EVALUATING MACHINE LEARNING MODELS FOR PREDICTING SLEEP DISORDERS IN A LIFESTYLE AND HEALTH DATA CONTEXT

Gregorius Airlangga[1]

[1]Atma Jaya Catholic University of Indonesia
*Email: [1]gregorius.airlangga@atmajaya.ac.id

Abstract

Sleep disorders pose significant challenges to public health due to their complex etiology and substantial impact on life quality. Accurately diagnosing these disorders is crucial yet difficult, owing to the diverse and overlapping symptoms influenced by various lifestyle and health factors. This study addresses the problem of identifying the most effective machine learning (ML) models for diagnosing sleep disorders, given the variability and complexity of related health data. We employed a dataset comprising 400 individual records, categorized by sleep disorder status (none, insomnia, sleep apnea) and featuring demographic, sleep, lifestyle, and health metrics. Our objective was to evaluate and compare a broad spectrum of ML models, including logistic regression, decision trees, ensemble methods like RandomForest and GradientBoosting, support vector machines, and neural networks, in terms of their diagnostic accuracy. These models were assessed using metrics such as accuracy, precision, recall, and F1 score. Results indicated that ensemble methods, particularly RandomForest and XGBClassifier, significantly outperformed others, achieving metrics as high as 0.93. These findings suggest that advanced ensemble techniques are most effective in managing the complex data involved, thereby enhancing the reliability and accuracy of sleep disorder diagnostics in clinical settings. The study advocates for the integration of these robust ML models into clinical practices and calls for further research to optimize these tools for real-world applications.

Keywords: *Machine Learning, Sleep Disorders, Lifestyle, Health Data, Predicting*

Corresponding Author: Gregorius Airlangga

1. INTRODUCTION

In the realm of healthcare, sleep is an essential yet often neglected component of overall health and wellbeing. The importance of sleep cannot be overstated, as it affects both physical and mental health profoundly [1]–[3]. Recent research has demonstrated that insufficient sleep is linked to a range of chronic conditions, from cardiovascular disease to diabetes, and mental health disorders such as depression and anxiety [4]–[6]. The complexity of sleep and its profound impact on human health necessitates sophisticated diagnostic tools to understand and mitigate sleep-related disorders effectively [7]–[9]. Recent advances in technology and data analysis have opened up new avenues in the diagnosis and management of sleep disorders [10]–[12]. Traditionally, sleep studies have been conducted in clinical settings using polysomnography, which, while effective, is costly and inconvenient for patients [13]–

[15]. With the advent of machine learning (ML) and data-driven methodologies, researchers have begun exploring less invasive and more scalable approaches [16]–[18]. Several studies have highlighted the potential of machine learning in sleep research, offering insights into patterns that may not be immediately apparent to clinicians or traditional models [19]–[21].

For instance, ML techniques have been employed to analyze sleep quality and duration through wearable technology data, revealing correlations with various health outcomes [22]. These studies typically leverage large datasets to train models that can predict sleep disorders from an array of inputs such as heart rate, blood pressure, and activity levels [23]. Notably, algorithms such as Random Forests, Support Vector Machines, and Neural Networks have shown promise in identifying sleep disorders like Insomnia and Sleep Apnea with high precision. The urgency for innovative diagnostic tools in sleep medicine is underscored by

the global rise in sleep disorders [24]. This increase is often attributed to modern lifestyle factors such as increased screen time, irregular work hours, and higher stress levels [25]. The state of the art in sleep disorder diagnosis involves an integration of machine learning models that can process vast amounts of data to detect subtle patterns indicative of sleep abnormalities [26]. These advanced models not only offer greater accuracy but also enhance the efficiency of diagnosis, thus potentially reducing the healthcare burden associated with untreated sleep disorders [27]–[29]. The primary goal of this research is to evaluate the effectiveness of various machine learning models in predicting sleep disorders using a comprehensive dataset. This dataset includes not only sleep metrics but also lifestyle and cardiovascular health data, providing a holistic view of factors that influence sleep health. By comparing different ML models, the research aims to identify the most effective algorithms for diagnosing sleep disorders in diverse populations.While significant progress has been made in the application of ML in sleep research, gaps remain in the predictive accuracy and generalizability of these models across different demographic groups. Additionally, there is a lack of comprehensive studies that integrate lifestyle and cardiovascular data into sleep disorder diagnostics at scale. Most existing research focuses on isolated aspects of sleep, without considering the multifactorial nature of sleep disorders.

