ABSTRACT

Sleep is a vital component of overall health, yet sleep disorders have become increasingly common due to modern stress, lifestyle habits, and irregular routines. This project aims to build an intelligent two-stage classification system to detect and categorize sleep disorders using machine learning and deep learning techniques. In the first stage, the model identifies whether an individual has a sleep disorder or not, while the second stage determines the specific disorder type such as Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, or Narcolepsy. This multi-layered approach provides a more accurate and interpretable analysis, making it suitable for real-world applications in health monitoring and digital wellness systems.

The dataset contained records combining lifestyle, physiological, and behavioral features related to sleep health. The first-stage classification revealed that 7,383 individuals showed no signs of a disorder, while 2,617 were diagnosed with at least one sleep disorder. Among these, Sleep Apnea emerged as the most frequent condition (865 cases), followed by Insomnia (626 cases), Restless Leg Syndrome (496 cases), Parasomnias (436 cases), and Narcolepsy (194 cases). Such variation across disorders emphasizes the importance of data-driven healthcare in identifying diverse sleep related issues and supporting early detection.

Traditional machine learning models such as Random Forest, XGBoost, and Gradient Boosting achieved high accuracies, with XGBoost of 98.8% and the ML_Hybrid_RF_XGB model of 99% outperforming others. Deep learning models like MLP of 95% and CNN of 92.7% effectively captured complex, non-linear relationships in the data. Hybrid deep learning methods such as CNN-DNDF showed the potential to combine the strengths of both convolutional learning and decision-based reasoning. The integration of ensemble and neural network models improved the robustness, interpretability, and generalization of predictions across both stages. The two-stage predictive framework demonstrates the capability of hybrid deep learning systems to accurately identify and classify sleep disorders.

Keywords:

Sleep Disorder Prediction, Machine Learning, Deep Learning, Two-Stage Classification, Sleep Apnea , Insomnia , Restless Leg Syndrome , Parasomnias and Narcolepsy

CHAPTER I

INTRODUCTION

Sleep disorders are problems that disturb a person's normal sleep, making it hard to get proper rest. People with sleep disorders often feel tired during the day, have trouble focusing, and may feel irritated or low in energy. There are many types of sleep disorders, such as insomnia, sleep apnea, restless leg syndrome, parasomnias, and narcolepsy. Insomnia makes it difficult to fall asleep or stay asleep, while sleep apnea causes short pauses in breathing during sleep. Restless leg syndrome gives an uncomfortable feeling in the legs, and parasomnias involve unusual actions like sleepwalking. Narcolepsy causes people to suddenly fall asleep during the day. These problems can affect both physical and mental health, so understanding and treating them early is very important for overall well-being.

This project focuses on creating a smart system that can detect and classify sleep disorders accurately. It uses important personal, lifestyle, and sleep-related information to make predictions. The system looks at features like age, gender, height, weight, caffeine intake, smoking and alcohol habits, screen time, diet quality, and daily sleep patterns. It also considers key sleep details such as how long a person sleeps, the time it takes to fall asleep, the number of times they wake up at night, sleep quality, daytime sleepiness, the urge to move legs before sleep, daytime sleep attacks, irregular work schedules, and unusual sleep behaviors. By analyzing all these factors together, the system can provide a personalized view of a person's sleep health and identify potential problems early. This makes it possible to give helpful recommendations, monitor health more effectively, and take timely steps to improve sleep and overall well-being.

Understanding Sleep Health Across Age Categories

Sleep disorders can affect people of all ages, and their impact often depends on a person's stage of life. In children and teenagers, poor sleep can interfere with learning, memory, and emotional growth. It may lead to difficulty concentrating in school, mood swings, or behavioral problems. Early identification of sleep issues in young people is important to support healthy development. By predicting potential disorders, parents and caregivers can take timely action to improve bedtime routines, sleep habits, and overall health. Preventing long term sleep problems at a young age can reduce the risk of future physical and mental health complications. Healthy sleep in childhood sets the foundation for better energy levels and cognitive function. Early intervention can also improve academic performance and social well being. In adults, sleep problems can reduce work productivity, increase stress, and raise the risk of conditions like diabetes, obesity, and heart disease.

Sleep apnea, insomnia, and restless leg syndrome are common in this age group and often go unnoticed. Predicting sleep disorders allows adults to make lifestyle changes, seek medical advice, or adopt therapies to improve sleep quality. For older adults, sleep disturbances are more frequent due to age-related changes, medications, or chronic illnesses. Poor sleep in elderly individuals can lead to memory loss, weakened immunity, and a higher risk of falls. Early detection helps implement solutions such as sleep hygiene improvements, medical treatments, or daily routine adjustments. Monitoring sleep across all age groups ensures better overall health, energy, and quality of life. Proactive sleep disorder prediction can prevent complications and support long-term well-being.

Challenges of Sleep Disorders

Sleep disorders pose several challenges for both individuals and healthcare systems. One major challenge is that many people are unaware they have a problem, as symptoms like daytime fatigue, mood changes, or concentration difficulties are often ignored. This delayed recognition can prevent timely treatment and worsen overall health. Diagnosing sleep disorders can also be complex because they have diverse causes, ranging from lifestyle habits and stress to underlying medical conditions. Traditional diagnostic methods, like polysomnography, are accurate but costly, time-consuming, and require specialized equipment and trained professionals.

Another challenge is managing sleep disorders effectively. Treatment often requires personalized approaches, including lifestyle changes, therapy, or medications, and results can vary widely between individuals. Adherence to treatment plans can be difficult, especially when symptoms are mild or intermittent. Additionally, coexisting health problems, such as diabetes, heart disease, or mental health issues, can complicate both diagnosis and treatment. The societal impact of sleep disorders, including reduced productivity, increased accidents, and healthcare costs, also highlights the importance of early detection and intervention.

Sleep Disorder Labels

A two-stage labeling approach is used to classify sleep disorders more accurately. The first stage focuses on a simple question: does the individual have a sleep disorder or not? This helps separate healthy individuals from those who may require further evaluation. By dividing the data into “No Disorder” and “Has Sleep Disorder,” the model can focus on identifying general patterns associated with sleep problems before trying to classify specific types. This stage ensures that the system avoids misclassifying healthy individuals as having a disorder.

The second stage goes deeper by identifying the specific type of sleep disorder for those flagged in the first stage. Here, individuals labeled as “Has Sleep Disorder” are further categorized into types such as Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, or Narcolepsy. This stage allows the model to analyze detailed patterns in lifestyle, physiological, and behavioral data that distinguish one disorder from another. Accurate classification in this stage is crucial for personalized recommendations and targeted interventions.

The two-stage labeling system improves the overall prediction accuracy and interpretability of the model. By first filtering out healthy individuals, the second stage can concentrate on subtle differences between disorders, reducing errors and increasing confidence in predictions. This hierarchical approach also mirrors real-world clinical decision-making, where doctors first determine if a problem exists and then identify its type. Overall, it provides a structured, reliable framework for detecting and classifying sleep disorders effectively.

Approaches to Improve Sleep Disorder Prediction

To improving sleep disorder prediction is to collect high-quality and diverse data. Including a wide range of personal, lifestyle, and sleep-related features such as age, gender, diet, caffeine intake, sleep duration, and night awakenings helps the model learn more accurate patterns. Consistent and reliable data from various age groups and health backgrounds ensures that predictions are robust and applicable to real-world scenarios. Additionally, handling missing values, outliers, and imbalanced datasets can significantly improve the model’s performance. Better data quality directly contributes to higher accuracy and more reliable detection of sleep disorders.

To use advanced machine learning and deep learning techniques. Combining traditional algorithms like Random Forest, XGBoost, and Gradient Boosting with deep learning models such as MLP and CNN allows the system to capture both simple and complex patterns in the data. Hybrid models like CNN-DNDF can merge neural network learning with decision-based reasoning, enhancing prediction accuracy. Feature selection and dimensionality reduction techniques can also help the models focus on the most important factors, reducing noise and improving interpretability. Continuous experimentation and tuning of model parameters are crucial for achieving optimal results.

To implement a two-stage or hierarchical prediction system. The first stage filters out healthy individuals, and the second stage focuses on classifying the specific type of sleep disorder. This reduces errors and increases confidence in predictions. Integrating personalized recommendations based on the predicted disorder can improve user engagement and health outcomes. Furthermore, periodic retraining of models with new data ensures that predictions remain up-to-date with changing lifestyle patterns and emerging sleep disorder trends. Combining these strategies leads to a more reliable, interpretable, and practical sleep disorder prediction system.

Implications of Results

The findings from this project show that the two-stage prediction system, using models like Random Forest, XGBoost, Gradient Boosting, MLP, CNN, and the hybrid CNN-DNDF, can accurately detect whether a person has a sleep disorder and identify its specific type. This means individuals can receive early warnings about potential sleep problems and take steps to improve their sleep, such as adopting better sleep routines, adjusting lifestyle habits, or consulting healthcare professionals. Early detection can prevent minor sleep issues from developing into serious health problems, improving overall well-being.

For healthcare providers, these results suggest that predictive models such as XGBoost and CNN-DNDF can support faster and more accurate diagnosis. Doctors can focus on patients at higher risk, saving time and resources compared to traditional sleep tests. The combination of machine learning and deep learning models enhances reliability and interpretability, allowing for better treatment plans and improved patient outcomes.

On a larger scale, the study highlights how data-driven systems can benefit public health and digital wellness initiatives. By analyzing lifestyle, behavioral, and physiological factors with models like Random Forest, Gradient Boosting, and MLP, health platforms can offer personalized recommendations and preventive strategies. Overall, these results demonstrate the potential of using advanced predictive systems to monitor sleep, improve quality of life, and reduce the impact of sleep-related disorders in society.

CHAPTER II

LITERATURE REVIEW

TITLE: “Applying Machine Learning Algorithms for the Classification of Sleep Disorders”**[1]**
AUTHORS: Talal Sarheed Alshammar

This paper explored the use of machine learning techniques to classify different types of sleep disorders, aiming to evaluate how effectively various algorithms could detect sleep problems using a structured dataset of 400 records with 13 features related to sleep patterns, lifestyle habits, and physiological measurements. The study tested several algorithms, including k-Nearest Neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), and Artificial Neural Networks (ANN), and found that the ANN model achieved the highest accuracy of 92.92%, along with strong precision, recall, and F1-score, demonstrating its ability to correctly identify true cases while minimizing false predictions.

To further improve performance, a genetic algorithm was applied for parameter optimization, fine-tuning model settings and ensuring robustness and reliability. The research highlighted that machine learning models can capture complex, non-linear relationships between features such as sleep duration, number of night awakenings, and lifestyle factors—which traditional statistical methods might overlook. By selecting relevant features and optimizing models, the study achieved high prediction accuracy, showing that data-driven approaches can support early detection of sleep disorders and help clinicians identify high-risk individuals more efficiently, potentially reducing the need for time-consuming and costly traditional diagnostic tests.

The paper also emphasized the potential of artificial intelligence and machine learning in sleep healthcare, illustrating that predictive models can not only aid in diagnosis but also provide personalized interventions, health monitoring, and preventive strategies. Overall, the findings demonstrate that integrating machine learning into sleep disorder management offers valuable insights, enhances early detection, and contributes to improved well-being. Incorporating real-time monitoring data from wearable devices may further enhance predictive accuracy and provide continuous health insights. Finally, the combination of machine learning with personalized feedback systems could empower individuals to make lifestyle adjustments that improve sleep quality and overall health outcomes.

TITLE: "Dynamic Impact of the Sleep Disorder, Depression and Anxiety on the Cognitive Function in the First-Episode Depressive Patients" **[2]**

AUTHORS: W. Zhang, N. Zhou, & J. Li

Zhang et al. conducted a study to examine how sleep disorders, depression, and anxiety dynamically affect cognitive function in patients experiencing their first episode of major depressive disorder (MDD). The study included 173 patients who were followed over an 8-week period, during which data on sleep quality, depressive and anxiety symptoms, and cognitive performance were collected. Latent Growth Curve Models (LGCM) were used to analyze the relationships between these variables and track changes over time.

The results indicated that baseline sleep disturbances significantly predicted both the level of cognitive function ($p = 0.043$) and its trajectory over the 8-week period ($p = 0.016$). Furthermore, improvements in depressive and anxiety symptoms were positively associated with enhancements in cognitive function. This suggests that managing sleep problems alongside mood symptoms can directly contribute to cognitive recovery in MDD patients. The study emphasizes the importance of early intervention to address sleep disorders and psychological distress in patients with major depression. By monitoring and improving sleep quality and mood symptoms, healthcare providers can help prevent cognitive decline and support overall mental health. These findings highlight the critical role of integrated care approaches that consider both physiological and psychological factors in the treatment of first-episode depressive patients.

Additionally, the research underscores the potential benefits of personalized treatment plans that simultaneously target sleep, mood, and cognitive function. Incorporating strategies such as cognitive behavioral therapy for insomnia, mindfulness-based interventions, or structured sleep hygiene programs could further enhance cognitive outcomes. The study also suggests that regular assessment of sleep quality and psychological symptoms should become a standard part of care for patients with MDD, as it allows for timely adjustments to treatment. Overall, these findings provide strong evidence that addressing sleep disorders is not only essential for mental health recovery but also plays a key role in improving long-term cognitive functioning in depressive patients.

TITLE: "Advanced Sleep Disorder Detection Using Multi-Layered Ensemble Learning and Advanced Data Balancing Techniques" **[3]**

AUTHORS: M. M. Monowar, S. M. N. Nobel, M. Afroj, M. A. Hamid, M. Z. Uddin, M. M. Kabir, & M. F. Mridha

Monowar et al. proposed a multi-layered ensemble learning approach to enhance the detection of sleep disorders, addressing challenges such as data imbalance and low interpretability in traditional models. The study used a combination of machine learning models and applied techniques like thresholding, predictive scoring, and converting Softmax labels into multidimensional feature vectors. These strategies improved both the understanding of model decisions and the reliability of predictions.

The ensemble methods, including voting and stacking, allowed multiple models to collaborate in making final predictions, ensuring higher accuracy and robustness compared to individual models. Additionally, the study addressed the problem of class imbalance using SMOTE, which generates synthetic samples for underrepresented classes. This ensured that the model could learn patterns effectively across all categories of sleep disorders, preventing bias toward majority classes. The proposed method achieved outstanding results, with an accuracy of 96.88% on the SMOTE-balanced dataset and 95.75% on the original dataset. Eight-fold cross-validation demonstrated the model's stability, reaching up to 99.5% accuracy, highlighting its robustness and generalizability. These results confirm that combining ensemble strategies with data balancing significantly improves the performance of sleep disorder detection systems.

