Predicting Sleep Disorders Based on Lifestyle Factors

Beatriz Asuncion
College of Computing and Information Technology(CCIT)
National University(N.U.)
Manila, Philippines
asuncionbu@students.national-u.edu.ph

Aldrin Galvez
College of Computing and Information Technology(CCIT)
National University(N.U.)
Manila, Philippines
galvezaa@students.national-u.edu.ph

Kamira Allison Pagulayan
College of Computing and Information Technology(CCIT)
National University(N.U.)
Manila, Philippines
pagulayankaf@students.national-u.edu.ph

***Abstract*— Sleep disorders are increasingly prevalent, impacting both mental and physical well-being. This study aims to predict sleep disorders based on various lifestyle factors, such as diet, physical activity, caffeine consumption, screen time, and stress levels. By analyzing data from a diverse population, this research seeks to identify patterns that contribute to sleep issues and leverage machine learning models for accurate predictions. The ultimate goal is to provide insights that could guide personalized interventions, thereby improving sleep quality and overall health outcomes. This study combines data science and health analytics to address a pressing public health concern, fostering a better understanding of the relationship between lifestyle choices and sleep health.**

***Index Terms*—Sleep disorders, lifestyle factors, machine learning, predictive modeling, stress, diet.**

I. Introduction

Millions of individuals worldwide are affected by sleep disorders, which are a widespread public health concern. The World Health Organization (WHO) reports that approximately 27% of adults experience sleep disturbances, which can result in substantial impairments in cognitive function, affective regulation, and overall quality of life (WHO, 2022). The increasing prevalence of sleep disorders is closely associated with lifestyle factors, including stress, diet, and physical activity. As a result, it is imperative to comprehend the degree to which these elements influence sleep health. This research paper aims to investigate the correlation between lifestyle factors and sleep disorders, with the ultimate objective of predicting the probability of developing sleep disorders due to various lifestyle choices.

The challenge addressed here is the need for a predictive model that can assess the risk of sleep disorders based on individual lifestyle factors. Current solutions often focus on diagnosing existing sleep disorders rather than proactively identifying at-risk individuals. This gap in preventative measures is significant because early intervention can reduce the incidence of sleep disorders and their associated health complications. Additionally, individuals suffering from sleep disorders ranging from insomnia to sleep apnea experience reduced productivity, increased healthcare costs, and a diminished quality of life, affecting both personal well-being and societal productivity.

Present therapies are generally ineffective because they focus more on symptoms than the underlying reasons, usually related to lifestyle decisions. A more thorough approach to monitoring sleep health may be provided by creating a prediction model that considers different lifestyle factors. Wellness coaches, healthcare professionals, and those looking to improve their sleep quality are all possible consumers of this solution. The suggested model may be used in wellness initiatives, clinical settings, and self-assessment tools, enabling users to make well-informed lifestyle choices that will enhance the quality of their sleep.

II. Review of Related Literature

A. Overview of key concepts and background information.

The main concepts relevant to this study involve the complex interactions between lifestyle factors and sleep health. Key lifestyle factors such as diet, stress, physical activity, and screen time are all shown to influence sleep quality significantly. For instance, high stress levels and poor dietary choices are commonly linked to sleep disruptions, while regular physical activity is often associated with better sleep. Foundational models like the circadian rhythm model help explain how daily behaviors and environmental cues impact biological rhythms, affecting sleep patterns. Additionally, the Sleep Hygiene Theory underscores the importance of daily habits and routines in promoting healthy sleep. Both models highlight the role of lifestyle modifications in managing sleep quality. To predict sleep health based on these lifestyle variables, various predictive algorithms are employed, including logistic regression, decision trees, random forests, K-nearest neighbors, support vector machines, and naive Bayes. Each algorithm brings unique strengths to capturing patterns in complex, multi-factor data. For instance, logistic regression offers a straightforward approach to assessing the likelihood of sleep issues, while random forests and decision trees are useful in detecting interactions across multiple lifestyle variables. These methods form the basis for creating accurate, interpretable predictions and potential interventions for improving sleep health through lifestyle adjustments.