This research contributes to the field by implementing a comparative analysis of multiple machine learning models on a dataset encompassing a wide range of variables related to sleep and daily habits. By including models ranging from logistic regression to more complex ensemble and deep learning models, this study provides a broad evaluation of the predictive capabilities across various algorithmic approaches. Moreover, the research also delves into the feature importance analysis for tree-based models, offering insights into the most critical factors influencing the predictive models' performance.The remainder of this article is organized as follows: The next section describes the methodology, detailing the data preparation, model training, and evaluation metrics used in the study. This is followed by the results section, where the performance of each model is analyzed and discussed. The discussion section then interprets these findings, considering the implications for clinical practice and future research. Finally, the conclusion summarizes the key takeaways and suggests directions for further investigation in the domain of machine learning applications in sleep medicine.

2. RESEARCH METHOD

The objective of this research is to assess and compare the efficacy of multiple machine learning (ML) models in predicting sleep disorders using a detailed and expansive dataset. This dataset encompasses variables related to demographic, lifestyle, and health-related metrics that potentially influence sleep quality and disorders. This section details the comprehensive methodology adopted for data handling, model development, and evaluation, adhering to robust scientific principles to ensure reliability and validity in the findings.

2.1 Data Collection and Description

The Sleep Health and Lifestyle Dataset serves as the foundational data source for this research, encompassing a detailed collection of 400 individual records, each delineated into 13 specific columns. The dataset can be downloaded from [30] and it is a meticulous amalgamation of both numerical and categorical data, which extensively covers a spectrum of variables integral to our study. These include demographic information such as gender, age, and occupation, which are pivotal in examining the social and economic influences on sleep health.

Demographic variables are instrumental for analyzing trends and disparities in sleep patterns across different groups. For instance, age and occupation are not merely indicative of an individual's stage in life and economic activity but are also reflective of varying stress levels and lifestyle choices which can significantly influence sleep quality and duration. Sleep metrics detailed in the dataset include the duration and quality of sleep. These metrics are quantified, with duration measured in hours and quality assessed on a scale from 1 to 10. This quantification allows for a nuanced analysis of sleep patterns and their direct correlation with health outcomes.

Lifestyle factors recorded encompass the daily physical activity levels and self-reported stress levels of individuals, both of which are measured in quantifiable terms—minutes per day for physical activity and a 1 to 10 scale for stress levels. These factors are critical as they offer insights into the daily behaviors and environmental stresses that can affect sleep quality and overall health. Health parameters in the dataset include Body Mass Index (BMI) categories, blood pressure readings (both systolic and diastolic), heart rate in beats per minute, and daily step count. These metrics provide a broader understanding of the physiological and cardiovascular health of individuals, which are often closely linked with sleep health.

Moreover, the dataset categorizes individuals based on their sleep health, identifying those with no sleep disorder, those suffering from insomnia, and those afflicted with sleep apnea. This classification is vital for the targeted analysis of sleep disorders, enabling the application of machine learning models to predict and understand the prevalence and impact of these disorders. By harnessing this dataset, the research delves into the intricate relationships between lifestyle choices, demographic factors, and their cumulative effects on sleep, thereby providing a robust

platform for predictive analysis and a deeper understanding of sleep disorders in the context of public health.

2.2 Data Preprocessing

The initial step in preprocessing involved a thorough examination of the dataset for missing values and inconsistencies, which are common in real-world data and can significantly skew or bias the results of machine learning models. To address this, a strategy of imputation was adopted where missing continuous variables were filled with the mean of their respective columns. This approach helps maintain the overall distribution of the variable while providing a reasonable approximation for the missing data. For categorical variables, the mode of the column was used for imputation, thus ensuring that the most frequent category was assigned in cases of missing data.

