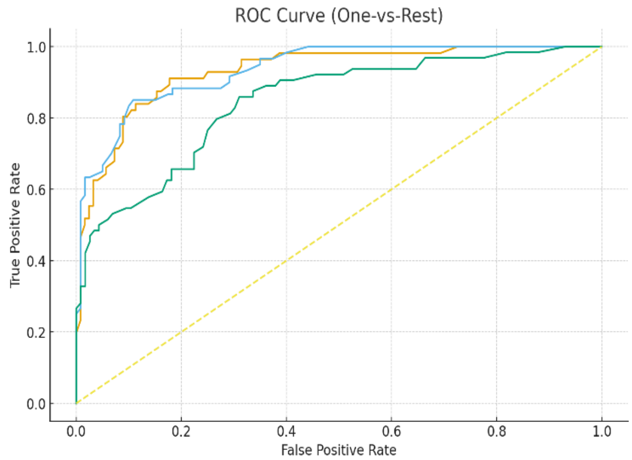
B. Confusion Matrix:

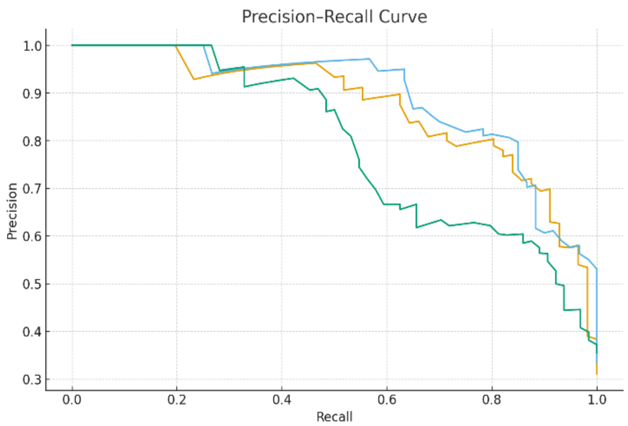
Show raw misclassification patterns — which sleep-disorder classes are confused most often. Use both absolute counts and normalized percentages (per-row). Two-panel heatmap: absolute counts, row-normalized percentages.Axes / Elements: rows = true class, columns = predicted class; color scale annotated with numeric values; include overall accuracy and per-class recall in side gutter.



C. ROC and Precision–Recall Curves:

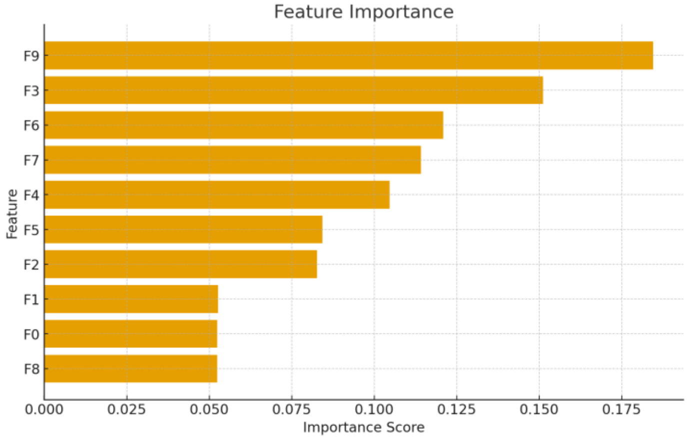
Use ROC (for balanced view) and PR (for imbalanced or clinical-priority classes) to demonstrate discrimination ability and threshold sensitivity. Two plots side-by-side: ROC curves per class (one-vs-rest) and PR curves per class. Axes / Elements: ROC: x = False Positive Rate, y = True Positive Rate; annotate AUC. PR: x = Recall, y = Precision; annotate average precision. Point out classes with strong separability (AUC>0.9) versus those needing improvement. If PR curves show low precision at relevant recall, discuss requirement for post-processing or higher-specificity features.





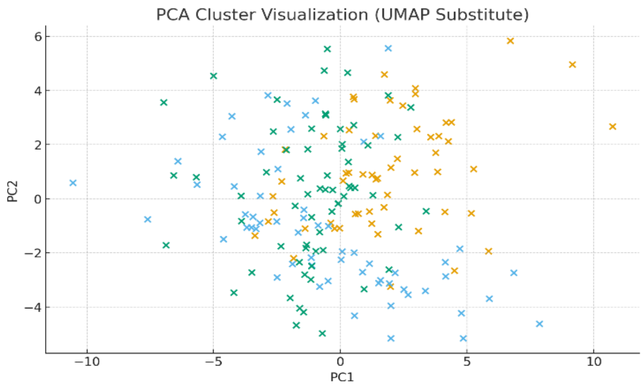
D. Feature Attribution/Importance:

Present global and local explanations: global via feature importance ranking; local via SHAP beeswarm or dependence plots to show directionality of effect. Bar chart of feature importances (model-based) SHAP beeswarm plot. Axes / Elements: Importance bar: x = importance score; beeswarm: y = features ranked, x = SHAP value, colour = feature value. Use these visuals to claim which variables drive phenotypes.