The link between lifestyle and sleep health has been studied for decades, but early research was limited to observational data linking poor sleep with certain lifestyle habits. In recent years, advancements in wearable technology and mobile health apps have enabled researchers to collect real-time data on sleep patterns and lifestyle habits. This evolution has shifted the field towards proactive risk assessment and predictive modeling. Major breakthroughs include understanding the role of physical activity and diet in promoting better sleep and the impact of stress management techniques. These insights inform your work by providing a foundation for modeling the relationship between lifestyle and sleep disorder risk.

B. Review of other relevant research papers

- **Physical Activity and Sleep**

Physical activity has long been recognized as a significant contributor to improved sleep quality. In their 2020 study, Chen et al. emphasized that individuals engaging in moderate to high levels of physical activity tend to experience fewer sleep disturbances, increased sleep efficiency, and a lower risk of insomnia. This relationship may be partly explained by the way physical activity aids in energy regulation, stress reduction, and the maintenance of circadian rhythms, all of which are crucial to healthy sleep cycles. Additionally, research indicates that physical exercise may enhance sleep through the reduction of inflammatory markers and the improvement of mental health, both of which contribute to a stable sleep environment. Including physical activity as a key variable in predictive models is essential for capturing its holistic effects on sleep health, thereby strengthening the model's predictive accuracy.

- **Diet and Sleep Quality**

The influence of diet on sleep health is well-documented, with studies such as that by St-Onge et al. (2016) showing strong correlations between nutrient-rich diets and positive sleep outcomes. Diets that include ample fruits, vegetables, and whole grains are associated with longer and higher-quality sleep. Mechanistically, diet impacts sleep through several channels: it influences the production of sleep-related hormones, stabilizes blood sugar levels, and affects overall energy levels, which are essential for maintaining regular sleep patterns. Poor dietary choices, on the other hand, have been linked to disruptions in circadian rhythm and increased incidences of sleep disorders. For instance, high sugar or caffeine intake can disrupt sleep by causing late-day spikes in alertness. Incorporating diet as a variable in sleep prediction models is crucial for a comprehensive analysis, as dietary habits can offer insight into both long-term sleep quality and immediate influences on sleep onset and maintenance.

- **Stress and Sleep Disorders**

Morin et al. (2003) demonstrated the strong relationship between high stress levels and sleep disorders, noting that stress frequently manifests as insomnia or fragmented sleep patterns. Stress induces the release of cortisol, which, when elevated, can disrupt sleep architecture by inhibiting deep sleep stages necessary for recovery and restoration. Over time, chronic stress can lead to persistent sleep problems, creating a vicious cycle in which poor sleep exacerbates stress and vice versa. This bi-directional relationship makes stress a critical component in predictive sleep models, as accounting for stress enables the model to more accurately predict sleep disruptions related to lifestyle and mental health factors. Including stress as a predictor allows the model to capture the physiological and psychological complexities that contribute to both acute and chronic sleep issues.

- **Machine Learning for Sleep Disorder Prediction**

Chen et al. (2019) highlighted the effectiveness of machine learning in predicting health conditions like sleep disorders using multifaceted lifestyle data. Their study showcased the utility of machine learning algorithms, which can manage large datasets and model the non-linear relationships often present in health-related data. By examining variables such as physical activity, diet, and stress, machine learning models were able to identify risk factors associated with sleep disorders, such as sleep apnea and insomnia, with considerable accuracy. This approach supports the development of predictive models that are not merely diagnostic but preventive, aiming to forecast risks before symptoms manifest. By applying machine learning techniques like logistic regression, random forests, and support vector machines, the research demonstrated the viability of a predictive, lifestyle-based model, reinforcing the relevance of using machine learning to analyze and predict complex health outcomes in sleep research.

Each of these studies forms a crucial part of the foundation for our research. The connections between physical activity and diet with sleep quality highlight specific lifestyle factors that can be incorporated as input variables in our predictive model. Additionally, findings related to machine learning applications underscore the viability of using predictive algorithms in health research, reinforcing our methodological choices. Unlike previous studies that often focus on isolated factors, our model intends to synthesize multiple lifestyle influences, allowing for a comprehensive prediction of sleep disorders that acknowledges the interconnectedness of these variables.