In addition to accuracy, the model showed improved precision, recall, and F1-scores, indicating reliable identification of true positive cases and fewer false predictions. By converting Softmax outputs into multidimensional feature vectors, the study also enhanced interpretability, which is crucial for clinical adoption where understanding the reasoning behind predictions is necessary. The study demonstrates that combining multiple machine learning models with thoughtful data handling and interpretability strategies can significantly enhance sleep disorder detection. This approach provides a reliable, accurate, and clinically relevant tool for early identification and management of sleep disorders, potentially improving patient outcomes and supporting digital health initiatives.

TITLE: "Improving Sleep Disorder Diagnosis Through Optimized Machine Learning Approaches" **[4]**

AUTHORS: M. A. Rahman, I. Jahan, M. Islam, T. Jabid, M. S. Ali, M. R. A. Rashid.

Rahman et al. proposed an optimized machine learning framework aimed at enhancing the diagnosis of sleep disorders such as insomnia and sleep apnea using advanced classification algorithms. The study utilized the Sleep Health and Lifestyle Dataset, which included a variety of demographic, behavioral, and physiological features influencing sleep quality. The researchers trained 15 machine learning models, including Gradient Boosting, CatBoost, Voting, and Stacking classifiers, to evaluate their performance in accurately identifying sleep disorders.

Feature selection was performed using the Mean Decrease Impurity (MDI) method, which helped determine the most significant predictors affecting sleep health. This step improved model efficiency by eliminating redundant features and enhancing interpretability. The optimized models achieved impressive results, with an overall accuracy of 97.33%, and precision, recall, and F1-score values all reaching 0.9733, while specificity stood at 0.9569. These metrics indicate that the models were both highly accurate and consistent in distinguishing between healthy and disordered sleep patterns.

Among all tested models, Gradient Boosting demonstrated the best performance, achieving the highest AUC score of 0.9953. It also exhibited faster processing time compared to other ensemble methods, highlighting its balance between computational efficiency and predictive power. This makes it particularly suitable for real-time or large-scale sleep health analysis applications.

Overall, Rahman et al. demonstrated that optimized machine learning techniques can play a critical role in modern sleep healthcare. Their results showcase how data-driven models can reduce diagnostic errors, improve early detection, and aid in designing personalized intervention plans. The study provides a strong foundation for future research that could integrate deep learning methods, wearable sensor data, or hybrid AI architectures to further refine sleep disorder prediction and real-time health monitoring systems.

TITLE: "A Systematic Review on Sleep Stage Classification and Sleep Disorder Detection Using Artificial Intelligence" **[5]**

AUTHORS: T. U. Wara, A. H. Fahad, A. S. Das, & M. M. H. Shawon

Wara et al. conducted a comprehensive systematic review to examine the applications of artificial intelligence (AI) in sleep stage classification and the detection of sleep disorders. The study analyzed 81 research papers published between 2016 and 2023, focusing primarily on EEG signals to understand and classify different sleep stages and disorders. The review aimed to identify the most effective AI models and highlight trends in performance, interpretability, and applicability in real-world sleep health monitoring.

Among the AI models reviewed, Convolutional Neural Networks (CNNs) were the most frequently used (27%), followed by Long Short-Term Memory networks (LSTM) at 11%, and traditional machine learning methods like Support Vector Machine (SVM) and Random Forest (RF), each at 6%. Recurrent Neural Networks (RNNs) accounted for 5% of the studies, indicating a growing interest in deep learning architectures capable of handling sequential sleep data. These models were evaluated using metrics such as accuracy, F1-score, sensitivity, and specificity, with reported accuracies reaching up to 83.75%.

The review highlighted that CNNs were particularly effective in capturing spatial and temporal patterns in EEG data, while hybrid approaches combining CNNs with LSTM or attention mechanisms showed potential for improving temporal understanding and robustness. The authors also emphasized the importance of balanced datasets, advanced feature extraction, and interpretability methods to enhance model reliability and clinical applicability.

Wara et al. concluded that AI techniques have strong potential to improve the accuracy, efficiency, and scalability of sleep disorder detection systems. Their work underscores the growing role of deep learning and machine learning in sleep healthcare and provides valuable insights for future research focused on automated, real-time monitoring, early diagnosis, and personalized interventions for patients with sleep disorders.

TITLE: "Prediction of Sleep Disorders Using Novel Decision Support Neutrosophic-Based Machine Learning Models" **[6]**

AUTHORS: N. R. Panda, S. Paramanik, P. K. Raut, & R. Bhuyan

Panda et al. proposed a novel decision support system to predict sleep disorders by integrating neutrosophic logic with conventional machine learning models. The approach was designed to handle uncertainty, indeterminate data, and inconsistencies commonly present in medical datasets, thereby improving both interpretability and robustness of predictions. This framework also allows for better decision-making in clinical settings where data ambiguity is a common challenge.

The study employed classifiers such as Decision Trees, Naive Bayes, and Support Vector Machines, enhanced with neutrosophic logic to process ambiguous or conflicting information effectively. By combining neutrosophic reasoning with machine learning, the system could better account for incomplete or uncertain sleep-related data, which is often a limitation in traditional predictive models. The authors noted that this method could be adapted to other medical domains with similar uncertainty challenges.

The proposed models achieved high accuracy in detecting sleep disorders, demonstrating that neutrosophic-based machine learning can significantly enhance medical diagnostics. The study highlighted that integrating uncertainty modeling with predictive analytics not only improves performance but also increases the reliability of recommendations provided by decision support systems. The results suggest that such models could be particularly valuable in early screening programs where precise data collection is difficult.

Panda et al. demonstrated that neutrosophic-based approaches offer a promising avenue for developing robust, interpretable, and accurate predictive systems in healthcare. This framework has the potential to support clinicians in early detection of sleep disorders, guide personalized interventions, and manage complex or ambiguous patient data more effectively. The study encourages further exploration of neutrosophic logic in combination with deep learning models to expand its applicability and predictive power.

TITLE: "Predicting Sleep Disorders: Leveraging Sleep Health and Lifestyle Data with Dipper Throated Optimization Algorithm for Feature Selection and Logistic Regression for Classification" **[7]**

AUTHORS: E. M. El-Kenawy, A. Ibrahim, A. A. Abdelhamid, N. Khodadadi, L. Abualigah, & M. M. Eid

El-Kenawy et al. proposed an intelligent predictive model for detecting sleep disorders by integrating the Dipper Throated Optimization Algorithm (DTOA) with Logistic Regression. The DTOA was used for effective feature selection, helping to identify the most significant variables related to sleep health and lifestyle factors that contribute to the development of sleep disorders. This optimization method mimics the behavior of dipper-throated birds to efficiently search and select optimal features, reducing noise and computational complexity in the dataset.

After selecting the key features, Logistic Regression was employed as the classification model to determine whether an individual had a sleep disorder or not. This combination ensured both interpretability and efficiency, as Logistic Regression offers simplicity and strong generalization capabilities, especially when paired with optimized inputs. The approach also enhanced the model's ability to handle nonlinear relationships between lifestyle habits and sleep outcomes, making it suitable for practical healthcare applications.

The experimental results showed that the proposed hybrid model achieved a high accuracy of 95%, along with excellent sensitivity and specificity values. These metrics demonstrate the robustness and consistency of the model in correctly identifying individuals with sleep disorders while minimizing false positives. The findings suggest that integrating metaheuristic optimization with classical machine learning algorithms can lead to substantial improvements in diagnostic accuracy.

El-Kenawy et al. demonstrated that combining DTOA with Logistic Regression provides a powerful yet computationally efficient solution for early prediction of sleep disorders. The study contributes to the growing field of intelligent healthcare by showing how nature-inspired optimization can enhance traditional classification models. Furthermore, the authors suggested that future research could extend this framework by incorporating deep learning architectures and real-time data from wearable devices to improve predictive performance and real-world usability.

TITLE: “Machine Learning-Based Prediction of Sleep Disorders from Lifestyle and Physiological Data: A Cross-Occupational Study” **[8]**

AUTHORS: H. K. Sari, S. Shoelarta, T. O. Pratama, G. N. Sajida, G. M. Krista, Y. F. Ferawati, & T. Taufiqurrahim

Sari et al. conducted a comprehensive cross-occupational study aimed at predicting sleep disorders using machine learning models trained on both lifestyle and physiological data. The study included individuals from various professions, recognizing that occupational differences can significantly influence sleep quality and the risk of developing sleep disorders. This multidisciplinary approach highlights the importance of incorporating occupational and lifestyle diversity when building predictive health models.

The dataset used in this research provided a wide range of behavioral and physiological indicators, making it well-suited for evaluating the impact of professional routines on sleep health. Data preprocessing and normalization were performed to ensure consistency, while advanced machine learning techniques were applied to extract meaningful insights. This allowed the researchers to handle both structured and unstructured lifestyle data effectively.

Several machine learning algorithms were tested, including XGBoost, Random Forest, and Support Vector Machine (SVM). Among these, the XGBoost model achieved the highest accuracy of 90%, outperforming Random Forest (87.5%) and SVM (82.5%). The high performance of XGBoost was attributed to its ability to manage feature interactions and handle imbalanced data efficiently. These results underscore the model’s robustness and its potential to generalize across diverse occupational groups, offering reliable predictions in real-world health monitoring scenarios.

This study demonstrates that machine learning can play a vital role in understanding and predicting sleep disorders across different professional backgrounds. By leveraging features related to lifestyle, work behavior, and physiological conditions, the model offers a powerful diagnostic and preventive tool for healthcare providers. The authors also suggested that integrating real-time data from wearable devices and expanding the study to include larger and more diverse populations could further enhance prediction accuracy. This research highlights the growing relevance of AI-driven health analytics in improving sleep quality, productivity, and overall well-being across various occupational sectors.

TITLE: "Sleep Disturbances and Depression Risk: A Meta-Analysis of Longitudinal Studies"**[9]**

AUTHORS: M. Li, Y. Zhang, J. Li, L. Ma, et al.

Li et al. conducted a comprehensive meta-analysis of longitudinal studies to explore the association between sleep disturbances and the risk of developing depression over time. The research examined various types of sleep problems, including insomnia, poor sleep quality, and irregular sleep durations, defined as short (<6 hours) or long (>9 hours) sleep periods. The analysis included a large number of participants across different demographic groups, allowing for a more reliable and generalized understanding of how sleep health influences mental well-being. By combining results from multiple studies, the authors aimed to clarify inconsistencies in previous findings and provide a robust estimate of the risk relationship.

Using a random-effects model alongside subgroup and sensitivity analyses, the study found that individuals with sleep disturbances had a significantly higher likelihood of developing depression. The pooled risk ratio (RR) was calculated at 2.01 (95% CI: 1.58–2.55), indicating that people experiencing sleep problems were twice as likely to develop depression compared to those with healthy sleep patterns. Among the sleep disturbances examined, insomnia symptoms posed the greatest risk (RR = 2.27), followed closely by poor sleep quality (RR = 1.92). Both short and long sleep durations were also linked to elevated depression risk, suggesting that maintaining an optimal sleep range is crucial for mental health stability.

The findings were consistent across different subgroups, including gender and age categories, reinforcing the conclusion that the relationship between sleep and depression is both strong and universal. The study highlights the critical role of sleep regulation in mental health maintenance and suggests that early identification and management of sleep problems may serve as an effective preventive strategy against depression. The authors also emphasized the importance of integrating sleep-focused interventions in psychological and psychiatric care.

This meta-analysis provides valuable evidence that sleep disturbances are not just symptoms but significant predictors of depression. By identifying sleep irregularities as a key modifiable risk factor, the research underscores the potential for early lifestyle and behavioral interventions to improve sleep quality and reduce the incidence of mood disorders. These insights support the growing recognition of sleep health as an essential component of preventive mental healthcare and overall psychological resilience.

TITLE: "Classification of Sleep Disorders Using Random Forest on Sleep Health and Lifestyle Dataset" **[10]**

AUTHORS: I. A. Hidayat

Hidayat et al. conducted a study focused on classifying sleep disorders using the Random Forest algorithm applied to the Sleep Health and Lifestyle Dataset. The primary objective was to determine how effectively Random Forest could identify different categories of sleep disorders by analyzing lifestyle and physiological factors. The dataset included attributes such as sleep duration, quality, physical activity, and screen time, allowing the model to learn diverse behavioral and health patterns. The research emphasized the importance of data preprocessing to ensure accurate results, including the treatment of missing values, normalization, and categorical encoding.

The Random Forest model was implemented using 200 decision trees, each with a maximum depth of 10, and the Gini Index was employed to determine optimal feature splits. This configuration allowed the model to balance accuracy and interpretability while preventing overfitting. The model's performance was evaluated using a confusion matrix, which revealed that the algorithm could reliably distinguish between individuals with and without sleep disorders.

Further analysis included feature importance evaluation and correlation visualizations, which provided deeper insights into how different lifestyle factors influenced sleep disorders. The study also highlighted noticeable gender-based patterns in sleep health, demonstrating that certain habits and physiological conditions varied significantly between males and females. These insights emphasized the potential for personalized approaches in addressing sleep-related health issues.

The study demonstrated that Random Forest is not only an accurate and robust algorithm for sleep disorder prediction but also an interpretable and practical choice for real-world healthcare applications. The findings reinforce that machine learning models, particularly ensemble techniques like Random Forest, can effectively support early diagnosis and intervention by analyzing lifestyle and health-related data. Hidayat et al. concluded that integrating such predictive models into healthcare systems could help develop personalized treatment plans and promote better sleep hygiene across diverse populations.

TITLE: "Enhancing Sleep Disorder Diagnosis with a Machine Learning Approach Using Ensemble Neural Networks" **[11]**

AUTHORS: M. S. Alom, S. M. Jeba, A. Debnath, T. T. Aurpa, & R. Siddiqua

Alom et al. conducted a study aimed at improving the accuracy of sleep disorder diagnosis using ensemble Artificial Neural Networks (ANNs) on the Sleep Health and Lifestyle Dataset. The main objective was to develop an intelligent predictive model capable of classifying sleep disorders such as Insomnia and Sleep Apnea by analyzing various behavioral and physiological factors. The dataset incorporated parameters like sleep duration, physical activity, caffeine intake, and stress level, providing a comprehensive understanding of lifestyle influences on sleep health. The authors focused on refining preprocessing steps, including missing value treatment, outlier removal, normalization, and categorical encoding, to ensure the model's robustness and accuracy.