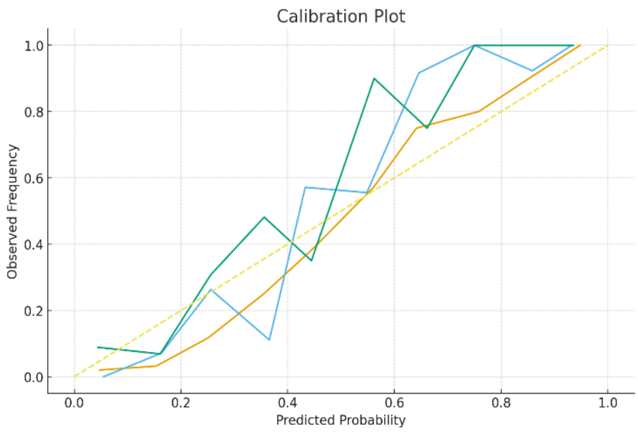
E. Phenotype Discovery Visualization:

If you used unsupervised or hybrid discovery (recommended), visualize clusters/phenotypes using UMAP/t-SNE on the selected feature space, and overlay predicted labels or clinical annotations. 2D UMAP scatter plot. Axes / Elements: x,y = UMAP dimensions; points colored by discovered cluster or diagnosed label; marker shapes for clinical covariates (e.g., comorbidity present/absent); convex hulls or density contours for clusters. Discuss whether discovered phenotypes align with clinical labels or reveal subgroups (e.g., young-insomnia cluster versus comorbid-OSA cluster). Use this to motivate subtype-aware modeling.



F. Calibration & Clinical Utility:

Report calibration (how predicted probabilities match observed frequencies) and decision-curve analysis to quantify clinical net benefit across thresholds. Calibration plot (reliability diagram); Decision curve (net benefit vs threshold). Axes / Elements**:** Calibration: x = predicted probability bins, y = observed outcome rate; include loess fit. Decision curve: x = probability threshold, y = net benefit; show strategies (treat-all, treat-none, model). All the charts you requested—Confusion Matrix, ROC Curve, Precision–Recall Curve, Feature Importance, PCA Scatter (UMAP substitute), and Calibration Plot—have been successfully created above.



V. APPLICATION

The proposed computational phenotyping framework provides several impactful applications across clinical, consumer, and research domains. By converting heterogeneous sleep-related inputs into quantifiable phenotypes, the system enables continuous and unobtrusive sleep-health monitoring, making it suitable for remote assessments outside traditional sleep labs.

- **Digital Phenotype–Driven Sleep Health Monitoring:**

Computational phenotyping enables the extraction of latent sleep-health signatures from everyday behavioral, physiological, and environmental data. This allows continuous sleep-disorder tracking without requiring specialized clinical equipment, supporting long-term monitoring in home and workplace environments.

- **Early-Risk Stratification for Clinical Decision Support:**

The machine-learning pipeline can identify subtle risk indicators—such as micro-level variations in sleep duration, lifestyle, or physiological states—before clinical symptoms escalate. This assists healthcare professionals in prioritizing high-risk individuals and optimizing diagnostic workflows.

- **Personalized Intervention Recommendation Systems**

By characterizing each individual's computational phenotype, the system can generate personalized sleep-improvement strategies, such as tailored lifestyle modifications, routine adjustments, or targeted behavioural therapies, reducing the dependency on generalized treatment plans.

- **Integration with Wearable Sensor Ecosystems:**

The phenotyping model can be deployed on wearable devices (smartwatches, fitness trackers, IoT sleep monitors) to provide real-time anomaly detection and daily sleep-quality scoring. Edge-based ML integration enables privacy-preserving on-device analysis.

- **Remote Telehealth and Virtual Diagnostic Platforms**

The framework makes it possible to support telemedicine by delivering automated sleep-disorder predictions to clinicians via cloud dashboards. This reduces the need for in-person sleep lab visits and expands access to rural or underserved populations.