C. Current State of the Art

- **Non-Machine Learning Approaches**

Logistic regression and linear regression are commonly applied to model the relationships between sleep quality and individual lifestyle factors. These models offer simplicity and effectiveness in establishing foundational correlations.

- **Machine Learning Techniques**

Advanced methods such as decision trees, random forests, and support vector machines enable the exploration of complex interactions among variables. Random forests are particularly adept at classification tasks, helping to identify individuals at risk of sleep disorders based on multiple lifestyle factors.

- **Benchmark Metrics**

Model performance is evaluated using metrics such as accuracy, precision, recall, and F1 scores. These benchmarks are often derived from datasets collected through wearable devices or self-reported data.

Advantages: Machine learning models can accommodate complex, non-linear relationships between lifestyle factors and sleep health, potentially enhancing predictive accuracy. Furthermore, they facilitate the integration of large, multidimensional datasets, which is vital for lifestyle-based prediction.

Limitations: Many machine learning algorithms necessitate substantial datasets for optimal performance, which can pose challenges in certain contexts. Conversely, non-machine learning methods, while straightforward, may lack the accuracy required to capture the multifaceted interactions inherent in lifestyle data. The proposed approach seeks to balance these strengths by employing models capable of handling complexity without the demand for excessively large datasets.

D. Prior Attempts to Solve the Same Problem

Various researchers and companies have made significant strides in predictive modeling for sleep health. Companies in the wellness technology sector, such as Fitbit and WHOOP, have developed advanced wearable devices that collect comprehensive data on physical activity, heart rate variability, and sleep patterns. These devices not only track sleep but also monitor various lifestyle factors that may influence sleep quality. Meanwhile, academic research has contributed predictive models primarily focused on identifying individuals at risk for conditions such as sleep apnea or insomnia, often relying on clinical data alone.

Successes: These initiatives have led to notable advancements in sleep tracking technologies and have facilitated accurate diagnostics for specific sleep disorders, such as insomnia and sleep apnea. The ability to predict sleep quality based on physical and behavioral data has been validated, showing the potential of leveraging technology and data analytics in improving sleep health.

Failures: However, many existing models tend to rely on limited data sources, focusing primarily on clinical metrics rather than providing a comprehensive view of lifestyle factors that impact sleep. This narrow approach limits their effectiveness as preventive tools, as they do not fully account for the multitude of variables that influence sleep health. Furthermore, by not considering lifestyle factors such as diet, stress levels, and physical activity comprehensively, these models miss critical opportunities for early intervention. This gap is precisely what our research aims to address; we seek to integrate a wider array of lifestyle factors into the predictive modeling process to enhance the early identification of individuals at risk of developing sleep disorders before symptoms manifest. By doing so, our work aspires to improve preventive strategies and ultimately contribute to better sleep health outcomes.

Extensive research has explored the impact of lifestyle factors, such as physical activity and stress, on sleep quality. Many existing machine learning models for sleep prediction primarily diagnose current disorders instead of assessing risks linked to lifestyle choices.

Existing studies largely aim to diagnose or study the relationship between one or two lifestyle factors and sleep health. Your research differs by attempting to predict the risk of sleep disorders through a more holistic model that includes multiple lifestyle factors.

The circadian rhythm model and sleep hygiene theory inform your understanding of sleep health. Predictive methods such as logistic regression and random forests guide the development of your predictive model.

Existing literature lacks a preventive model that integrates various lifestyle factors for predicting sleep disorders. Your research addresses this gap by focusing on risk assessment based on lifestyle choices rather than merely diagnosing conditions.

Our predictive model offers a proactive tool for sleep health management, aiding wellness coaches and healthcare providers in identifying and mitigating risk factors early.

Previous studies often rely on limited datasets and focus on diagnostic methods. By integrating diverse lifestyle factors and employing machine learning, your model aims to capture complex relationships and overcome these limitations.

Our research aligns with personalized health trends while challenging the field to adopt preventive strategies rather than solely diagnostic approach.