Three ensemble ANN techniques Bagging, Boosting, and Weighted Average—were implemented to enhance the classification performance. Among these, both ANN Bagging and ANN Boosting achieved the highest accuracy of **94.7%**, while the Weighted Average Ensemble attained **92.9%**. These ensemble strategies effectively reduced overfitting and improved model generalization by combining the outputs of multiple neural networks.

The study highlighted the superior performance of ensemble ANNs compared to traditional machine learning algorithms such as Decision Trees and Logistic Regression. By leveraging the collective decision-making of multiple neural networks, the models achieved enhanced reliability and stability in predictions. Additionally, feature correlation and weight visualization helped interpret how lifestyle variables contributed to sleep disorder risks, demonstrating the explainability of deep ensemble approaches.

The research established that ensemble ANN models provide a powerful and accurate method for detecting sleep disorders and can be effectively used for real-world healthcare applications. The findings suggest that integrating ensemble neural networks into clinical decision support systems can assist medical professionals in early diagnosis, personalized treatment, and preventive interventions for sleep-related health issues. Alom et al. concluded that their proposed ensemble framework could significantly enhance the precision and dependability of AI-driven sleep disorder prediction systems.

CHAPTER III

DATASET DESCRIPTION

The dataset is designed to analyze and predict sleep disorders and overall sleep health based on lifestyle, physiological, and behavioral factors. It includes a wide range of attributes covering personal details, daily habits, health indicators, and environmental conditions that may influence sleep quality. Each record in the dataset represents an individual's sleep and health profile, offering valuable insights for machine learning and deep learning models.

Attributes and Their Descriptions

- **Name :** Represents the participant's identifier or name, used for record differentiation.
- **Age :** The age of the individual in years, reflecting demographic variation and its impact on sleep health.
- **Gender :** Specifies the participant's gender, useful for identifying sleep pattern differences between males and females.
- **Height_cm :** The individual's height measured in centimeters, essential for calculating BMI and analyzing body composition effects on sleep.
- **Weight_kg :** The individual's weight in kilograms, paired with height to assess body mass index (BMI).
- **Chronic_conditions :** Indicates the presence of chronic illnesses such as diabetes or hypertension that can influence sleep quality.
- **Mental_conditions :** Describes any mental health issues like anxiety or depression that may disrupt normal sleep cycles.
- **Medication :** Lists whether the participant is on medication, as certain drugs can affect sleep duration and quality.
- **Caffeine_per_day :** The number of caffeinated drinks consumed per day, a key factor influencing alertness and sleep latency.
- **Smoking_status :** Specifies whether the person smokes, since nicotine is known to interfere with healthy sleep patterns.
- **Alcoholic_drinks_per_week :** The average number of alcoholic beverages consumed weekly, another factor that may disturb sleep architecture.

- **Exercise_duration :** Describes the duration or frequency of physical activity, as regular exercise is correlated with improved sleep quality.
- **Screen_time_hrs :** Average daily screen exposure in hours, used to assess the impact of electronic device usage on sleep onset and quality.
- **Diet_quality :** Represents the nutritional balance of the individual's diet, which can influence overall health and sleep consistency.
- **Sleep_duration :** The total hours of sleep per night, a primary variable for evaluating sleep sufficiency and labeling sleep health categories.
- **Time_to_fall_asleep :** The time (in minutes) it takes for the participant to fall asleep after going to bed, indicating potential sleep onset issues.
- **Night_awakenings :** The number of times the individual wakes up during the night, reflecting sleep fragmentation.
- **Weekday_bedtime :** Usual bedtime during weekdays, providing insight into regular sleep routines.
- **Weekday_wakeup :** The typical wake-up time on weekdays, used to determine total weekday sleep duration.
- **Weekend_bedtime :** The bedtime during weekends, helping identify changes in sleep patterns due to lifestyle or social habits.
- **Weekend_wakeup :** Wake-up time on weekends, often later than weekdays, indicating possible sleep compensation.
- **Sleep_quality :** A subjective measure (e.g., Poor, Fair, Good) representing the participant's overall sleep satisfaction.
- **Bedroom_noise_level :** A numeric rating indicating how noisy the sleeping environment is, as higher noise levels reduce sleep quality.
- **Bedroom_light_level :** Measures the amount of light in the bedroom, since excessive brightness can disrupt melatonin production.
- **Comfort_level :** Rates the comfort of the sleeping surface (mattress and pillows), influencing sleep depth and restfulness.
- **Daytime_sleepiness :** A score indicating how sleepy the participant feels during the day, a common symptom of poor nighttime sleep.
- **Memory_issues :** Notes whether the individual experiences forgetfulness or concentration problems related to sleep deprivation.

- **Blood_pressure :** The participant's blood pressure readings, as poor sleep is linked to hypertension and cardiovascular strain.
- **Heart_rate :** Resting heart rate, which can signal stress levels or underlying health conditions affecting sleep.
- **Daily_steps :** The average number of steps taken per day, reflecting physical activity level and overall lifestyle balance.
- **Other_health_conditions :** Mentions additional health issues not covered in other fields that may influence sleep quality.
- **Urge_to_move_leg_score :** A numeric indicator of leg movement urges during rest, associated with Restless Leg Syndrome.
- **Daytime_sleep_attacks :** Notes the occurrence of sudden sleep episodes during the day, commonly linked with narcolepsy.
- **Irregular_work_patterns :** Indicates whether the person has irregular or shift-based work schedules, which can disrupt circadian rhythms.
- **Unusual_behaviors :** Captures abnormal nighttime behaviors such as sleepwalking, talking, or movements indicative of parasomnias.

Demographic Information

Demographic information provides fundamental insights into the population characteristics of the dataset. It includes attributes such as age, gender, height, and weight, which are essential for understanding how sleep patterns may vary across different groups. Age can influence the duration and quality of sleep, while gender-related differences may affect susceptibility to certain sleep disorders. Height and weight are often used to calculate BMI, which can be associated with conditions like sleep apnea. Collectively, demographic data helps in identifying population-specific trends and tailoring predictive models to account for individual differences.

Lifestyle Factors

Lifestyle factors capture daily habits and routines that impact sleep health. This includes caffeine consumption, alcohol intake, smoking habits, exercise duration, screen time, and diet quality. High caffeine or alcohol consumption, excessive screen time, and irregular exercise patterns can interfere with normal sleep cycles, while a balanced diet and regular physical activity promote better sleep. By analyzing these habits, models can link lifestyle behaviors to sleep disturbances and provide actionable recommendations for improving sleep quality.

Physiological Attributes

Physiological attributes reflect the internal health conditions that may influence sleep patterns. Key features in the dataset include chronic health conditions, mental health conditions, medications, blood pressure, heart rate, and daily steps. Chronic illnesses and mental health issues such as anxiety or depression are known to disrupt sleep, while medication use can alter sleep architecture. Heart rate, blood pressure, and physical activity levels indicate overall health status and its relationship with restorative sleep. Including physiological data ensures that predictive models consider both internal and external influences on sleep.

Behavioral Features

Behavioral features provide information about sleep-related actions and environmental conditions affecting sleep. These include sleep duration, time to fall asleep, night awakenings, bedtime and wake-up times, daytime sleepiness, urge to move legs, daytime sleep attacks, irregular work schedules, unusual sleep behaviors, bedroom noise and light levels, and comfort levels. Such behaviors and environmental factors are critical in identifying patterns linked to sleep disorders like insomnia, sleep apnea, or restless leg syndrome. Monitoring these behaviors allows predictive models to detect irregularities, assess sleep quality, and support targeted interventions to improve overall sleep health.

Derived Features for Model Training

In this project, several derived features were created to enhance predictive analysis of sleep disorders. One key derived feature is BMI (Body Mass Index), calculated using height and weight. BMI provides a standardized measure of body fat and is an important indicator of health, as higher BMI values are often associated with conditions like sleep apnea and other sleep-related problems. Including BMI allows the model to capture the relationship between body composition and sleep quality, enabling more accurate predictions.

Another important derived feature is the sleep disorder classification. This is implemented in a two-stage labeling system. In the first stage, the model determines whether an individual has a sleep disorder or not. In the second stage, for those diagnosed with a disorder, the system further classifies the type into one of five categories: Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, or Narcolepsy. This multi-stage approach ensures a detailed understanding of sleep problems and allows for personalized recommendations and interventions. By using these derived features alongside original demographic, lifestyle, physiological, and behavioral data, the project can provide a comprehensive and data-driven assessment of sleep health.

Significance of the Dataset

The dataset is significant because it provides a comprehensive view of the factors influencing sleep health, combining demographic, lifestyle, physiological, and behavioral attributes. It allows researchers and healthcare professionals to analyze how individual characteristics and daily habits impact sleep quality and the risk of developing sleep disorders. By including features such as age, gender, BMI, caffeine intake, exercise, sleep duration, and environmental factors, the dataset enables a multi-dimensional understanding of sleep patterns.

The inclusion of derived features like BMI and two-stage sleep disorder classification adds predictive power, allowing models to detect not only whether an individual has a sleep disorder but also the specific type. This level of detail supports personalized health monitoring, early detection of potential sleep problems, and tailored intervention strategies. Overall, the dataset serves as a valuable resource for developing machine learning and deep learning models aimed at improving sleep health, informing clinical decisions, and promoting overall well-being.

The dataset also facilitates research into the relationships between lifestyle choices, environmental factors, and physiological conditions with sleep outcomes. By providing rich, multi-faceted data, it enables the development of predictive models that can uncover hidden patterns and correlations, which may not be evident through traditional analysis methods. This makes it a powerful tool for preventive healthcare, as it allows early identification of at-risk individuals and the design of targeted interventions. Moreover, the dataset can support the creation of digital health solutions, such as personalized sleep monitoring apps and decision support systems for clinicians, ultimately contributing to better sleep management and overall public health outcomes.

CHAPTER IV

METHODOLOGY

The methodology for this project was designed to accurately predict sleep disorders and classify their types using demographic, lifestyle, physiological, and behavioral data. The process involved multiple stages, including data collection, preprocessing, exploratory data analysis (EDA), feature engineering, model implementation, performance evaluation, and future recommendations for deployment. Each step was carried out to ensure that the models are robust, interpretable, and applicable to real-world sleep health monitoring.

Data Preparation

- **Data Cleaning:** Handled missing values using appropriate imputation techniques to maintain consistency and reduce bias. Numerical variables like BMI and sleep duration were imputed with the median, while categorical variables such as gender, smoking status, and mental conditions were imputed with the mode.
- **Encoding Categorical Variables:** Transformed categorical features (e.g., gender, smoking status, diet quality) using one-hot or label encoding to ensure compatibility with machine learning algorithms. One-hot encoding prevented the introduction of ordinal relationships where none exist.
- **Scaling Numerical Features:** Standardized numerical features such as age, BMI, caffeine intake, sleep duration, and night awakenings to ensure uniform scale and prevent dominance of certain features. Min-max scaling was applied to bring values into a 0 - 1 range.
- **Outlier Detection and Treatment:** Outliers were detected using Z-score and IQR methods and were treated to minimize their impact on model performance. Extreme values in features like sleep duration, BMI, and screen time were adjusted to improve model stability.
- **Feature Engineering:** Derived new features such as BMI, difference between weekday and weekend sleep duration, and sleep irregularity scores to enhance prediction. The dataset also included two-stage labels for sleep disorders: first identifying whether a disorder exists, then classifying it into one of five types (Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, Narcolepsy).

- **Data Splitting:** The dataset was split into training and testing sets, with stratified sampling to maintain class distribution, ensuring models could generalize well to unseen data.

Approaches to Improve Performance

- **Exploratory Data Analysis (EDA):** Performed correlation analysis, heatmaps, and pair plots to examine relationships between features like sleep duration, caffeine intake, physical activity, screen time, and sleep quality. EDA helped identify patterns, trends, and potential predictive features, while also detecting outliers, missing values, and inconsistencies in the data.
- **Model Implementation:** Applied a wide range of machine learning models including Random Forest, XGBoost, Gradient Boosting, Logistic Regression, KNN, and SVM to benchmark predictive performance. Deep learning models such as Multi-Layer Perceptron (MLP) and Convolutional Neural Networks (CNN) were implemented to capture complex patterns in the data. Hybrid models, like CNN-DNDF, were also developed to combine convolutional feature extraction with decision-based reasoning, enhancing generalization.
- **Cross-Validation:** Used k-fold cross-validation to assess model robustness and generalizability, monitoring performance variance across folds to detect overfitting or underfitting. Stratified sampling was applied to maintain proportional distribution of sleep disorder classes, and both stages of classification were validated separately to ensure balanced evaluation.
- **Performance Evaluation:** Models were evaluated using accuracy, precision, recall, F1-score, confusion matrices, and ROC-AUC curves. Per-class performance was assessed to avoid neglecting rare sleep disorder classes. Learning curves and error analysis helped identify patterns in misclassification and refine feature engineering, ensuring reliable predictions.
- **Hyperparameter Tuning:** Optimized model parameters using grid search, random search, and Bayesian optimization. Neural network-specific parameters such as dropout rates and early stopping were tuned to prevent overfitting. Each change was evaluated using validation metrics, allowing iterative improvements in predictive performance.
- **Model Selection and Integration:** Selected the best-performing models based on multi-metric evaluation, considering both accuracy and computational efficiency.