As the dataset contained several categorical variables such as gender, occupation, and BMI category, it was crucial to transform these into a format suitable for machine learning algorithms. One-hot encoding was employed for this purpose, converting each category within the variable into a new binary column. This encoding method facilitates the model's ability to distinguish between different categories without imposing any ordinal relationship between them. Normalization of numerical features was another critical step, aimed at standardizing the range of continuous data features so that each one contributes equally to the analysis. Variables such as age, sleep duration, physical activity level, stress level, blood pressure, heart rate, and daily steps were scaled using the StandardScaler. This scaler adjusts the data to have a mean of zero and a standard deviation of one, effectively normalizing the distribution of these features. Such normalization is pivotal, especially in algorithms that are sensitive to the scale of input data, as it aids in the convergence of the model during training and prevents any single feature from dominating the influence on the predictive outcomes.

The initial models included all available features, providing a baseline from which to assess the relative importance of each variable. However, to enhance the performance and interpretability of the machine learning models, a feature selection process was undertaken. This process was informed by exploratory data analysis and initial feature importance rankings derived from preliminary runs of tree-based models. Features that were found to contribute minimally to the predictive power of the models were pruned from the dataset. This selective exclusion helps in reducing the complexity of the model, improving computational efficiency, and mitigating the risk of overfitting by eliminating noise and less relevant information from the analysis.

2.3 Model Development

The central endeavor of this study was the deployment and meticulous evaluation of a diverse spectrum of machine learning models, each selected to cater to the specificities of the Sleep Health and Lifestyle Dataset. This range included straightforward models like logistic regression and extended to more intricate ensemble methods and neural networks, creating a comprehensive framework for analysis. At the foundational level, logistic regression and decision trees were utilized as baseline models. These models, known for their simplicity and swiftness in training, are not only accessible but also crucial for establishing performance benchmark. This benchmark is essential for evaluating the efficacy and complexity of more advanced algorithms implemented later in the research. Logistic regression, with its direct probabilistic approach and inherent simplicity, and decision trees, with their intuitive decision-making process, provided the initial insights into the data's behavior and outlined the fundamental relationships within the variables.

To enhance the predictive power and address potential overfitting, a variety of ensemble methods were incorporated. These included Random Forest, Gradient Boosting, AdaBoost, and XGBoost. Each of these techniques leverages the concept of integrating multiple weak learners to form a robust strong learner. The ensemble methods, by design, improve model performance and generalizability as they combine the strengths of numerous simple models to enhance accuracy and control overfitting, making them well-suited for dealing with complex datasets with intricate patterns.

In scenarios demanding high performance in classification tasks, especially within high-dimensional spaces, Support Vector Machines (SVM) were employed. SVMs are particularly adept at creating optimal hyperplanes in an n-dimensional space, which maximizes the margin between different classes. For this dataset, both linear and radial basis function (RBF) kernels were used, facilitating the model's ability to handle linear separations and complex, non-linear boundaries between classes respectively. Neural networks, represented in this study by the Multi-Layer Perceptron (MLP), were used to model non-linear and complex relationships through their layered architecture and activation functions. The MLP is capable of learning deep representations and intricate patterns in the data, which might be missed by less sophisticated models. This ability makes neural networks an invaluable tool for capturing the nuanced interactions of features that affect sleep health.

Lastly, the study explored instance-based learning through the K-Nearest Neighbors (KNN) algorithm. KNN operates on the principle that similar data points can be found near one another in the feature space. By examining the labels of the nearest data points, KNN makes predictions, offering insights into the local structure of the data. This model's inclusion helped assess how proximity-based decision-making could be leveraged to enhance our understanding of sleep

patterns and disorders. This varied ensemble of models was not just a methodological choice but a strategic approach to dissect the dataset from multiple angles, ensuring a robust, comprehensive exploration of potential predictive relationships within the data. Each model contributed uniquely to the understanding of the dataset, providing a layered understanding of the factors influencing sleep health.