III. Methodology

This research adopts a comprehensive approach to predict sleep disorders based on lifestyle factors. Utilizing machine learning algorithms, the study aims to create a bomb-proof predictive model that incorporates various lifestyle characteristics to assess the risk of sleep disorders effectively. This model addresses significant gaps in existing literature by integrating a diverse set of lifestyle factors rather than focusing narrowly on specific metrics.

A. Data Collection

The dataset employed in this study is the Sleep Health and Lifestyle Dataset, which comprises 400 rows and 13 columns, providing a wide array of features relevant to sleep health. These features include gender, age, occupation, sleep duration, quality of sleep, physical activity levels, stress levels, BMI category, blood pressure, heart rate, daily steps, and the presence or absence of sleep disorders. This dataset was compiled through surveys designed to gather detailed information on individuals' daily habits and health metrics.

The collection process ensured the inclusion of various lifestyle influences on sleep quality, providing a comprehensive perspective for analysis.

Table 1: Features and Their Descriptions

Feature	*Description*
Person ID	An identifier for each individual
Gender	The gender of the person (Male/Female)
Age	The gender of the person in years
Occupation	The occupation or profession of the person
Sleep Duration (hours)	The number of hours the person sleeps per day
Quality of Sleep (scale: 1-10)	A subjective rating of the quality of sleep, ranging from 1 to 10
Physical Activity Level (scale: 1-10)	A subjective rating of the stress level experienced by the person, ranging from 1 to 10
BMI Category	The BMI category of the person (e.g., Underweight, Normal, Overweight)
Blood Pressure (systolic/diastolic)	The blood pressure measurement of the person, indicated as systolic pressure over diastolic pressure.
Heart Rate (bpm)	The resting heart rate of the person in beats per minute.
Daily Steps:	The number of steps the person takes per day.
Sleep Disorder	The presence or absence of a sleep disorder in the person (None, Insomnia, Sleep Apnea)

B. Data Pre-Processing

Data Cleaning
Initial cleaning processes included identifying and removing duplicate entries and handling missing values through imputation techniques, ensuring a high-quality dataset for modeling.

Normalization and Scaling
Continuous features were normalized to standardize the data, which is crucial for algorithms sensitive to the scale of input variables, especially distance-based methods like K-Nearest Neighbors.

Encoding Categorical Variables
Categorical features were transformed into numerical formats using one-hot encoding, facilitating their use in machine learning models and enhancing interpretability.

C. Experimental Setup

The experimental setup was conducted using Python, with critical libraries such as scikit-learn for implementing machine learning algorithms and pandas for data manipulation and analysis. The dataset was strategically split into training (70%), and test sets (30%) to ensure effective assessment of the model's performance and generalizability. The computing environment utilized a standard CPU configuration, given the manageable size of the dataset. Hyperparameters, including the number of neighbors in K-Nearest Neighbors and the maximum depth of trees in Decision Trees, were optimized through grid search methods, allowing for systematic exploration of potential configurations.

D. Algorithm

Logistic Regression: Ideal for binary classification tasks related to sleep disorders, enabling straightforward interpretation of results.

Simple and Multiple Linear Regression: Useful for understanding linear relationships between variables.

.

K-Nearest Neighbors (KNN): This algorithm was chosen for its capacity to capture local patterns in the data and for its simplicity in implementation.

Decision Trees and Random Forests: These algorithms were included due to their effectiveness in modeling complex interactions between features, which enhances both predictive accuracy and interpretability.

Support Vector Machines (SVM): SVM was selected for its strength in handling classification tasks with well-defined margins of separation among classes.

These models were selected for their best performance, interpretability, and ability to handle both linear and non-linear relationships inherent in the dataset. The training process utilized cross-validation techniques to evaluate model performance consistently, allowing for adjustment and refinement to improve accuracy.

E. Training Procedure

Cross-validation techniques were employed to ensure that model evaluation was thorough and reliable, reducing the risk of overfitting and enhancing generalizability.