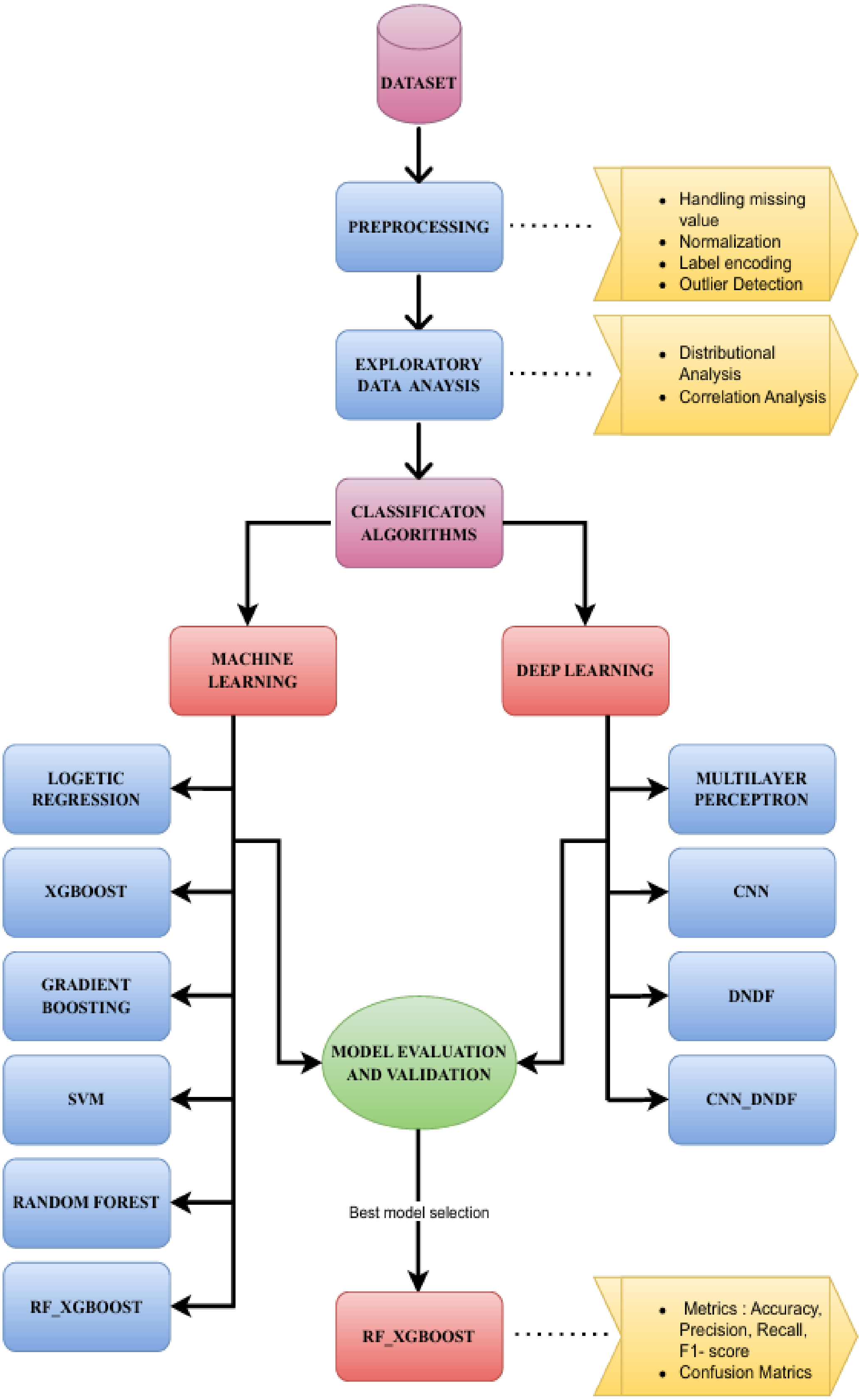

Figure1. Architecture of working methodology

Models Implemented

- **Logistic Regression (LR):** Logistic Regression was implemented as a baseline model to classify individuals based on the probability of having a sleep disorder. Despite its simplicity, it provides valuable interpretability by showing how each feature contributes to the prediction. It works well for binary and multiclass problems and helps in understanding linear relationships between predictors and outcomes.
- **Support Vector Machine (SVM):** SVM was used to find the optimal hyperplane that separates different sleep disorder classes. With kernel functions like RBF and polynomial, SVM effectively handled non-linear relationships between sleep-related features. It provided high accuracy and was especially efficient in dealing with complex decision boundaries, making it suitable for health data with overlapping characteristics.
- **Random Forest (RF):** Random Forest, an ensemble of multiple decision trees, was implemented to improve prediction stability and accuracy. It reduces overfitting by averaging multiple trees trained on different data subsets. The model provided insights into feature importance, highlighting influential variables such as sleep duration, caffeine intake, and stress levels. Its robustness made it one of the most reliable models in the project.
- **XGBoost (Extreme Gradient Boosting):** XGBoost, a highly efficient gradient boosting algorithm, was employed to enhance predictive power through sequential learning. It corrects errors from previous iterations and optimizes model performance using gradient descent. This model achieved high accuracy and excellent generalization, making it a top performer in both binary and multiclass classification of sleep disorders.
- **Gradient Boosting Classifier (GBC):** GBC was implemented as another ensemble model that builds trees sequentially to minimize classification errors. It offers strong predictive accuracy and resistance to overfitting. The model captured subtle relationships between behavioral and physiological features, improving the detection of complex sleep disorder types.
- **Multi-Layer Perceptron (MLP):** The MLP, a type of feedforward neural network, was applied to capture non-linear and high-dimensional relationships in the data. It consisted of multiple hidden layers with activation functions like ReLU and sigmoid, allowing the model to learn complex patterns between lifestyle, physiological, and

behavioral attributes. MLP improved prediction in cases where traditional models struggled with multi-feature interactions.

- **Convolutional Neural Network (CNN):** Although CNNs are typically used for image and sequential data, in this project they were adapted for tabular data to detect local feature dependencies. One-dimensional convolutions were applied across feature sets to extract hierarchical representations. This approach allowed CNNs to learn complex associations between features such as sleep duration, heart rate, and activity level.
- **CNN-DNDF (Deep Neural Decision Forest):** A hybrid model combining CNN and Decision Neural Decision Forest (DNDF) was implemented to leverage both deep feature extraction and decision-based interpretability. The CNN component automatically learned abstract features, while the DNDF performed probabilistic decision-making using tree-like structures. This hybrid design enhanced classification accuracy and interpretability, making it particularly effective for multi-stage sleep disorder detection.

Dataset extraction

The dataset for this project was created using a combination of real-time survey data and synthetic data generation to ensure diversity and balance. Real-time data was collected through a Google Form survey, where participants provided details about their demographics, lifestyle habits, physiological measures, and sleep-related behaviors, including factors like sleep duration, caffeine intake, physical activity, screen time, and bedtime routines. This real world data served as the foundation for understanding sleep patterns and health behaviors. To enhance the dataset and address issues such as class imbalance and limited sample size, synthetic data was generated using Python-based statistical methods while maintaining realistic feature distributions and logical consistency. The final dataset included derived features such as Body Mass Index (BMI) and a two-stage sleep disorder classification first identifying whether a person had a sleep disorder, and then categorizing it into one of five types: Insomnia, Sleep Apnea, Restless Leg Syndrome, Parasomnias, or Narcolepsy. This combined dataset provided a rich, comprehensive, and balanced representation of sleep health factors, forming a strong foundation for machine learning and deep learning model development.

Data Preprocessing

Handling Missing Values:

To maintain the integrity and reliability of the dataset, missing values were carefully identified and treated using suitable imputation strategies. For numerical features such as sleep duration, time to fall asleep, and night awakenings, missing values were replaced using the mean imputation method, ensuring that the overall data distribution remained unaffected. For text-based health-related attributes like chronic conditions, mental health conditions, and other health conditions, missing entries were filled with the label "NONE" to indicate the absence of information. Similarly, binary variables such as daytime sleep attacks, irregular work patterns, and unusual sleep behaviors were filled with "No" to maintain consistency and completeness in categorical fields.

Renaming and Cleaning Columns:

To ensure better readability and compatibility for further processing, column names were standardized using a consistent naming convention. All spaces, newline characters, and special symbols were removed, and columns were renamed with descriptive identifiers such as sleep_duration, screen_time_hrs, and time_to_fall_asleep. This step improved dataset usability, eliminated redundant formatting errors, and enabled smooth handling during feature engineering and model training.

Encoding Categorical Variables:

Categorical attributes such as gender and lifestyle indicators were standardized to uniform formats (e.g., converting "male/female" variations to "Male" and "Female"). Label encoding was performed to convert categorical data into numerical form for algorithm compatibility. For example, gender and sleep quality were encoded to ensure that models could interpret them numerically while preserving relationships among categories.

Outlier Detection and Treatment:

Outliers in numerical features, especially in time to fall asleep, were identified using the Interquartile Range (IQR) method. Values beyond realistic ranges (e.g., individuals taking over 3 hours to fall asleep) were capped at a logical threshold (180 minutes) to prevent distortion in model training. This ensured the dataset represented realistic human sleep patterns and reduced noise in learning.

Feature Engineering and Derived Features:

New derived attributes were introduced to enhance model interpretability and prediction performance. One key engineered feature was Body Mass Index (BMI), calculated using height and weight to assess obesity-related risks in sleep disorders like Sleep Apnea. This additional feature contributed significantly to the disorder classification logic and provided a meaningful health indicator within the dataset.

Sleep Disorder Labeling:

A rule-based function was developed to assign disorder labels. The first stage identified whether an individual had a sleep disorder (Yes/No) using parameters like caffeine intake, smoking, alcohol consumption, sleep duration, and night awakenings. The second stage classified the disorder type into categories Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, and Narcolepsy based on specific combinations of physiological and behavioral attributes. This two-step labeling process prepared the dataset for effective supervised learning.

Exploratory Data Analysis (EDA)

Exploratory Data Analysis was performed to understand the overall structure, patterns, and relationships within the dataset. Various visualizations were used to gain insights into demographic characteristics, behavioral patterns, and their association with different types of sleep disorders.

1. Distribution of Sleep Disorder Types

A count plot was used to examine the frequency of each sleep disorder category. The results revealed that Sleep Apnea and Insomnia were the most common disorders, while Parasomnias and Narcolepsy appeared less frequent. This imbalance indicates that the dataset is slightly skewed towards certain disorder types, which is important to consider during model training and evaluation.

2. Gender Distribution across Disorders

The gender-based analysis showed differences in how sleep disorders are distributed among males and females. Using a grouped bar chart, it was observed that Sleep Apnea was more prevalent among males, whereas Insomnia appeared more common in females. This insight aligns with medical studies that associate gender-specific factors with certain sleep disorders.

3. Age Distribution and Age-wise Disorder Analysis

The age distribution plot indicated that most participants fell between the ages of 20 and 50, representing a young to middle-aged population. When categorized into age groups, it was observed that Sleep Apnea cases were more common in the older age categories (40+), while Insomnia and Restless Leg Syndrome were more frequent among younger adults (20–40 years). These findings emphasize how age influences the type of sleep disturbance experienced.

4. Sleep Duration across Disorders

A boxplot comparison showed that individuals with Insomnia and Restless Leg Syndrome generally had shorter sleep durations, while those with No Disorder exhibited longer and more consistent sleep times. This variation clearly demonstrates how sleep duration can serve as an important predictive feature in detecting sleep disorders.

5. Correlation Analysis

A correlation heatmap was created to identify relationships between numerical variables such as sleep duration, caffeine intake, BMI, alcohol consumption, and daytime sleepiness. It was observed that:

- Sleep duration negatively correlated with screen time and caffeine intake.
- BMI showed a positive correlation with daytime sleepiness, highlighting its connection to Sleep Apnea.

These correlations provided deeper insights into how lifestyle and physiological factors interact in influencing sleep health.

6. Sleep Quality and Related Factors

The boxplot of sleep quality across disorders indicated significantly lower quality scores among individuals with Insomnia and Sleep Apnea. Additionally, scatter plots between caffeine intake, screen time, and sleep duration revealed that higher caffeine consumption and excessive screen exposure were both associated with reduced sleep duration.

7. BMI and Diet Quality Analysis

The BMI distribution across disorders revealed that individuals with Sleep Apnea tended to have higher BMI levels, supporting its known association with obesity. Furthermore, the diet quality analysis showed that participants reporting good to excellent diets generally experienced fewer sleep issues, while poor dietary habits were linked with Insomnia and Restless Leg Syndrome.

8. Daytime Sleepiness and Night Awakenings

Histograms and boxplots illustrated the distribution of daytime sleepiness and night awakenings across disorders. Individuals with Narcolepsy showed high daytime sleepiness scores, while those with Insomnia and Parasomnias exhibited frequent night awakenings. These behavioral indicators proved valuable in differentiating disorder categories.

9. Gender and Sleep Duration Relationship

A grouped bar chart comparing average sleep duration by gender and disorder type revealed that females, on average, reported slightly longer sleep durations in No Disorder and Insomnia categories, while males showed more variation across disorder types. These gender-specific trends highlight behavioral and biological differences in sleep health.

Classification

The classification process in this project involved applying advanced machine learning algorithms to predict sleep disorders and their specific types based on various demographic, lifestyle, physiological, and behavioral factors. The models implemented included Random Forest, XGBoost, and Gradient Boosting, all known for their strong predictive capabilities and adaptability to complex datasets. Random Forest, an ensemble-based algorithm, was used for its ability to handle large feature sets, reduce overfitting, and provide feature importance insights. XGBoost, a highly efficient gradient boosting algorithm, was chosen for its speed, accuracy, and ability to handle missing data effectively. Gradient Boosting, another ensemble method, helped improve prediction performance by sequentially building models that corrected the errors of previous ones, ensuring a more refined classification outcome.

These algorithms worked to classify individuals into sleep disorder and non-disorder categories, followed by identifying specific disorder types in the second stage. The ensemble nature of these models ensured robustness, while their interpretability made it possible to understand which features most influenced the predictions. Through this approach, the classification process not only produced reliable results but also provided valuable insights into how various factors like sleep habits, physical activity, and mental health contribute to sleep quality. This multi-model strategy created a strong foundation for developing an intelligent, data-driven system for sleep disorder prediction and health recommendations.

Logestic Regression

Logistic Regression is a type of supervised machine learning algorithm used to classify data into categories. Unlike linear regression, which predicts numerical values, logistic regression estimates the probability that an input belongs to a particular class, producing outputs between 0 and 1. Depending on the problem, it can handle binary classification, such as determining whether a person has a sleep disorder (Yes/No), or multiclass classification, like categorizing sleep quality into Poor, Moderate, Good, or Excellent. The algorithm works by examining the relationship between independent variables (features) like sleep duration, study hours, screen time, caffeine intake, and physical activity, and the dependent categorical variable (sleep quality or disorder).

In this project, Logistic Regression was applied to predict students' sleep quality and identify potential sleep disorders. It calculated the likelihood of a student falling into each category and allowed early detection of poor sleep patterns, providing actionable insights for recommendations. The core of the method is the sigmoid function, which converts a weighted sum of features into a probability between 0 and 1. Mathematically, it is represented as

$$\hat{y} = \frac{1}{1+e^{-z}},$$

where $z = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n$,

$b_0$is the intercept,

$b_i$are the feature coefficients.

Using a simple decision rule, predictions above 0.5 are assigned to class 1 (e.g., Sleep Disorder = Yes), and those below 0.5 to class 0 (e.g., Sleep Disorder = No). This approach also provided interpretable results, showing how features like increased screen time or low physical activity could raise the probability of poor sleep quality.