2.4 Model Training and Development

The dataset was split into a training set (80%) and a testing set (20%) using stratified sampling to maintain the proportion of categories in both sets. Each model was trained on the training set using a 5-fold cross-validation approach to ensure the models' generalizability. The hyperparameters for each model were optimized using GridSearchCV, which searches through a predefined grid of parameters and selects the combination that performs best based on cross-validated accuracy.

2.5 Evaluation Metrics

The evaluation of machine learning models involves several key performance metrics that are crucial for assessing the effectiveness and accuracy of the models in predicting outcomes. Below are the definitions and formulas for each of these metrics such as accuracy, precision, recall and F1 score. Accuracy measures the proportion of total correct predictions made by the model out of all predictions. It is defined mathematically as presented in equation 1.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions Made}} \quad (1)$$

Furthermore, precision is the ratio of correctly predicted positive observations to the total predicted positives. It is a measure of the exactness or quality of the classifier. The formula for precision is presented in equation 2.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

Recall, also known as sensitivity, measures the ratio of correctly predicted positive observations to all observations in the actual class - yes. The formula for recall is presented in equation 3.

$$(3)$$
$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

The F1 Score is the harmonic mean of Precision and Recall, and it is used when the balance between Precision and Recall is required. The F1 Score is especially useful in situations of uneven class distribution. The formula for calculating the F1 Score is presented in equation 4.

$$(4)$$
$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

These metrics provide a comprehensive framework for evaluating the performance of machine learning models, not only on the training dataset to monitor and adjust for overfitting but also on unseen test data to assess generalization capabilities.

2.6 Model Training and Development

The dataset was split into a training set (80%) and a testing set (20%) using stratified sampling to maintain the proportion of categories in both sets. Each model was trained on the training set using a 5-fold cross-validation approach to ensure the models' generalizability. The hyperparameters for each model were optimized using GridSearchCV, which searches through a predefined grid of parameters and selects the combination that performs best based on cross-validated accuracy.

2.7 Feature Importance Analysis and Software Tools

For tree-based models, feature importance was extracted, which indicates the relative importance of each feature in making accurate predictions. This analysis helps in understanding the contribution of different variables to sleep disorder predictions and provides insights that could influence future data collection and feature engineering efforts. In terms of Software and Tools, the analysis was performed using Python, leveraging libraries such as Pandas for data manipulation, Scikit-learn for machine learning tasks, and Plotly for visualization. This choice of software and tools ensures a robust analytical environment, suitable for complex model development and extensive data analysis.

3. RESULT AND DISCUSSION

The study aimed to evaluate and compare various machine learning models to predict sleep disorders using a comprehensive dataset from the Sleep Health and Lifestyle Study. The models tested include RandomForest, SVC (Support Vector Classifier), LogisticRegression, KNeighborsClassifier, GradientBoostingClassifier, DecisionTreeClassifier, AdaBoostClassifier, XGBClassifier, and MLPClassifier (Multi-Layer Perceptron). As presented in the table 1, the performance of these models was assessed using several key metrics: accuracy, precision, recall, and F1 score. The RandomForestClassifier and XGBClassifier exhibited the highest best scores (approximately 0.923), suggesting their strong capability in handling the diversity and complexity of the data. Similarly, these two models, along with the GradientBoostingClassifier, achieved the highest test accuracy and test F1 scores (around 0.931), highlighting their effectiveness in generalization on unseen data. The SVC and LogisticRegression models

demonstrated comparable performance, with test accuracies and F1 scores slightly above 0.900. These results are commendable, considering these models are generally less complex than ensemble methods and can be more interpretable.

The KNeighborsClassifier showed a slightly lower performance in comparison, with an accuracy of 0.875 and an F1 score of 0.874. This might suggest a sensitivity to the dataset's feature space dimensionality or the choice of 'k', indicating that instance-based learning may require careful tuning of parameters or might be less suitable for this particular dataset's characteristics. The DecisionTreeClassifier and AdaBoostClassifier also performed reasonably well, with accuracies just over 0.916 and 0.902, respectively. This performance indicates that while these models are capable, they might not capture the full complexity or the patterns in the data as effectively as the top-performing models.