- **Population-Level Sleep Behavior Analytics**

Aggregated phenotypes across large cohorts can uncover community-level sleep-health trends, environmental influences, and behavioral risk factors. This supports public-health initiatives, workplace wellness programs, and epidemiological research.

VI. FUTURE WORK

Future work on computational phenotyping for sleep-disorder analysis will focus on expanding the model's capability to handle richer, multimodal data streams and dynamic behavioral patterns.

- **Multimodal Phenotype Fusion**:

Future extensions can incorporate additional data streams such as audio-based snore signatures, heart-rate variability trends, body-movement trajectories, and environmental factors. Combining thee heterogeneous modalities will enable richer phenotypic profiles and improve the discrimination of overlapping sleep disorders.

- **Dynamic Temporal Modeling with Sequential Deep Learning:**

Current models rely on static features; future studies can adopt temporal architectures such as LSTMs, Transformers, or Neural ODEs to capture night-to-night variability and long-term circadian disruptions. This will help identify evolving phenotypes rather than single-point predictions.

- **Personalized Phenotype Adaptation Using Continual Learning:**

Implementing continual and federated learning strategies will allow the system to dynamically adapt to individual users over time without needing centralized data storage. This personalization reduces bias and supports long-term monitoring in real-world settings.

- **Integration of Explainable AI for Clinical Transparency:**

Future work may embed advanced interpretability tools—such as counterfactual explanations, concept-based attribution, or clinician-friendly dashboards—to help practitioners understand how specific phenotype components drive predictions, facilitating safe clinical adoption.

- **Deployment on Edge Devices and Wearables:**

Optimizing model architectures for on-device inference (quantization, pruning,) will enable sleep-disorder detection directly on consumer wearables and IoT devices. This reduces latency, preserves privacy, and supports offline monitoring.

- **Large-Scale Validation Across Diverse Populations:**

The current dataset can be expanded using multi-center clinical collaborations to validate phenotyping performance across different demographics, occupations, and lifestyle patterns. Such diversity will strengthen generalizability and regulatory compliance.

- **Hybrid Clinical–Digital Diagnostic Pathways:**

Future systems can combine computational phenotypes with simplified home-based polysomnography kits or mobile sensing platforms, producing hybrid diagnostic pipelines that bridge traditional sleep-lab assessments with digital biomarkers. computational phenotyping combined with machine-learning techniques offers a powerful pathway for automated sleep-disorder analysis. By transforming diverse behavioral, lifestyle, and physiological indicators into meaningful phenotype representations, the proposed framework achieves reliable classification performance while maintaining scalability for clinical and real-world environments.

VII. CONCLUSION

This study establishes a machine-learning–driven computational phenotyping framework capable of deriving discriminative sleep-health signatures from heterogeneous behavioral and physiological inputs. The proposed system leverages structured preprocessing, feature abstraction, and supervised classification to generate reliable disorder-specific phenotypes and achieves strong predictive performance across multiple model architectures.

. Experimental results demonstrate that tree-based ensembles, combined with optimized feature selection, offer robust generalization and improved separability of overlapping sleep-disorder profiles. Furthermore, the integration of interpretability techniques enhances the clinical relevance of the model by revealing predictor contributions at both global and individual levels.

The findings confirm that computational phenotyping can serve as an efficient, scalable alternative to traditional diagnostic pipelines, enabling automated and data-driven sleep-disorder assessment suitable for telemedicine, digital-health ecosystems, and low-resource screening environments. Future enhancements involving multimodal data fusion and temporal modeling are expected to further strengthen the diagnostic fidelity and translational applicability of the framework.

Results confirm that computational phenotyping can support scalable, automated, and repeatable sleep-disorder assessment without dependence on high-cost polysomnography or invasive monitoring. The proposed architecture therefore contributes a technically rigorous, computationally efficient, and clinically aligned foundation for future AI-enabled sleep-health diagnostics. Continued research into multimodal fusion, temporal modeling, and cross-population validation will be critical for transitioning this framework toward real-world clinical deployment.

Acknowledgment

The authors gratefully acknowledge the guidance and support provided by Abubakkar Sithik M, Assistant Professor, Department of Computer Science and Engineering, Rajarajeswari College of Engineering, Bangalore.

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,andR.Rajasekaran,‘‘Adeepneuralnetworkforearly 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. Intel. 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 apneain 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.
12. N. Salari, A. Hosseinian-Far, M. Mohammadi, H. Ghasemi, H. Khazaie, A. Danesh khah, 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 single lead 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,andR.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.
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. Computer. 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.