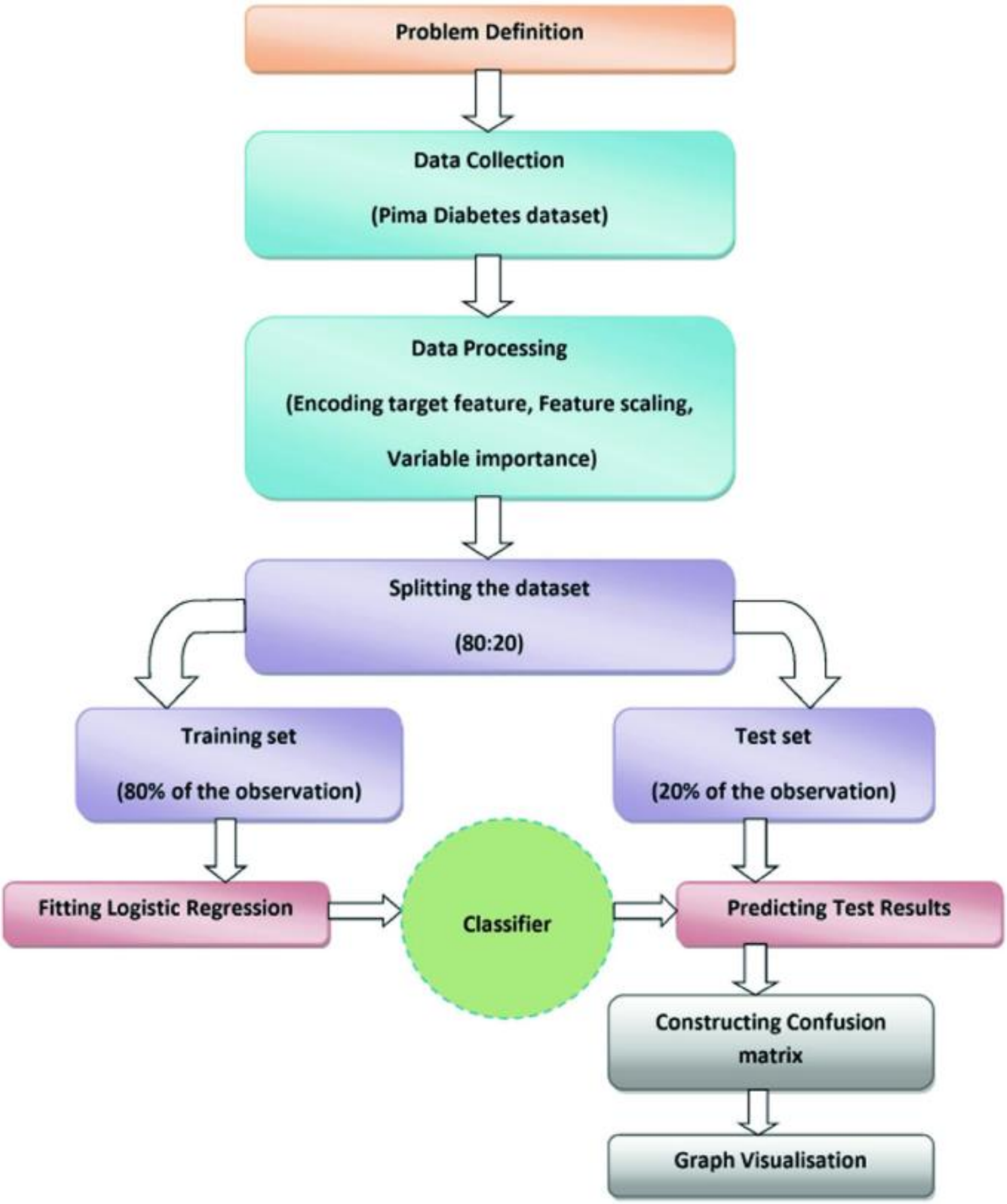

Figure 2. Work flow of Logestic Regression

Pseudocode for Logistic Regression Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target)

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize Logistic Regression model.

Step 8: Train the model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

Random Forest

Random Forest is an ensemble machine learning method that builds multiple decision trees on random subsets of data and features, then combines their predictions to produce a more accurate and stable result. For a two-stage sleep disorder prediction, Random Forest is well-suited because it can handle complex interactions between multiple physiological and lifestyle features, reduce overfitting compared to single decision trees, and provide reliable classification even when the dataset has noise or missing values. Its ability to output feature importance also helps identify which factors most strongly influence the onset or severity of sleep disorders, making it ideal for both stages of prediction.

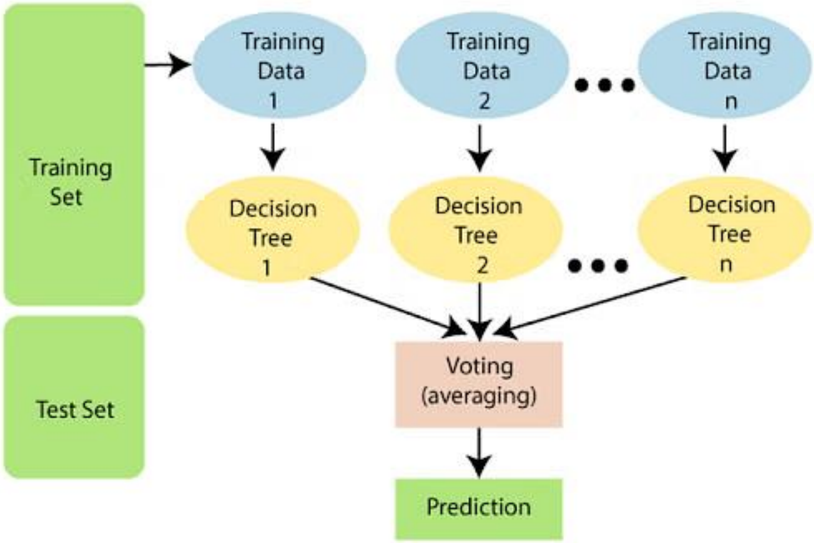

Figure 3. Work flow of Random Forest

Pseudocode for Random Forest Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target).

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize Random Forest model with suitable parameters.

Step 8: Train the model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

XG Boost

XGBoost (Extreme Gradient Boosting) is a powerful machine learning algorithm based on the concept of boosting, where multiple weak models (usually decision trees) are combined to form a strong predictive model. Unlike Random Forest, which builds trees independently, XGBoost builds trees one after another, with each new tree trying to correct the errors made by the previous ones. It uses gradient descent to minimize the loss function efficiently, making it both fast and accurate. The algorithm also includes regularization to prevent overfitting and handles missing data automatically, which makes it very practical for real-world datasets.

In this project, XGBoost was used to predict sleep quality and sleep disorder risk based on features like sleep duration, caffeine intake, screen time, and physical activity. It helped identify complex patterns and interactions among the variables, giving highly accurate predictions. By learning from the mistakes of earlier models, XGBoost provided better results than many traditional algorithms. It also helped highlight which factors most affected sleep health, making it useful for understanding the lifestyle habits linked to poor sleep. Overall, XGBoost proved to be an efficient and reliable model for analyzing and predicting sleep-related patterns.

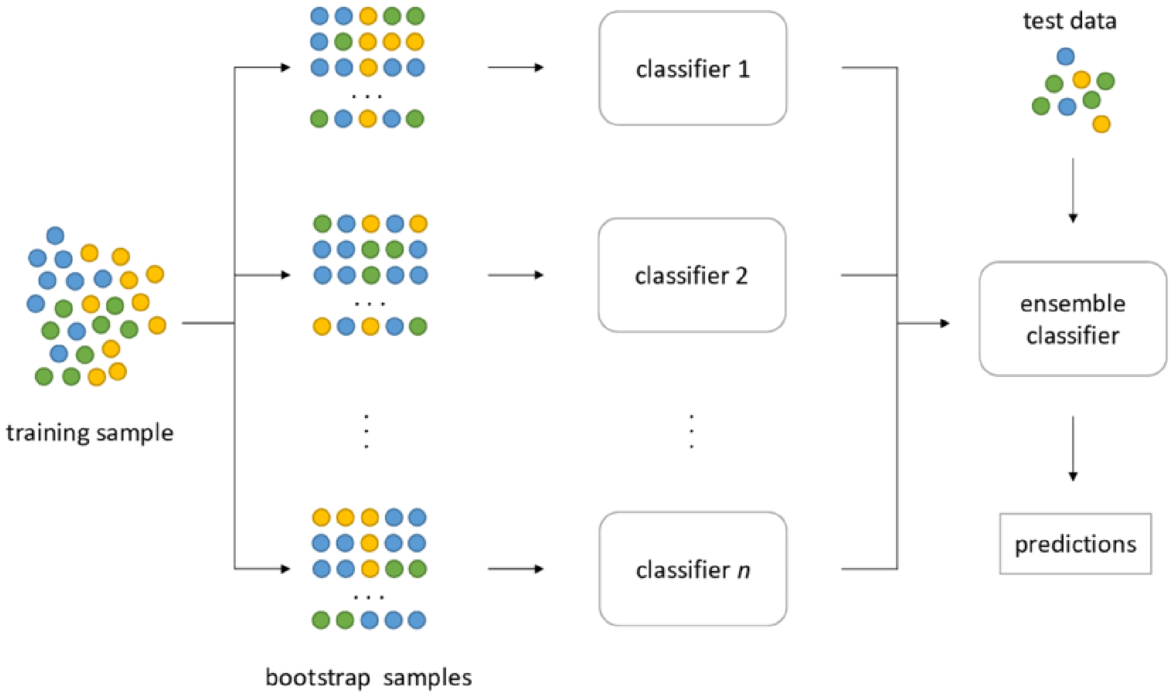

Figure 4. Work flow of XG Boost

Pseudocode for XG Boost Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target).

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize XGBoost model with suitable parameters.

Step 8: Train the model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

Gradient boosting

Gradient Boosting is a machine learning algorithm that builds multiple weak models, usually decision trees, in sequence where each new model corrects the errors made by the previous ones. It uses a gradient descent technique to minimize prediction errors and improve performance step by step. In this project, Gradient Boosting was applied to classify sleep quality and predict the risk of sleep disorders based on lifestyle and health-related factors such as sleep duration, screen time, physical activity, and stress level. The algorithm effectively captured complex relationships between these variables and sleep outcomes. With proper tuning of parameters like learning rate and tree depth, the model achieved strong accuracy and provided meaningful insights into how daily habits influence sleep health. Additionally, Gradient Boosting proved to be stable even with non-linear and noisy data, making it suitable for real-world sleep prediction. It also offered interpretability through feature importance scores, helping identify which habits most affect sleep quality.

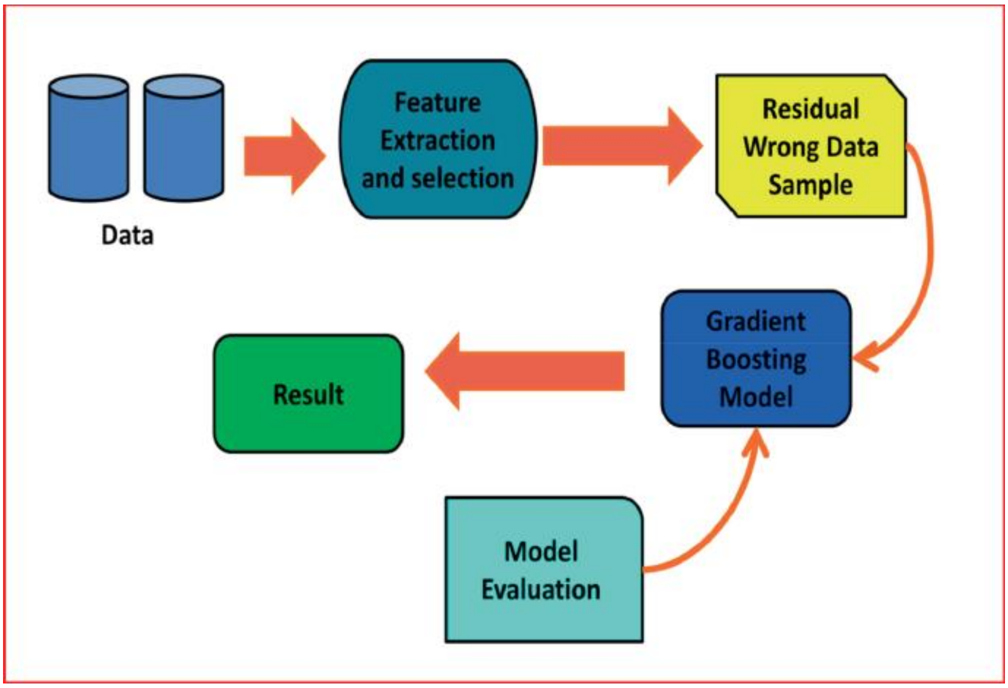

Figure 5. Work flow of Gradient Boosting

Pseudocode for Gradient Boosting Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target).

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize Gradient Boosting model with suitable parameters.

Step 8: Train the model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

Support vector Machine

Support Vector Machine (SVM) is a supervised machine learning algorithm used for both classification and regression tasks. It works by finding the optimal hyperplane that best separates data points of different classes in a high-dimensional space. In this project, SVM was used to classify sleep quality and detect potential sleep disorders based on various health and lifestyle features such as sleep duration, caffeine intake, screen time, and physical activity. The algorithm performed well in handling complex and non-linear data by using kernel functions like the Radial Basis Function (RBF), which maps input features into a higher-dimensional space for better separation. SVM is particularly effective for smaller datasets and avoids overfitting by focusing on the most critical data points, known as support vectors. It provided reliable predictions of sleep disorder categories while maintaining a strong balance between precision and recall. Overall, SVM contributed to improving the robustness of the model ensemble and offered valuable insights into how lifestyle factors influence sleep health.