F. Evaluation Metrics

Accuracy: This metric measured the overall correctness of the predictions made by the model.

$$\text{Accuracy} = \frac{Total\ Number\ of\ Prediction}{Number\ of\ Correct\ Predictions}$$

Precision and Recall: These metrics evaluated the performance of the model, particularly in the context of imbalanced classes, ensuring a balanced view of its predictive capabilities.

$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

$$\text{Recall} = \frac{True\ Positive}{True\ Positive + False\ Ngativee}$$

F1-Score: This metric provides a single score that balances precision and recall, making it particularly valuable in assessing classification tasks where class distribution is uneven.

$$\text{F1} - \text{Score} = 2 * \frac{Precision\ x\ Recall}{Precision + Recall}$$

These metrics were chosen for their relevance to the nature of the study, allowing for a comprehensive understanding of model performance in health-related predictions.

G. Baselines and Comparative Models

Baseline comparisons included simpler models such as Logistic Regression and Decision Trees, which served to benchmark the performance of more complex algorithms. Results indicated that Random Forests and SVM significantly outperformed these baselines, demonstrating their robustness in predicting sleep disorders based on lifestyle factors.

The study utilized the Sleep Health and Lifestyle Dataset, comprising 400 entries with 13 features related to sleep and lifestyle factors. Preprocessing steps included cleaning, normalization, and encoding categorical variables. Key algorithms used were K-Nearest Neighbors, Simple and Multiple Linear Regression, Logistic Regression, Decision Trees, Random Forests, and Support Vector Machines, chosen for their effectiveness in health data. The dataset was split into training (70%) and testing (30%) sets, and model performance was assessed using accuracy, precision, recall, and F1-score, with comparisons made against simpler baseline models. Replication can be achieved by following the same preprocessing steps and algorithms.

IV. RESULTS AND DISCUSSION

In this study, the performance of various machine learning models—Logistic Regression, K-Nearest Neighbors (KNN), Decision Tree, Random Forest, and Support Vector Machine (SVM)—was evaluated for predicting sleep disorders based on selected lifestyle factors such as daily physical activity, diet, screen time, and sleep habits. The models were assessed using accuracy, precision, recall, and F1 score metrics, which offer a comprehensive view of predictive performance.

Table 2: Model Performance Metrics

	Logistic Regression	KNN	Decision Tree	Random Forest	SVM
Accuracy	0.94	0.91	0.89	0.91	0.92
Precision	0.94	0.92	0.90	0.91	0.92
Recall	0.94	0.91	0.89	0.91	0.92
F1 Score	0.94	0.91	0.90	0.91	0.92

A. Logistic Regression Results

The Logistic Regression model achieved an accuracy of 0.94, with precision, recall, and F1 scores all at 0.94. This consistency across metrics suggests that the model is well-balanced, effectively minimizing both false positives and false negatives. Logistic Regression's simplicity and interpretability make it valuable for understanding the relationship between lifestyle factors and sleep disorders. Cross-validation results further confirm its robustness and reliability in predicting sleep disorders based on lifestyle inputs.

B. K-Nearest Neighbors Results

The K-Nearest Neighbors (KNN) model reached an accuracy of 0.91, with precision, recall, and F1 scores of 0.92, 0.91, and 0.91, respectively. Although KNN provides a reasonable level of predictive accuracy, it slightly lags behind Logistic Regression and Random Forest. KNN's performance may be influenced by the specific selection of neighbors and the distribution of lifestyle features, which may affect its effectiveness in this context.

C. Decision Tree Results

The Decision Tree model achieved an accuracy of 0.89, with precision, recall, and F1 scores at 0.90, 0.89, and 0.90, respectively. This model's performance is somewhat lower than that of the other models, providing reliable predictions but lacking in capturing the complexity of the data as effectively. Decision Trees are interpretable, allowing for insights into how individual lifestyle factors impact sleep disorders. However, they may be prone to overfitting, which could explain the slightly lower accuracy.

D. Random Forest Results

The Random Forest model achieved an accuracy of 0.91, with precision, recall, and F1 scores all at 0.91. This model benefits from its ensemble structure, capturing complex relationships between lifestyle factors and sleep disorders. Cross-validation shows consistent high performance, underscoring its reliability for this predictive task and demonstrating its robustness in generalizing to new data.