Table 1. Machine Learning Comparison Results

model	best_score	test_accuracy	test_precision	test_recall	test_f1
RandomForestClassifier	0.92323	0.916667	0.92029	0.916667	0.915337
SVC	0.919722	0.902778	0.908688	0.902778	0.9
LogisticRegression	0.912825	0.902778	0.907851	0.902778	0.902345
KNeighborsClassifier	0.905747	0.875	0.884622	0.875	0.874237
GradientBoostingClassifier	0.919722	0.930556	0.936089	0.930556	0.930031
DecisionTreeClassifier	0.905808	0.916667	0.92029	0.916667	0.915337
AdaBoostClassifier	0.916213	0.902778	0.909005	0.902778	0.900439
XGBClassifier	0.92323	0.930556	0.936089	0.930556	0.930031
MLPClassifier	0.92323	0.902778	0.907851	0.902778	0.902345

The superior performance of ensemble methods such as RandomForest, GradientBoosting, and XGBClassifier can be attributed to their ability to construct more robust models by combining multiple weak learners. These methods are particularly effective in reducing variance and bias, making them excellent for complex datasets with many features, such as the Sleep Health and Lifestyle Dataset. The robustness of RandomForest and XGBClassifier is also evident in their precision and recall measures, which are crucial for medical diagnostic tasks where the cost of false negatives and false positives can be high. High precision indicates a low rate of false positives, which is vital for not alarming patients unnecessarily. Conversely, high recall ensures that the model identifies most patients with sleep disorders, which is critical for effective treatment. The lower performance of KNeighborsClassifier suggests that proximity-based methods may be limited by the sparsity of the data or the scale of certain features, despite normalization. This model's sensitivity to the dataset's intrinsic dimensionality and noise could account for its relatively lower accuracy and F1 score.

4. CONCLUSION

This study embarked on an exploratory journey to discern the most effective machine learning models for predicting sleep disorders using a dataset rich with demographic, lifestyle, and health-related variables. The endeavor not only aimed at identifying the best predictive models but also sought to underscore the nuanced interplay between various health indicators and sleep quality. Through rigorous evaluation, the research delineated clear distinctions in performance among a diverse array of models, from simple logistic regression to more sophisticated ensemble methods.

The findings from the comparative analysis robustly indicate that ensemble methods, particularly the RandomForest and XGBClassifier, excel in terms of accuracy, precision, recall, and F1 score. Achieving scores around 0.93 on these metrics, these models demonstrated superior capability to handle the dataset's complexity, making them standout choices for predicting sleep disorders. Their strength lies in their ability to amalgamate multiple weak learners into a coherent model that effectively reduces both bias and variance, hence their remarkable performance. Moreover, the study also highlighted the limitations and strengths of other models. While logistic regression and SVM provided reasonable accuracy and are valued for their interpretability, they fell short of the ensemble methods in performance. The KNeighborsClassifier, although useful in certain contexts, appeared less suited to this particular dataset, possibly due to its high dimensionality and the nature of the data distribution.

This research contributes to the ongoing efforts in the medical field to integrate advanced data analytics into clinical practices. The ability of machine learning to parse through complex datasets and extract meaningful predictions can significantly aid in the early diagnosis and treatment of sleep disorders, potentially improving outcomes for a wide demographic afflicted by such conditions. As the study concludes, it paves the way for future investigations. There is ample scope to enhance the predictive models further by incorporating more granular data, experimenting with hybrid models, and applying advanced feature selection techniques. Additionally, integrating these models into real-time monitoring systems in clinical settings could revolutionize the diagnosis and management of sleep disorders. The promising results of this study advocate for a deeper integration of machine learning techniques in healthcare, emphasizing their critical role in advancing diagnostic tools and improving patient care in the realm of sleep medicine.

5. REFERENCE

[1] R. Kobe, *Mental Health and Wellbeing*. Richmond Kobe, 2023.