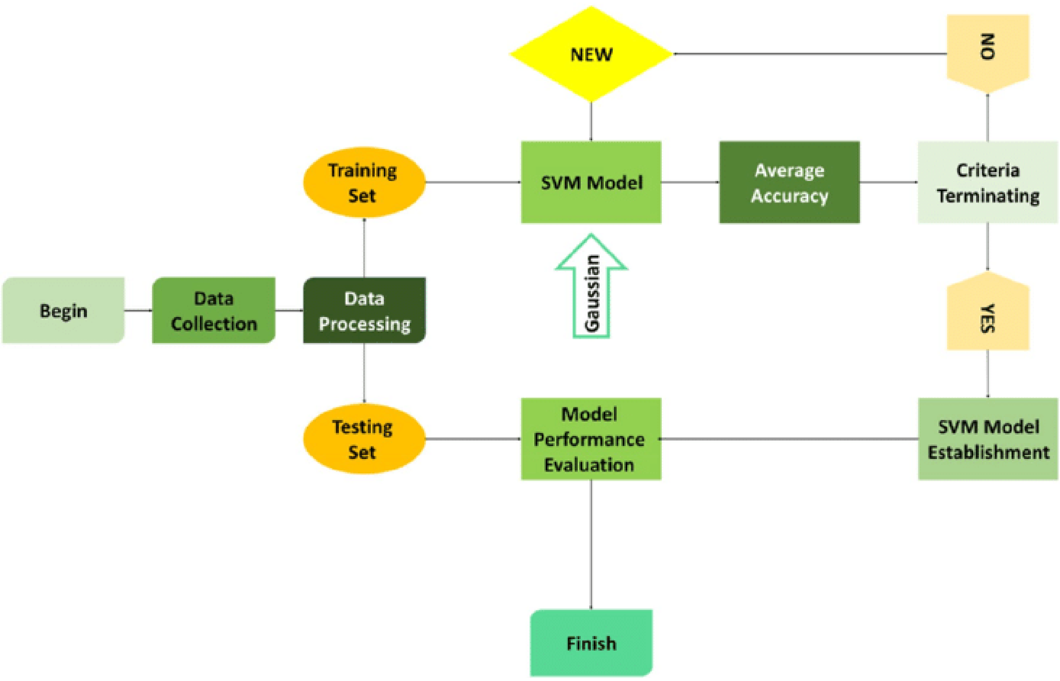

Figure 6. Work flow of SVM

Pseudocode for SVM Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target).

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize SVM model with suitable kernel and parameters.

Step 8: Train the model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

RF_XG Boost

RF–XGBoost Hybrid Model is a powerful ensemble learning approach that combines the strengths of Random Forest and XGBoost to achieve higher accuracy and robustness in prediction tasks. Random Forest focuses on reducing overfitting by averaging multiple decision trees built from random subsets of data, while XGBoost enhances predictive power through sequential learning and gradient-based optimization. In this project, the RF–XGBoost hybrid model was implemented to classify sleep disorders more accurately by leveraging both algorithms' strengths. Random Forest provided stable base predictions, and XGBoost refined these predictions by focusing on the difficult-to-classify cases. This combination helped capture both linear and non-linear relationships among features like sleep duration, caffeine intake, exercise level, and screen time. The hybrid model not only improved overall accuracy but also enhanced the interpretability of important sleep-related factors.

Pseudocode for RF_XG Boost Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target).

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features (if needed).

Step 7: Initialize XGBoost and Random Forest models with suitable parameters.

Step 8: Combine the models into a hybrid/stacking model.

Step 9: Train the hybrid model on training data.

Step 10: Predict outcomes on testing data.

Step 11: Evaluate performance: accuracy, precision, recall, F1-score.

Step 12: Generate confusion matrix and classification report.

Step 13: Measure total execution time.

Step 14: Display accuracy, confusion matrix, classification report, and execution time.

Multi layer Perceptron

Multilayer Perceptron (MLP) is a deep learning algorithm that belongs to the family of feedforward artificial neural networks. It consists of an input layer, one or more hidden layers, and an output layer, where each neuron is connected to every neuron in the next layer through weighted connections. MLP learns complex non-linear relationships in the data by adjusting these weights using backpropagation and optimization techniques such as gradient descent. In this project, MLP was used to predict sleep disorders by analyzing multiple lifestyle and health-related factors like sleep duration, screen time, caffeine intake, and physical activity. Its ability to learn complex feature interactions made it suitable for handling the diverse nature of the dataset. The model effectively captured patterns between behavioral and physiological parameters, improving the accuracy of sleep disorder classification. Moreover, MLP showed good generalization ability and adaptability, making it an effective model for real-world health monitoring applications.

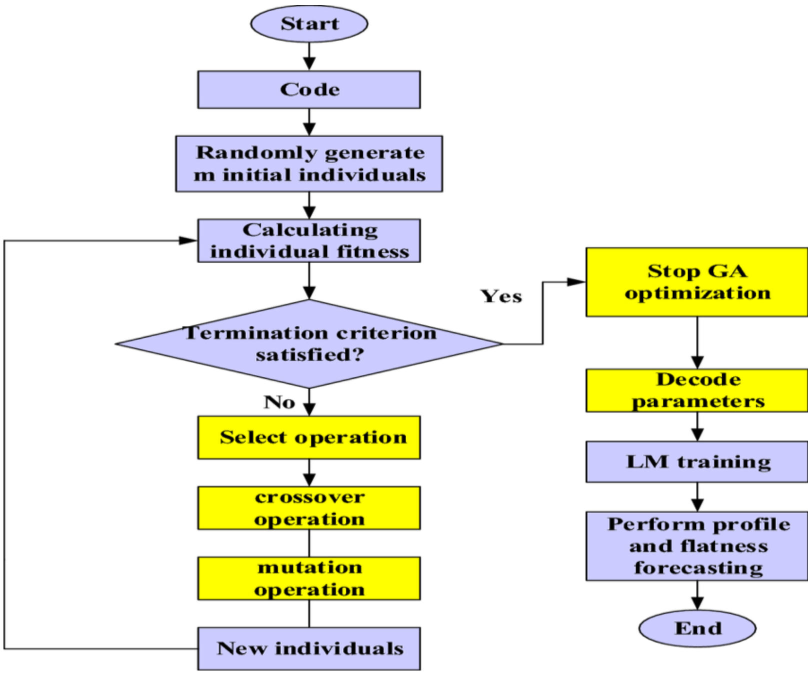

Figure 7. Work flow of MLP

Pseudocode for MLP Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define independent variables (features) and dependent variable (target).

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize MLP model with suitable parameters (hidden layers, activation function, solver).

Step 8: Train the model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

Convolutional neural network

Convolutional Neural Network (CNN) is a deep learning model primarily designed to process data with spatial or sequential relationships by using convolutional layers that automatically extract important patterns and features. Unlike traditional neural networks, CNNs use filters (kernels) that slide over the input data to detect local patterns, reducing the need for manual feature extraction. In this project, CNN was adapted for tabular sleep health data to identify complex relationships between lifestyle, physiological, and behavioral factors. The convolutional layers helped the model recognize subtle variations in features such as sleep duration, caffeine intake, and stress levels that contribute to sleep disorders. By capturing hierarchical feature representations, CNN improved classification accuracy and robustness. Additionally, its ability to generalize well on unseen data made it suitable for identifying patterns across different types of sleep disorders, supporting more accurate and efficient predictions in real-world scenarios.

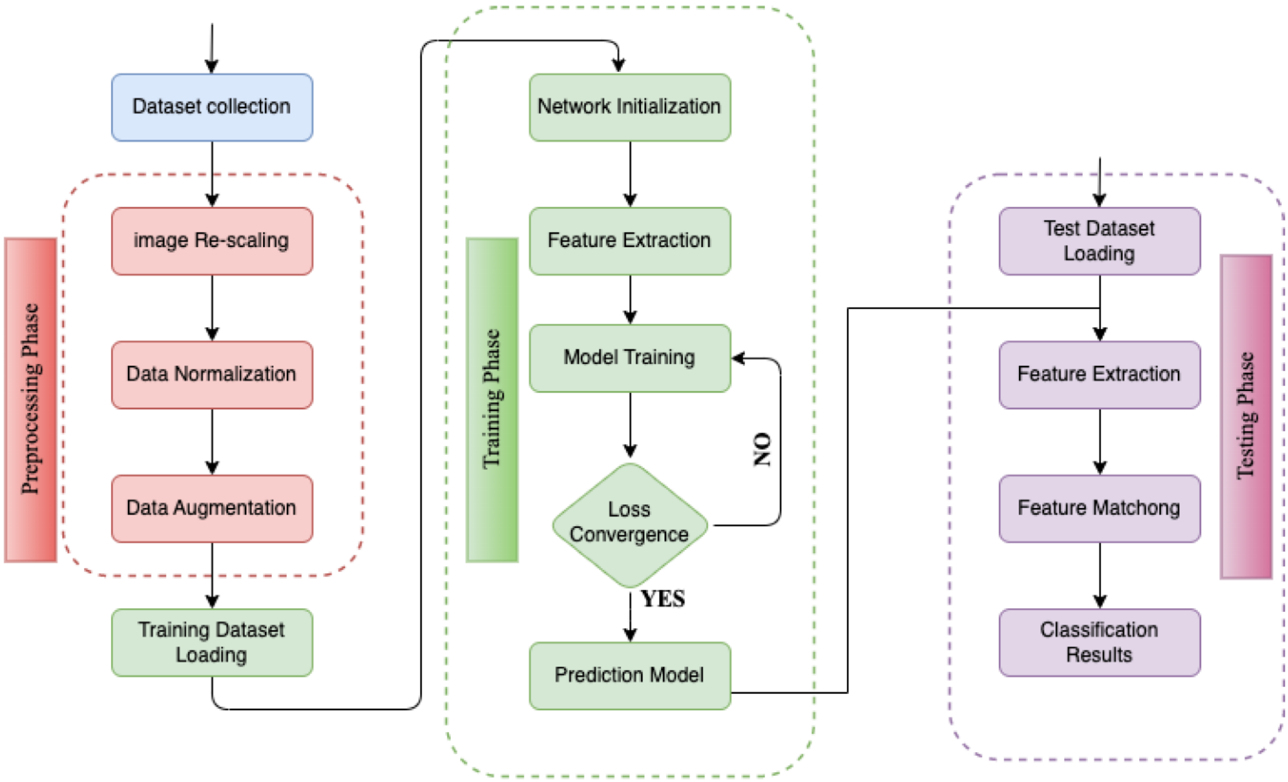

Figure 8. Work flow of CNN

Pseudocode for CNN Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, modeling (deep learning), and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features, and reshape data if needed.

Step 4: Define input features and target variable.

Step 5: Split data into training and testing sets.

Step 6: Standardize or normalize numerical features.

Step 7: Build the CNN architecture (convolutional layers, pooling layers, fully connected layers).

Step 8: Compile the CNN model with suitable optimizer, loss function, and metrics.

Step 9: Train the CNN model on training data.

Step 10: Predict outcomes on testing data.

Step 11: Evaluate performance: accuracy, precision, recall, F1-score.

Step 12: Generate confusion matrix and classification report.

Step 13: Measure total execution time.

Step 14: Display accuracy, confusion matrix, classification report, and execution time.

Deep neural decision forest

Deep Neural Decision Forest (DND-F) is a hybrid deep learning model that combines the representational power of neural networks with the structured decision-making of decision trees. Unlike traditional neural networks, DND-F integrates differentiable decision nodes that allow end-to-end training while capturing complex, non-linear relationships in the data. In this project, DND-F was applied to tabular sleep health data to classify sleep disorders by leveraging both continuous and categorical features. The neural network layers extracted high level feature representations from variables such as sleep duration, screen time, caffeine intake, and physical activity, while the decision forest component structured these features into interpretable decision paths. This hybrid approach enabled the model to detect subtle patterns and interactions among lifestyle, physiological, and behavioral factors, improving classification accuracy and robustness

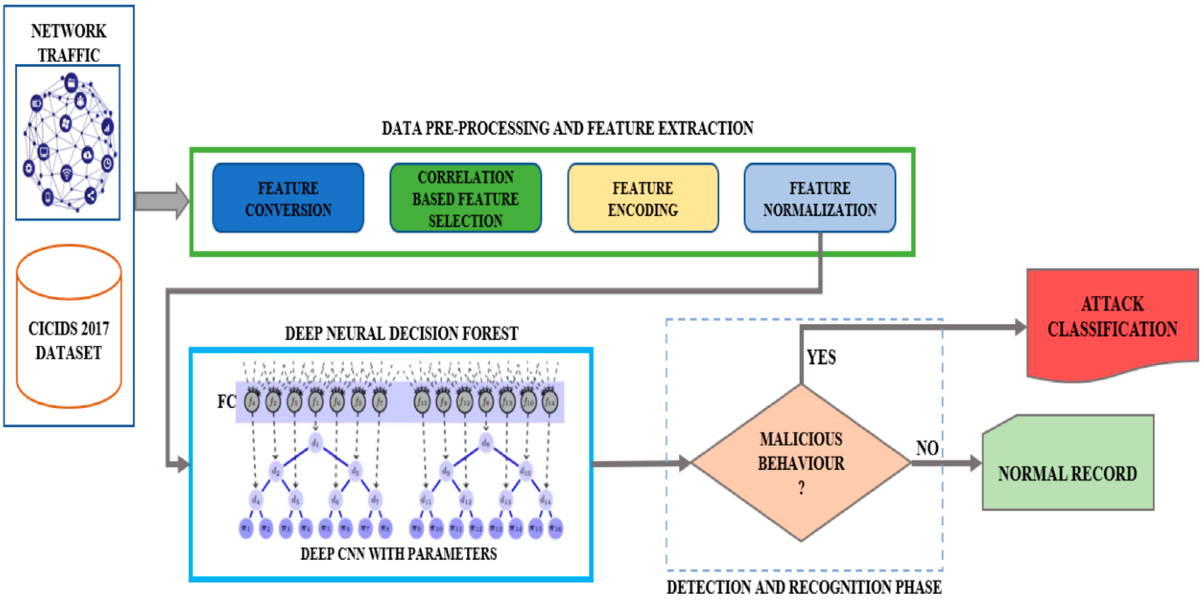

Figure 9. Work flow of DNDF

Pseudocode for DNDF Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, deep learning, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features.

Step 4: Define input features and target variable.

Step 5: Split data into training and testing sets.

Step 6: Standardize numerical features.

Step 7: Initialize the DNDF model: combine neural network layers with decision forest.

Step 8: Train the DNDF model on training data.

Step 9: Predict outcomes on testing data.

Step 10: Evaluate performance: accuracy, precision, recall, F1-score.

Step 11: Generate confusion matrix and classification report.

Step 12: Measure total execution time.

Step 13: Display accuracy, confusion matrix, classification report, and execution time.

CNN_DNDF

CNN-DNDF Hybrid Model is a deep learning framework that combines the feature extraction power of Convolutional Neural Networks (CNN) with the structured decision-making capabilities of Deep Neural Decision Forests (DNDF). In this project, the hybrid model was applied to tabular sleep health data to predict sleep disorders more accurately. The CNN layers first extracted hierarchical and subtle feature patterns from variables such as sleep duration, caffeine intake, screen time, stress levels, and physical activity. These high-level feature representations were then passed to the DNDF component, which structured the information into interpretable decision paths to classify sleep disorders effectively. By integrating both CNN and DNDF, the hybrid model captured complex interactions between lifestyle, physiological, and behavioral factors while maintaining robustness and generalization on unseen data. This synergy improved classification accuracy, enhanced model interpretability, and supported more precise predictions across different types of sleep disorders in real-world scenarios.

Pseudocode for CNN_DNDF Classification

Input: Sleep Disorder Dataset

Output: Model Accuracy, Classification Report, Confusion Matrix, and Execution Time

Step 1: Import libraries for data handling, deep learning, and evaluation.

Step 2: Load the dataset and check its structure.

Step 3: Preprocess data: handle missing values, remove duplicates, encode categorical features, and reshape if needed.

Step 4: Define input features and target variable.

Step 5: Split data into training and testing sets.

Step 6: Standardize or normalize numerical features.

Step 7: Build CNN layers to extract features from input data.

Step 8: Connect CNN output to DNDF layers to form the hybrid model.

Step 9: Train the hybrid CNN-DNDF model on training data.

Step 10: Predict outcomes on testing data.

Step 11: Evaluate performance: accuracy, precision, recall, F1-score.

Step 12: Generate confusion matrix and classification report.

Step 13: Measure total execution time.

Step 14: Display accuracy, confusion matrix, classification report, and execution time.

CHAPTER V

EXPERIMENTAL RESULTS

In this project, sleep disorder prediction was performed in two stages. The first stage identifies whether a person has a sleep disorder or not. The second stage classifies the type of disorder into five categories: Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, and Narcolepsy. The models were evaluated using metrics such as accuracy, precision, recall, and F1-score to assess their performance in both stages. These results provide insights into the models' effectiveness in detecting the presence of sleep disorders and correctly classifying their types.

Dataset Overview

The dataset used for this project contained critical attributes capturing the lifestyle, physiological, and sleep habits of participants. The data was collected to analyze factors affecting sleep quality and to predict the presence and type of sleep disorders. These attributes include:

- **Demographics:** Age, Gender, Height, Weight. These features help understand the basic characteristics of participants and how physiological factors like body mass may contribute to sleep disorders.
- **Lifestyle Factors:** Caffeine Intake, Smoking Status, Alcoholic Drinks per Week, Screen Time, Physical Activity. These features reflect participants' daily habits, which can significantly influence sleep patterns and overall sleep health.
- **Sleep Patterns:** Sleep Duration, Time to Fall Asleep, Night Awakenings, Sleep Quality, Daytime Sleepiness, Urge to Move Legs, Daytime Sleep Attacks, Irregular Work Patterns, Unusual Sleep Behaviors. These variables capture both the quantity and quality of sleep, as well as behavioral symptoms that may indicate specific sleep disorders.
- **Derived Features:** Body Mass Index (BMI) was calculated from height and weight to assess obesity-related risks, which are commonly associated with conditions such as Sleep Apnea.

Target Variables

- **Stage 1:** has_sleep_disorder classifies whether a participant has a sleep disorder:
  - **Yes:** Participant has a sleep disorder
  - **No:** Participant does not have a sleep disorder
- **Stage 2:** sleep_disorder_type classifies the type of disorder among participants identified with sleep disorders. The categories include:
  - Sleep Apnea
  - Insomnia
  - Restless Leg Syndrome
  - Parasomnias
  - Narcolepsy
  - No Disorder (for participants without sleep disorder)

Encoding Categorical Variables

- **Gender** was standardized to title case (Male, Female) for consistency.
- Binary fields such as daytime_sleep_attacks, irregular_work_patterns, and unusual_behaviors were encoded as Yes/No. This allows the model to process behavioral features effectively.

Handling Missing Values

- Numeric columns were imputed using mean values to maintain dataset integrity.
- Text-based health fields such as chronic or mental conditions were filled with "NONE" to indicate absence of specific data.
- Binary fields were filled with "No" for missing values, assuming the participant did not report these behaviors.

Feature Engineering

- **BMI:** Calculated as weight_kg / (height_cm / 100)^2 to evaluate obesity and its impact on sleep disorders.
- **Risk Scoring:** A cumulative score based on multiple behavioral and physiological factors was computed to detect potential sleep disorders.

Sleep Disorder Counts

- **Stage 1:** Classifies participants as either having a sleep disorder (Yes) or not (No).
- **Stage 2:** Classifies participants into specific types of sleep disorders: Sleep Apnea, Insomnia, Restless Leg Syndrome, Parasomnias, Narcolepsy, or No Disorder.

This dataset provides a comprehensive view of participant sleep habits, lifestyle, and physiological factors, enabling a detailed analysis of sleep disorders and supporting the development of predictive models for sleep health.

EDA

Categorical Analysis

- Count plots were used to visualize the distribution of categorical variables such as Sleep Disorder Type and other demographic features.
- These charts revealed the imbalance among different disorder categories, with a majority of participants showing no disorder compared to other types like Sleep Apnea or Insomnia.
- Such analysis helps in understanding population trends and guides the model to handle class imbalance effectively during prediction

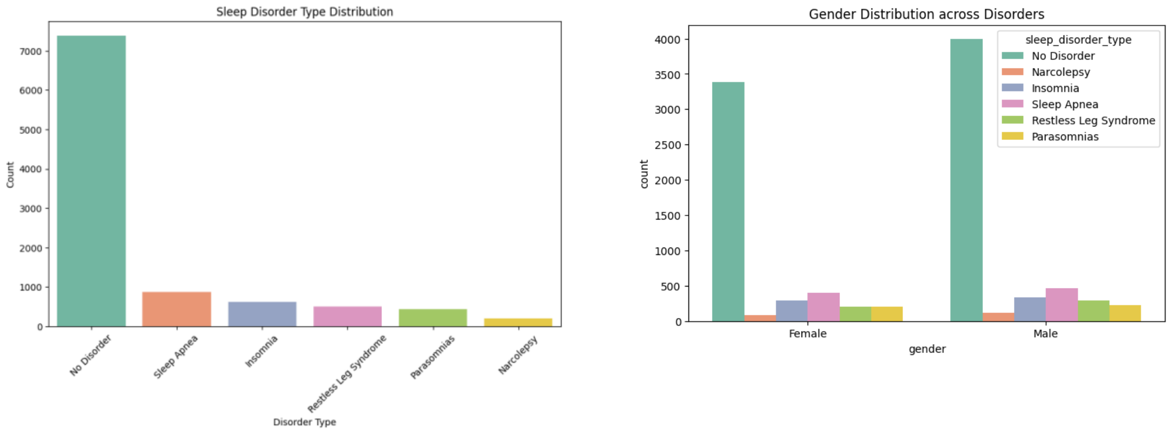

Figure 10. Distribution of sleep disorder

Distributional Analysis

- Lowest Sleep Duration: The Sleep Apnea group has the lowest median sleep duration (around 5.5 hours), significantly lower than all other groups.
- Most Common Duration: Most groups (No Disorder, Narcolepsy, RLS, Parasomnias) have a median sleep duration centered between 6.8 and 7.5 hours.
- High Variability: The "No Disorder" group has the widest range of very short and very long sleep times (outliers), indicating a high diversity in the general population's sleep habits.

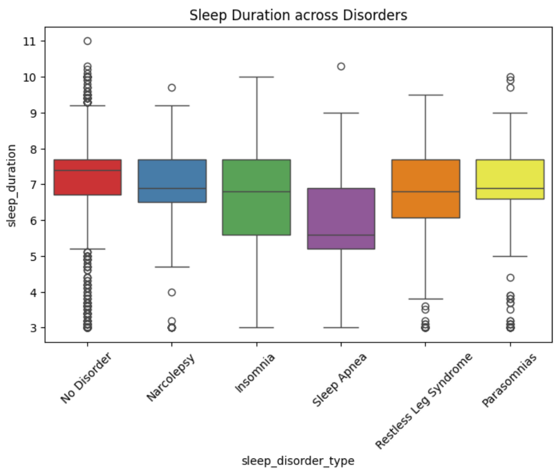

Figure 11. Distribution of sleep duration across disorder

Correlation Heatmap Analysis

- Strong Correlations: A high positive correlation between Weight and BMI (r = 0.81) and a negative correlation between Height and BMI (r = - 0.39), reflecting BMI's dependency on weight and height.

- Weak Relationships: Most sleep-related variables show negligible correlation with demographic and lifestyle factors (age, caffeine, alcohol, screen time), indicating minimal linear influence.

- Independent Sleep Metrics: Sleep variables such as sleep duration, night awakenings, and daytime sleepiness are largely independent, showing very weak internal correlations.

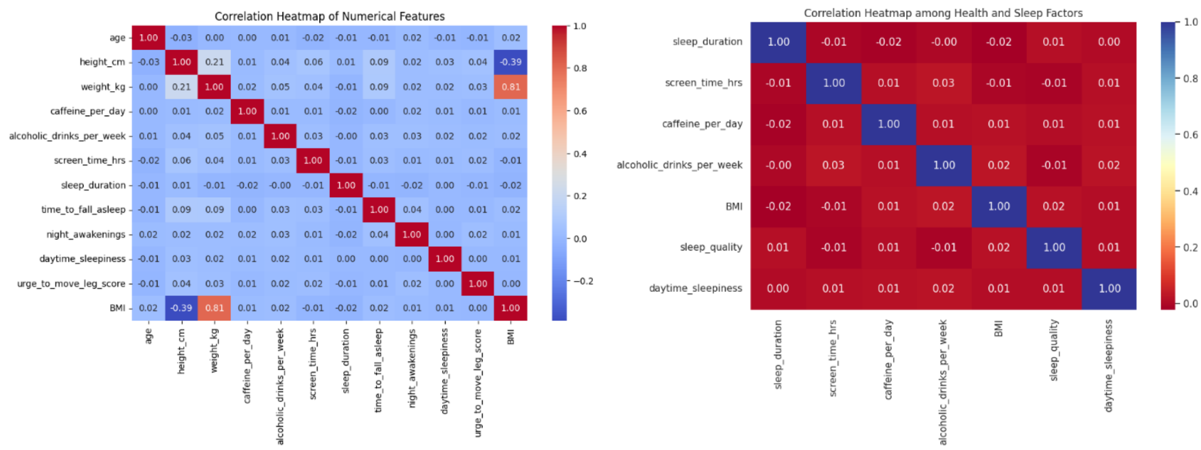

Figure 12. Correlation Heatmap

Distribution Analysis of Daytime Sleepiness

- Multimodal Pattern: The histogram displays multiple peaks, with the highest at score 3, followed by 1 and a smaller peak at 10, reflecting varied levels of daytime sleepiness among individuals.
- Prevalence of Low to Moderate Sleepiness: Most participants reported low to moderate sleepiness levels (scores 1–4), forming the majority in the dataset.
- High Sleepiness Subgroup: A noticeable smaller group reported extreme sleepiness (score 10), potentially representing individuals with underlying sleep-related issues or disorders.

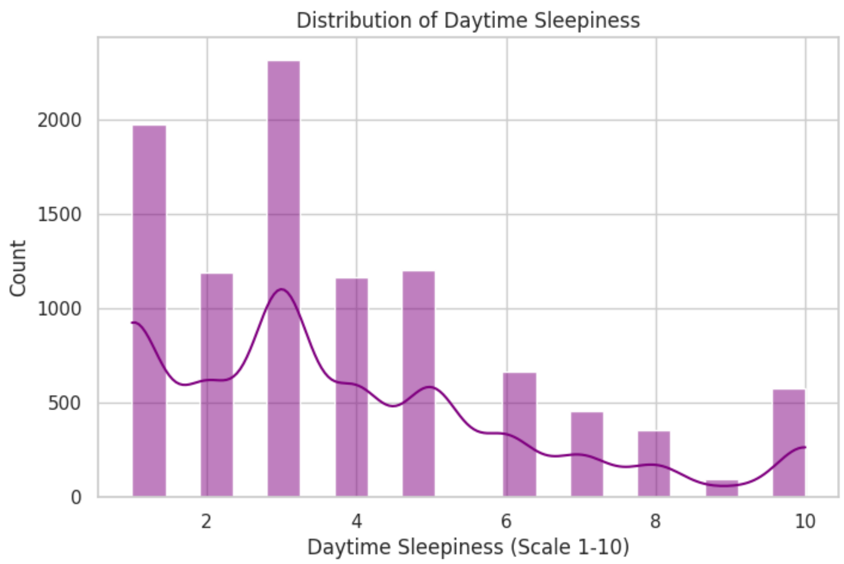

Figure 13. Distribution of daytime sleepiness

Logistic Regression

The Logistic Regression model was implemented in two stages to predict sleep disorder occurrence and classify specific disorder types. In Stage 1, the model achieved an accuracy of 90.30%, demonstrating strong performance in distinguishing between individuals with and without sleep disorders, supported by high precision (0.92) and recall (0.95) for the non-disorder class. In Stage 2, the model obtained an accuracy of 84.10%, effectively identifying multiple disorder categories with a weighted F1-score of 0.83. Although performance varied across individual classes, the results indicate that Logistic Regression provides a reliable baseline model capable of capturing essential patterns in the dataset for both detection and classification of sleep disorders.

Random Forest

The Random Forest model demonstrated strong performance in predicting sleep disorders across both stages. In Stage 1, it achieved an accuracy of 97.75%, with high precision and recall for both the non-disorder and disorder classes, indicating reliable discrimination between healthy and affected individuals. In Stage 2, which classified specific disorder types, the model achieved 95.35% accuracy, showing slightly lower performance for less represented classes but maintaining strong overall predictive capability. The overall accuracy across both stages was 96.15%, highlighting Random Forest's robustness and suitability for handling complex patterns in sleep disorder data.

SVM

The Support Vector Machine (SVM) model achieved moderate performance in predicting sleep disorders. In Stage 1, it reached an accuracy of 88.70%, with strong recall for the non-disorder class but lower performance for the disorder class, indicating some difficulty in distinguishing affected individuals. In Stage 2, which involved multi-class classification of specific sleep disorders, the accuracy dropped to 78.05%, reflecting challenges in handling imbalanced classes and complex feature interactions. The overall accuracy across both stages was 79.25%, suggesting that while SVM can be used as a baseline, ensemble or hybrid models may offer more reliable performance for sleep disorder prediction.

Gradient Boosting

The Gradient Boosting model demonstrated high predictive performance for sleep disorder classification. In Stage 1, it achieved an accuracy of 98.10%, effectively distinguishing between individuals with and without sleep disorders, supported by strong precision and recall for both classes. In Stage 2, which involved multi-class classification of specific sleep disorder types, the model maintained a high accuracy of 98.25%, showing reliable predictions even for less frequent classes. The overall accuracy across both stages was 98.20%, highlighting Gradient Boosting's robustness and effectiveness in capturing complex patterns in the dataset.

XGBoost

The XGBoost model exhibited the highest performance among the individual models for sleep disorder prediction. In Stage 1, it achieved an accuracy of 99.30%, accurately distinguishing between individuals with and without sleep disorders with near-perfect precision and recall. In Stage 2, which involved multi-class classification of specific disorder types, the model maintained an excellent accuracy of 98.65%, demonstrating its capability to handle complex and imbalanced class distributions. The overall accuracy across both stages was 98.75%, confirming XGBoost as a highly effective and reliable model for sleep disorder prediction.

ML Hybrid (Random Forest + XGBoost)

The ML hybrid model, combining Random Forest and XGBoost, achieved the best overall performance for sleep disorder prediction. In Stage 1, it reached an accuracy of 99.30%, effectively distinguishing between healthy and affected individuals with near-perfect precision and recall. In Stage 2, which classified specific sleep disorder types, the model maintained an accuracy of 98.70%, demonstrating robust predictions even for less represented classes. The overall accuracy across both stages was 98.95%, highlighting the advantage of hybrid ensemble methods in capturing complex patterns and improving reliability over individual models.

MLP

The Multi-Layer Perceptron (MLP) model showed strong performance in predicting sleep disorders. In **Stage 1**, it achieved an accuracy of **96.30%**, effectively distinguishing between individuals with and without sleep disorders with high precision and recall. In **Stage 2**, which involved multi-class classification of specific disorder types, the model maintained an accuracy of 93.25%, performing well across most classes while showing slightly lower performance for less frequent categories. The overall accuracy across both stages was 95.00%, indicating that MLP is a reliable deep learning approach for modeling complex patterns in sleep disorder data.

CNN

The Convolutional Neural Network (CNN) model demonstrated strong performance in predicting sleep disorders. In Stage 1, it achieved an accuracy of 96.05%, effectively distinguishing between individuals with and without sleep disorders with high precision and recall. In Stage 2, which involved multi-class classification of specific disorder types, the model achieved 92.75% accuracy, performing well across most classes while showing slightly lower results for less frequent categories. Overall, the CNN model provided reliable predictions, highlighting its ability to capture complex patterns in sleep disorder data.

DNDF

The Deep Neural Decision Forest (DNDF) model showed moderate performance in predicting sleep disorders. In Stage 1, it achieved an accuracy of 82.50%, indicating reasonable discrimination between individuals with and without sleep disorders. In Stage 2, which involved multi-class classification of specific disorder types, the accuracy decreased to 75.60%, with lower performance for less frequent classes. Overall, DNDF provided acceptable predictions but was outperformed by hybrid and deep learning approaches, suggesting that additional feature extraction or model combination could improve its reliability.

CNN + DNDF Hybrid

The CNN + DNDF hybrid model combined the feature-extraction capabilities of CNN with the structured decision-making of DNDF to improve predictive performance. In Stage 1, it achieved an accuracy of 85.65%, showing moderate effectiveness in distinguishing between individuals with and without sleep disorders. In Stage 2, which involved multi-class classification of specific disorder types, the model achieved 80.35% accuracy, demonstrating better performance than DNDF alone but lower than standalone CNN and other hybrid ML models. Overall, the CNN + DNDF hybrid provided improved predictions over DNDF, highlighting the benefits of integrating deep learning with decision forest structures, though it did not outperform traditional ML hybrids.

Overall Accuracy

Model	Stage 1 Accuracy	Stage 2 Accuracy	Overall Accuracy
RandomForest	0.9775	0.9535	0.9615
XGBoost	0.9930	0.9865	0.9875
GradientBoosting	0.9810	0.9825	0.9820
LogisticRegression	0.9030	0.8410	0.8580
SVM	0.8870	0.7805	0.7925
ML_Hybrid_RF_XGB	0.9930	0.9870	0.9895
MLP	0.9630	0.9325	0.9500
CNN	0.9605	0.9275	0.9275
DNDF	0.8250	0.7560	0.7670
CNN-DNDF	0.8565	0.8035	0.7945

Figure 14. Overall accuracy

Model Comparison and Insights

The hybrid approach demonstrated the best predictive performance across both stages of sleep disorder classification. While individual or simpler models provided reasonable results, the hybrid method was more reliable in capturing complex patterns and handling multi-class predictions. This indicates that combining multiple learning algorithms enhances accuracy and robustness, making hybrid models the most effective choice for this project.

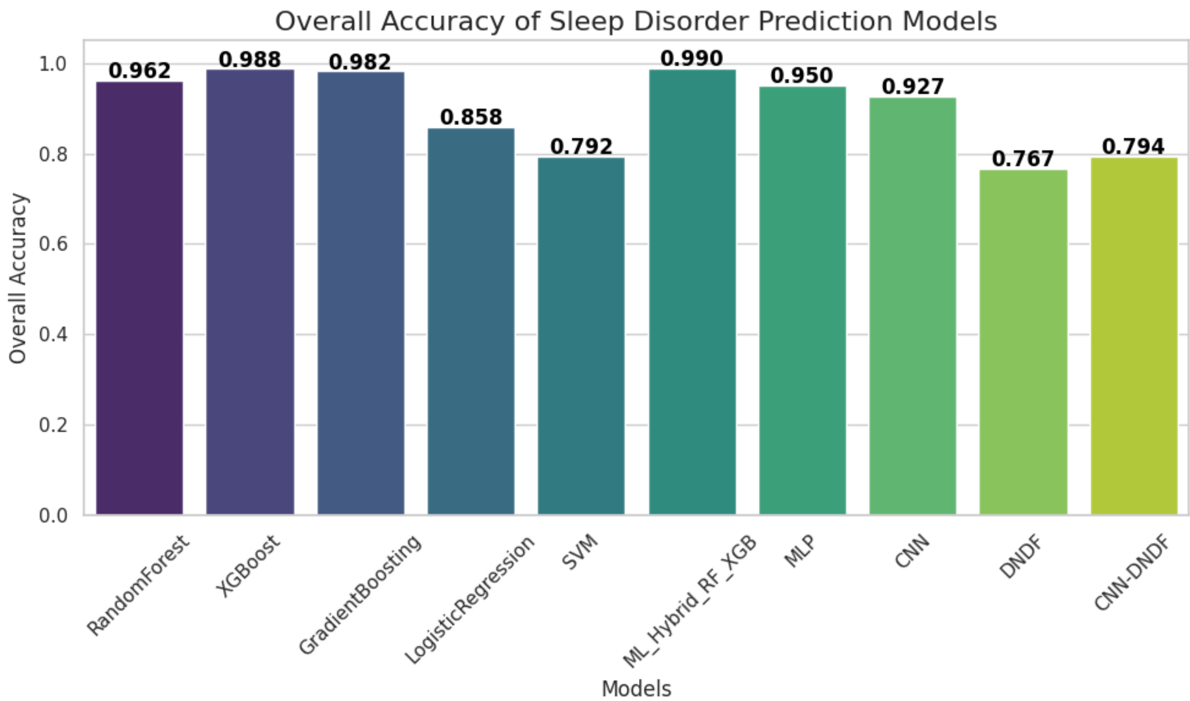

Figure 15. comparison of model accuracy

DISCUSSION OF FINDINGS

Table 1. Model performance on different scenarios

MODELS	STRENGTHS	WEAKNESS
Random Forest	• Handles large datasets and high-dimensional data effectively. • Reduces overfitting by averaging predictions of multiple decision trees. • Provides high accuracy and stability. • Highlights feature importance, aiding in the identification of critical factors.	• Computationally intensive and slow, especially with large datasets. • Complex and less interpretable compared to individual decision trees. • Difficult to explain model decisions to non-technical users.
XGBoost	• Known for speed, scalability, and superior predictive performance. • Handles missing values efficiently. • Incorporates regularization to minimize overfitting. • Ideal for complex classification problems with large datasets.	• Requires careful hyperparameter tuning to prevent overfitting. • Computationally demanding and resource-intensive. • Complex and harder to interpret, limiting transparency.
Logistic Regression	• Simple, easy to implement, and interpretable. • Ideal for binary classification tasks. • Computationally efficient	• Struggles with capturing non-linear relationships in the data. • Performs poorly with complex patterns. • Sensitive to outliers

Gradient Boosting	• Gradient Boosting gives high accuracy by reducing errors step by step. • It handles complex patterns and relationships in data effectively. • It works well for both classification and regression problems.	• It can overfit if not tuned properly. • Training takes more time compared to simpler models. • It needs careful tuning of parameters for the best performance.
Support Vector Machine (SVM)	• Performs well in high-dimensional spaces. • Effectively separates classes using a hyperplane with maximum margin. • Versatile due to various kernel functions for linear and non-linear classification.Suitable for complex datasets with clear class boundaries.	• Computationally intensive, especially with large datasets. • Requires significant resources and time for training. • Sensitive to parameter tuning; choosing the right kernel and hyperparameters is challenging. • Less interpretable, making it difficult to explain decision boundaries to end-users.
Multi layer Perceptron (MLP)	• MLP can learn complex, non-linear relationships between features. • It works well with large datasets and multiple input variables. • It can generalize well after proper training and tuning.	• It requires more computational power and time to train. • It may overfit if not regularized properly. • It works like a “black box,” making interpretation difficult.

Convoluti\ional neural network (CNN)	• CNN automatically extracts important features without manual selection. • It captures spatial and hierarchical patterns effectively. • It performs well with large and complex datasets.	• It requires a large amount of data for effective training. • It demands high computational resources • It can be difficult to interpret or explain model decisions.
Deep neural decision forest (DNDF)	• DNDF combines deep learning and decision forests for better accuracy. • It captures both linear and non-linear relationships in the data. • It performs well on complex datasets with mixed feature types.	• It requires high computational power and longer training time. • It may overfit if not properly tuned. • The model structure can be complex and difficult to interpret.

CHAPTER VI

CONCLUSION

The experimental analysis of various machine learning and deep learning models for sleep disorder prediction demonstrated that different architectures excel in different aspects of classification. Traditional ensemble models such as Random Forest, Gradient Boosting, and XGBoost achieved high accuracy and reliability, with XGBoost standing out as the best individual performer due to its ability to handle complex feature interactions and imbalanced data effectively. The ML Hybrid model (Random Forest + XGBoost) surpassed all other models, achieving the highest overall accuracy and showcasing the strength of combining ensemble techniques for better generalization and stability.

Deep learning models like the MLP and CNN also delivered strong results, effectively capturing nonlinear relationships and subtle sleep pattern variations, proving suitable for complex data structures. However, models such as DNDF and the CNN DNDF hybrid showed moderate results, performing better than simpler classifiers but not reaching the precision of ensemble-based methods suggesting that further tuning or feature enhancement could improve their performance. The experiments highlight that hybrid ensemble approaches offer the most robust and consistent performance for sleep disorder prediction. These models can effectively classify both the presence of a disorder and the specific disorder type, making them highly valuable for early detection systems, personalized healthcare monitoring, and improving sleep quality assessment in real-world applications.

REFERENCE

[1]. Alshammari, T. S. (2023). “Applying Machine Learning Algorithms for the Classification of Sleep Disorders”. Department of Information and Computer Science, University of Ha’il, Saudi Arabia.

[2]. Zhang, W., Zhou, N., & Li, J. (2025). “Dynamic Impact of the Sleep Disorder, Depression and Anxiety on the Cognitive Function in the First-Episode Depressive Patients”. Psychology Research and Behavior Management.

[3]. Monowar, M. M., Nobel, S. M. N., Afroj, M., Hamid, M. A., Uddin, M. Z., Kabir, M. M., & Mridha, M. F. (2025). “Advanced sleep disorder detection using multi-layered ensemble learning and advanced data balancing techniques”. Frontiers in Artificial Intelligence.

[4]. Rahman, M. A., Jahan, I., Islam, M., Jabid, T., Ali, M. S., Rashid, M. R. A., Islam, M. M., Ferdaus, M. H., Rasel, M. M. K., Jahan, M. R., Rimy, T. A., Talukder, A. S., Sharmin, S., & Rimi, T. A. (2025). “Improving sleep disorder diagnosis through optimized machine learning approaches”. IEEE Access.

[5]. Wara, T. U., Fahad, A. H., Das, A. S., & Shawon, M. M. H. (2024). “A Systematic Review on Sleep Stage Classification and Sleep Disorder Detection Using Artificial Intelligence”.

[6]. Panda, N. R., Paramanik, S., Raut, P. K., & Bhuyan, R. (2025). “Prediction of sleep disorders using novel decision support neutrosophic-based machine learning models”. Neutrosophic Sets and Systems.

[7]. El-Kenawy, E. M., Ibrahim, A., Abdelhamid, A. A., Khodadadi, N., Abualigah, L., & Eid, M. M. (2024). “Predicting sleep disorders: Leveraging sleep health and lifestyle data with Dipper Throated Optimization Algorithm for feature selection and Logistic Regression for classification”. Computational Journal of Mathematical and Statistical Sciences.

[8]. Sari, H. K., Shoelarta, S., Pratama, T. O., Sajida, G. N., Krista, G. M., Ferawati, Y. F., & Taufiqurrahim, T. (2025). “Machine Learning-Based Prediction of Sleep Disorders from Lifestyle and Physiological Data”: A Cross-Occupational Study. Journal of Technology.

[9]. Li, M., Zhang, Y., Li, J., Ma, L., et al. (2020). “Sleep Disturbances and Depression Risk: A Meta-Analysis of Longitudinal Studies”. BMC Psychiatry.

[10]. Hidayat, I. A. (2023). “Classification of Sleep Disorders Using Random Forest on Sleep Health and Lifestyle Dataset”. Institut Teknologi Telkom Purwokerto, Indonesia.

[11]. Alom, M. S., Jeba, S. M., Debnath, A., Aurpa, T. T., & Siddiqua, R. (2025). “Enhancing Sleep Disorder Diagnosis with a Machine Learning Approach Using Ensemble Neural Networks”. International Journal of Artificial Intelligence in Healthcare.