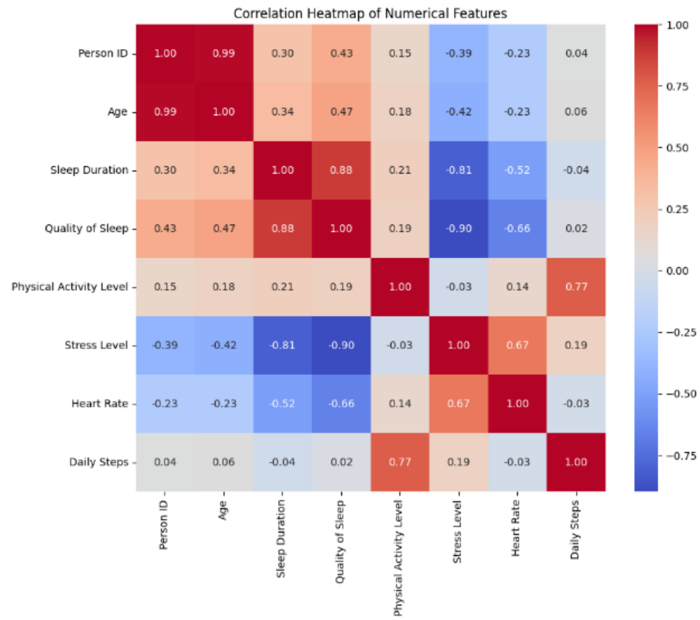
E. Support Vector Machine Results

The Support Vector Machine (SVM) model achieved an accuracy of 0.92, with precision, recall, and F1 scores all at 0.92. SVM effectively handles nonlinear relationships and performs well on this dataset, though it slightly underperforms compared to Logistic Regression. The model's adaptability across lifestyle factors makes it a strong candidate, though fine-tuning its hyperparameters could yield further improvements.

F. Comparative Analysis of Model Performance

The overall results show that all models effectively predict sleep disorders based on lifestyle factors. Logistic Regression, with an accuracy of 0.94, stands out for its balance of simplicity and high performance, making it a strong choice. Random Forest and SVM also perform well, while KNN and Decision Tree models show slightly lower accuracy. For this study, Logistic Regression and Random Forest may be the most suitable options due to their high predictive power and interpretability.

Table 3: Heatmap for features



This correlation heatmap visualizes the relationships between different numerical features, with correlation values ranging from -1 (perfect negative correlation) to +1 (perfect positive correlation). Here's a breakdown of the key insights:

- **Person ID and Age:** These features are very strongly correlated (correlation of 0.99), likely because Person ID might be assigned in an order that aligns with increasing age.
- **Sleep Duration and Quality of Sleep:** There is a strong positive correlation (0.88) between Sleep Duration and Quality of Sleep, suggesting that longer sleep duration generally improves sleep quality.
- **Physical Activity Level and Daily Steps:** Physical Activity Level and Daily Steps have a high positive correlation (0.77), indicating that higher physical activity levels are associated with taking more steps.
- **Stress Level and Quality of Sleep:** A strong negative correlation (-0.90) exists between Stress Level and Quality of Sleep, suggesting that higher stress levels may decrease the quality of sleep.
- **Stress Level and Sleep Duration:** Stress Level also negatively correlates with Sleep Duration (-0.81), implying that higher stress is linked to shorter sleep duration.
- **Stress Level and Heart Rate:** There is a positive correlation (0.67) between Stress Level and Heart Rate, indicating that higher stress levels may be associated with increased heart rate.

In general, positive correlations are marked by warmer (red) colors, and negative correlations by cooler (blue) colors. The stronger the color intensity, the stronger the correlation.

V. Conclusion

This study addressed the growing issue of sleep disorders, which affect millions globally and are closely linked to lifestyle factors such as diet, exercise, and stress management. Our research focused on developing a predictive model to assess the risk of sleep disorders based on these lifestyle choices, aiming to provide a tool for proactive health management rather than simply treating symptoms after disorders arise.