[2] L. Chatre, *Serenity, Tranquility, Respire, Exercise, Sleep, Self-care*. Librinova, 2023.

[3] J. Sorensen, *The Art and Science of Lifestyle Medicine: A Practical Guide to Transforming Your Health*. eBookIt. com, 2024.

[4] X. Chen *et al.*, "Relationship between sleep duration and sociodemographic characteristics, mental health and chronic diseases in individuals aged from 18 to 85 years old in Guangdong province in China: a population-based cross-sectional study," *BMC Psychiatry*, vol. 20, pp. 1–10, 2020, doi: https://doi.org/10.1186/s12888-020-02866-9.

[5] R. Matsuda *et al.*, "Psychological disturbances and their association with sleep disturbances in patients admitted for cardiovascular diseases," *PLoS One*, vol. 16, no. 1, p. e0244484, 2021, doi: https://doi.org/10.1371/journal.pone.0244484.

[6] A. Riera-Sampol, M. Bennasar-Veny, P. Tauler, M. Nafr'ia, M. Colom, and A. Aguilo, "Association between depression, lifestyles, sleep quality, and sense of coherence in a population with cardiovascular risk," *Nutrients*, vol. 13, no. 2, p. 585, 2021, doi: https://doi.org/10.3390/nu13020585.

[7] V. K. Chattu *et al.*, "Precision medicine meets sleep medicine: how can sleep health aid to reduce the preventable burden of non-communicable diseases?," *Sleep Vigil.*, vol. 5, no. 2, pp. 179–188, 2021, doi: https://doi.org/10.1007/s41782-021-00149-1.

[8] R. De Fazio, V. Mattei, B. Al-Naami, M. De Vittorio, and P. Visconti, "Methodologies and wearable devices to monitor biophysical parameters related to sleep dysfunctions: an overview," *Micromachines*, vol. 13, no. 8, p. 1335, 2022, doi: https://doi.org/10.3390/mi13081335.

[9] A. A. Seixas *et al.*, "Benefits of community-based approaches in assessing and addressing sleep health and sleep-related cardiovascular disease risk: a precision and personalized population health approach," *Curr. Hypertens. Rep.*, vol. 22, pp. 1–18, 2020, doi: https://doi.org/10.1007/s11906-020-01051-3.

[10] N. F. Watson and C. R. Fernandez, "Artificial intelligence and sleep: Advancing sleep medicine," *Sleep Med. Rev.*, vol. 59, p. 101512, 2021, doi: https://doi.org/10.1016/j.smrv.2021.101512.

[11] J.-L. Pépin, S. Bailly, and R. Tamisier, "Big Data in sleep apnoea: Opportunities and challenges," *Respirology*, vol. 25, no. 5, pp. 486–494, 2020, doi: https://doi.org/10.1111/resp.13669.

[12] S. Xu *et al.*, "A review of automated sleep disorder detection," *Comput. Biol. Med.*, vol. 150, p. 106100, 2022, doi: https://doi.org/10.1016/j.compbiomed.2022.106100.

[13] B. C. Go and E. R. Thaler, "Home Sleep Testing versus Traditional Polysomnography: Pros and Cons," *Otolaryngol. Clin. North Am.*, 2023, doi: https://doi.org/10.1016/j.otc.2023.11.003.

[14] G. L. Weinhouse, E. Kimchi, P. Watson, and J. W. Devlin, "Sleep assessment in critically ill adults: established methods and emerging strategies," *Crit. Care Explor.*, vol. 4, no. 2, p. e0628, 2022, doi: https://doi.org/10.1097/CCE.0000000000000628.

[15] J. C. Suri, J. C. Suri, and T. M. Suri, "A Review of the Current Status of Home Sleep Apnea Testing vis-à-vis In-lab Polysomnography: Is Old Still Gold?," *Indian J. Sleep Med.*, vol. 17, no. 4, pp. 99–102, 2023, doi: https://doi.org/10.5005/jp-journals-10069-0106.

[16] M. Chromik, F. Lachner, and A. Butz, "Ml for ux?-an inventory and predictions on the use of machine learning techniques for UX research," in *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, 2020, pp. 1–11, doi: https://doi.org/10.1145/3419249.3420163.

[17] K. Setiawan and A. Wibowo, "Data Mining Implementation For Detection Of Anomalies In Network Traffic Packets Using Outlier Detection Approach," *JIKO (Jurnal Inform. dan Komputer)*, vol. 6, no. 2, 2023, doi: https://doi.org/10.33387/jiko.v6i2.6092.

[18] J. Egger *et al.*, "Medical deep learning—A systematic meta-review," *Comput. Methods Programs Biomed.*, vol. 221, p. 106874, 2022, doi: https://doi.org/10.1016/j.cmpb.2022.106874.

[19] V. Hovestadt, O. Ayrault, F. J. Swartling, G. W. Robinson, S. M. Pfister, and P. A. Northcott, "Medulloblastomics revisited: biological and clinical insights from thousands of patients," *Nat. Rev. Cancer*, vol. 20, no. 1, pp. 42–56, 2020, doi: https://doi.org/10.1038/s41568-019-0223-8.

[20] D. Bzdok and A. Meyer-Lindenberg, "Machine learning for precision psychiatry: opportunities and challenges," *Biol. Psychiatry Cogn. Neurosci. Neuroimaging*, vol. 3, no. 3, pp. 223–230, 2018, doi: https://doi.org/10.1016/j.bpsc.2017.11.007.

[21] A. M. Antoniadi *et al.*, "Current challenges and future opportunities for XAI in machine learning-based clinical decision support systems: a systematic review," *Appl. Sci.*, vol. 11, no. 11, p.

5088, 2021, doi: https://doi.org/10.3390/app11115088.

[22] A. Arora, P. Chakraborty, and M. P. S. Bhatia, "Intervention of wearables and smartphones in real time monitoring of sleep and behavioral health: an assessment using adaptive neuro-fuzzy technique," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1999–2024, 2022, doi: https://doi.org/10.1007/s13369-021-06078-5.

[23] B. Zhai, "Towards automated sleep stage assessment using ubiquitous computing technologies," Newcastle University, 2023.

[24] A. Bhatt *et al.*, "The Emergence of Inpatient Sleep Medicine: Screening for Sleep Disordered Breathing to Reduce the Burden of Cardiovascular Disease," *Curr. Sleep Med. Reports*, pp. 1–11, 2024, doi: https://doi.org/10.1007/s40675-024-00286-9.

[25] G. Thomas, J. A. Bennie, K. De Cocker, and S. J. H. Biddle, "Exploring contemporary screen time in Australian adolescents: A qualitative study," *Heal. Promot. J. Aust.*, vol. 32, pp. 238–247, 2021, doi: https://doi.org/10.1002/hpja.440.

[26] M. Sharma, H. Lodhi, R. Yadav, H. Elphick, and U. R. Acharya, "Computerized detection of cyclic alternating patterns of sleep: A new paradigm, future scope and challenges," *Comput. Methods Programs Biomed.*, p. 107471, 2023, doi: https://doi.org/10.1016/j.cmpb.2023.107471.

[27] J. Streatfeild, J. Smith, D. Mansfield, L. Pezzullo, and D. Hillman, "The social and economic cost of sleep disorders," *Sleep*, vol. 44, no. 11, p. zsab132, 2021, doi: https://doi.org/10.1093/sleep/zsab132.

[28] P. Huyett and N. Bhattacharyya, "Incremental health care utilization and expenditures for sleep disorders in the United States," *J. Clin. Sleep Med.*, vol. 17, no. 10, pp. 1981–1986, 2021, doi: https://doi.org/10.5664/jcsm.9392.

[29] S. Giménez, M. Altuna, E. Blessing, R. M. Osorio, and J. Fortea, "Sleep disorders in adults with Down syndrome," *J. Clin. Med.*, vol. 10, no. 14, p. 3012, 2021, doi: https://doi.org/10.3390/jcm10143012.

[30] U. of Michigan, "Sleep Health and Lifestyle Dataset." 2023.