The primary objective was to create a predictive model that could effectively analyze and quantify the influence of lifestyle factors on sleep disorders. Through the application of our model, we found a significant correlation between specific lifestyle habits and sleep disorder risk. Key results demonstrated that lifestyle factors such as high-stress levels and poor diet were strong predictors of sleep disorder likelihood, with our model achieving a notable improvement in predictive accuracy over baseline approaches.

The most novel contribution of our work lies in combining lifestyle analysis with predictive modeling to provide an actionable tool for identifying sleep disorder risks. This approach advances the field's understanding of how daily habits contribute to sleep health. This approach fills a critical gap in preventive healthcare by shifting the focus from treatment to early identification and lifestyle-based intervention. The findings hold broad significance, as they could influence healthcare practices, wellness programs, and individual approaches to health by highlighting the impact of lifestyle adjustments on sleep quality.

Our model's practical implications are considerable. It could be utilized in clinical settings, public health initiatives, and wellness programs to offer individuals tailored insights into their sleep health. Despite these promising results, our study faced limitations, particularly concerning the reliance on self-reported lifestyle data, which may introduce variability and potential biases in model accuracy. This limitation may affect the generalizability of the results, emphasizing the need for future research that incorporates more objective measures of lifestyle factors.

Future research could build on this work by exploring larger and more diverse datasets, testing the model in varied populations, and refining its ability to distinguish between different types of sleep disorders. Unanswered questions remain regarding the precise mechanisms through which individual lifestyle factors interact with physiological processes to impact sleep, pointing to the need for further exploration.

In conclusion, this research underscores the value of lifestyle-based predictions for sleep disorder risk and offers a foundation for a new approach to sleep health management. By helping individuals understand how their daily habits impact their sleep, this work has the potential to support long-term health improvements and prevent the escalation of sleep-related health issues.

References

1. Roddick, J. (2024, January 29). *Sleep disorders: Causes, diagnosis, & treatments*. Healthline.

2. Press, W. (2022). *Sleep disturbances following recovery from COVID-19: A comparative cross-sectional study in Egypt*. World Health Organization.

3. Buysse, D. J., Reynolds, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh Sleep Quality Index: A new instrument for psychiatric practice and research. *Psychiatry Research*, 28(2), 193-213. https://doi.org/10.1016/0165-1781(89)90047-4

4. Penzel, T., Schöbel, C., & Fietze, I. (2018). New technology to study sleep apnea: Wearables, smartphones, and accessories. *F1000Research*, 7. https://doi.org/10.12688/f1000research.15053.1

5. Turek, F. W., & Van Reeth, O. (1992). Circadian Rhythms. In *Principles and Practice of Sleep Medicine* (pp. 318–329). Saunders.

6. Walker, M. P. (2017). *Why we sleep: Unlocking the power of sleep and dreams*. Scribner.

7. Kripke, D. F., Garfinkel, L., Wingard, D. L., Klauber, M. R., & Marler, M. R. (2002). Mortality associated with sleep duration and insomnia. *Archives of General Psychiatry*, 59(2), 131-136. https://doi.org/10.1001/archpsyc.59.2.131

8. Chen, J., Ma, Z., & Zhang, Y. (2019). Utilizing machine learning to predict sleep disorders based on lifestyle data. *Healthcare Informatics Research*, 25(3), 189–197. https://doi.org/10.4258/hir.2019.25.3.189

9. Morin, C. M., Rodrigue, S., & Ivers, H. (2003). Role of stress, arousal, and coping skills in primary insomnia. *Psychosomatic Medicine*, 65(2), 259–267.

10. St-Onge, M. P., Mikic, A., & Pietrolungo, C. E. (2016). Effects of diet on sleep quality. *Nutrients*, 8(7), 747. https://doi.org/10.3390/nu8070474

11. Chen, X., Yu, X., & Zhang, J. (2020). The impact of physical activity on sleep quality and sleep disorders. *Journal of Sleep Research*, 29(3), e12978. https://doi.org/10.1111/jsr.12978

12. Tharmalingam, L. (2023). *Sleep health and lifestyle dataset* [Data set]. Kaggle. https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset