Sleep Disorder Prediction Using Advanced Machine Learning Techniques

Mrs P Sarala[1], Charitha Cherukuri[2], Kusuma Kala Challa[3], Yaswanthi Angadala[4], Abdul Haq[5]

[1]Assistant Professor, [2,3,4,5]Student
[1,2,3,4,5]Dhanekula Institute of Engineering & Technology, Ganguru, Vijayawada, Andhra Pradesh, India.
[1]saralapulivarthi@gmail.com, [2]charithacherukuri1710@gmail.com, [3]kusumakalac@gmail.com, [4]Yashuangadala7@gmail.com, [5]abdulshakeel241@gmail.com

Abstract
Sleep disorders, such as insomnia, sleep apnea, and restless legs syndrome, have a significant impact on global health because they increase the risk of cardiovascular diseases, cognitive decline, and a decreased quality of life. In addition to being expensive and resource intensive, polysomnography (PSG) and other conventional diagnostic methods require clinical nightly monitoring. This study recommends a machine learning based predictive model as a non-invasive, precise, and reasonably priced alternative to conventional techniques for identifying sleep issues.
By combining supervised learning, ensemble methods, and recurrent neural networks (RNN), the model evaluates a wide range of data sources, including physiological signals, medical histories, sleep patterns, and demographic data. When these algorithms are compared, it can be seen that deep learning models especially those that deal with temporal data achieve excellent predictive accuracy. Future research will concentrate on integrating other data sources, verifying the model with bigger datasets, and investigating real-time applications for home-based and clinical monitoring.

Keywords: prediction of sleep disorders, detection of sleep apnea, classification of insomnia, neural networks (RNN, LSTM), ensemble learning, and evaluation of sleep health.

1. Introduction

Sleep problems impact a large percentage of people worldwide and include a broad spectrum of illnesses that interfere with regular sleep patterns. Significant health issues, such as cardiovascular conditions, cognitive decline, compromised immune systems, and a general deterioration in quality of life, can result from these problems. The limits of conventional diagnostic techniques make it difficult to diagnose sleep problems, despite their high frequency. Patients must spend the night in specialized sleep laboratories for polysomnography (PSG), a resource-intensive and expensive test that is regarded as the gold standard for diagnosing sleep problems. This method not only raises healthcare expenses but also inconveniences patients because it doesn't replicate their normal sleeping environment.
Long wait periods are another consequence of the scarcity of PSG facilities and qualified personnel, which delays diagnosis and therapy. Numerous things can cause sleep disturbances, such as stress, illnesses, lifestyle choices, and underlying neurological or genetic influences. Sleep apnea, narcolepsy, insomnia, restless legs syndrome (RLS), circadian rhythm sleep disorders, and parasomnias are a few of the most

prevalent kinds. Since the symptoms and causes of each of these disorders vary, early discovery and precise diagnosis are crucial for successful treatment. There is a rising demand for new ways that provide accurate, non-invasive, and affordable solutions for the diagnosis of sleep disorders because of the difficulties with standard diagnostic techniques.

2. Literature survey

1. "Machine learning approach for anxiety and sleep disorders analysis during COVID-19 lockdown" is the title of the paper. 740 participants from various demographic groups participated in a survey using the General Anxiety Disorder (GAD-7) and Pittsburgh Sleep Quality Index (PSQI) evaluation measures. Results show that anxiety and sleep disruptions are significantly correlated, with students being especially at risk. Key stresses were discovered by machine learning techniques like Random Forest and K-means clustering, and they included social isolation, online learning challenges, and academic burden. Anxiety levels differed by profession, but occupation did not significantly affect sleep quality, according to statistical analysis using ANOVA and Kruskal-Wallis tests. The study emphasizes the necessity of focused treatments to lessen mental health issues made worse by stress brought on by pandemics.
2. In the IEEE paper "Automatic Detection and Classification of Sleep Disorders Using AILearning Models," a thorough analysis of the use of AI in the identification and classification of different sleep disorders is presented. In order to identify abnormalities suggestive of sleep disorders, the authors created a strong artificial intelligence system that examines polysomnographic data. The system classified disorders like narcolepsy, sleep apnea, and insomnia with excellent accuracy by using machine learning techniques. The study emphasizes how AI-driven methods may improve the accuracy of diagnoses and make early intervention in sleep medicine easier.

Proposed system

The suggested approach uses ensemble learning techniques (AdaBoost, Gradient Boosting) and recurrent neural networks (RNN) to create a machine learning-based predictive model for the classification of sleep disorders. It makes use of both clinical and non-clinical data, such as physiological signals, lifestyle factors, and sleep patterns, to offer a non-invasive and reasonably priced substitute for conventional diagnostic techniques like polysomnography (PSG).

Preprocessing the data, selecting features with PCA, and training the model with hyperparameter optimization (Grid Search & Genetic Algorithm) are important steps. Sequential sleep patterns are captured by RNN (LSTM), and classification accuracy is improved by AdaBoost and Gradient Boosting. The system is appropriate for clinical evaluations and wearable sleep monitoring devices since it guarantees excellent accuracy, scalability, and real-time application.

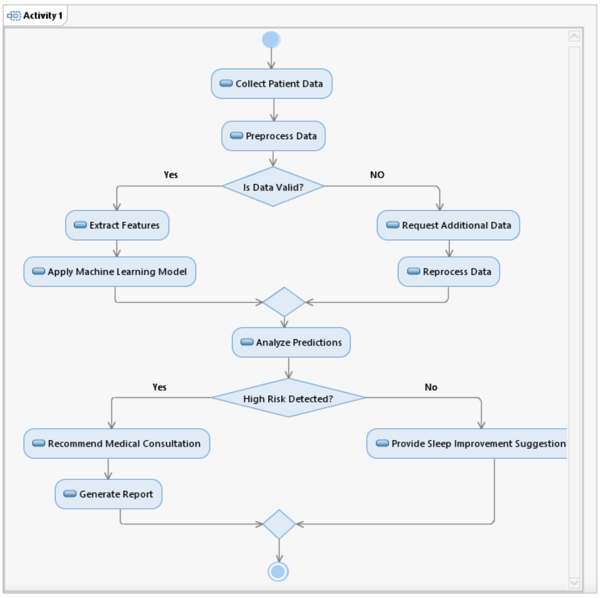
1. Algorithms used

Recurrent neural network

RNN is a particular kind of neural network intended for processing sequential input, including text, audio, and time series. RNNs may remember information from past inputs (memory) and learn temporal patterns because, in contrast to standard neural networks, their connections loop back on themselves. For tasks like language modelling, translation, and sleep problem analysis, this makes them ideal. Newer versions like LSTM and GRU are intended to solve problems like vanishing gradients that RNNs may encounter.

Ensemble learning

Ensemble Learning is a machine learning technique that enhances overall performance by combining the predictions of several models. By combining the advantages of multiple algorithms, ensemble techniques like stacking and boosting (like AdaBoost and Gradient Boosting) lower errors, improve accuracy, and strengthen resistance to noise and overfitting. These techniques produce more accurate predictions than individual models and are particularly useful in challenging tasks like fraud detection, sleep problem categorization, and image recognition.



2. **Working**

For a sleep disorder prediction system, data preparation is essential. Several crucial steps are involved in the process:

- Missing Value Handling: Statistical techniques such as mean or median imputation are used to fill in the gaps in physiological signals (heart rate, oxygen levels, etc.) and sleep patterns.
- Normalisation & Scaling: For consistency, continuous variables such as sleep duration, body movement data, and breathing rate are normalised using Min-Max Scaling.
- Categorical Encoding: One-Hot Encoding is used to translate categorical features (such as age groups, lifestyles, and sleep quality evaluations) into numerical representations.
- Time-Series Processing: A sliding window method records how the sequential sleep data evolves over time.

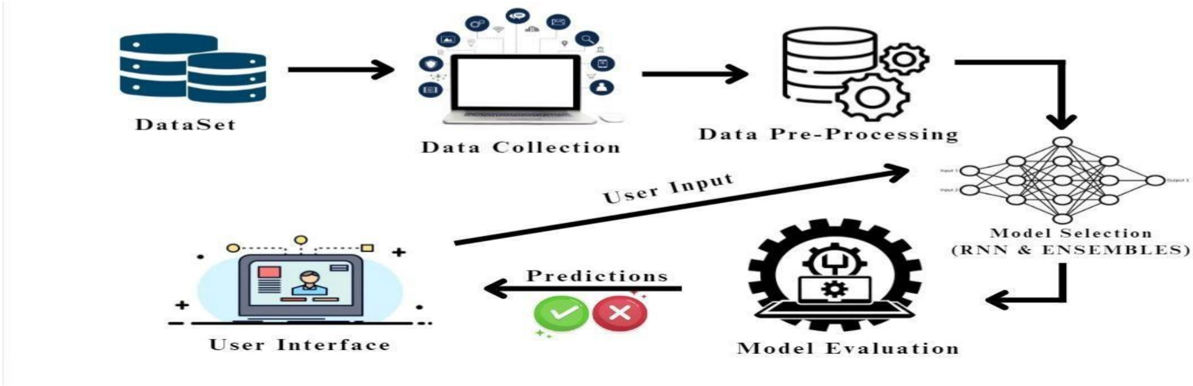
Extraction of Features

Long Short-Term Memory (LSTM), a kind of Recurrent Neural Network (RNN), is used for feature extraction in order to efficiently assess sequential sleep data. Because it can identify long-term dependencies in time-series data, like irregular breathing, heart rate changes, and sleep cycle oscillations, LSTM is especially helpful in forecasting sleep disorders. By understanding intricate correlations between various physiological inputs, this deep learning technique allows the machine to identify possible patterns of sleep disorders.

Data Division

The dataset is divided into training (70%) and testing (30%) sets to guarantee strong prediction performance. This enables the system to be assessed on unseen data and train efficiently. Five-fold cross-validation is used to make sure the model generalizes properly and avoid overfitting.

Architecture



Predicting Sleep Disorders using RNN and Ensemble Learning

To enhance predictive performance, the system combines ensemble learning methods with deep learning: Learns time-series patterns from sleep data and extracts useful features for prediction using **Recurrent Neural Networks (RNN) with LSTM.**

Adaptive Boosting (AdaBoost): Combines several weak learners and dynamically modifies weights in response to errors to improve prediction accuracy.

Gradient Boosting: This technique is useful for managing non-linear interactions in sleep data since it iteratively builds prediction models by lowering residual errors at each stage. AdaBoost and Gradient Boosting models receive the LSTM-extracted features, integrating ensemble learning and deep learning to increase prediction accuracy.

Interaction with End Users

Through the system's user-friendly interface, users can:You can manually enter sleep-related parameters or upload data from wearable sleep tracking devices. Get individualized sleep disorder predictions that assist users in determining their likelihood of developing diseases such as restless legs syndrome, sleep apnea, and insomnia. Obtain suggestions for enhancing sleep health based on anticipated outcomes.

Warning Signs of Possible Sleep Issues

When a high likelihood of a sleep issue is identified, the system notifies users. Users get alerts about the severity of their sleep problems and possible next measures, like speaking with a healthcare provider, based on anticipated results.

Model Performance Metrics

Based on accuracy, precision, recall, and F1-score, the performance metrics contrast Adaboost, Gradient Boosting, and LSTM. With the highest F1-score (95.1%) and the best accuracy of 96.0%, gradient boosting ensures a balanced precision-recall tradeoff. Adaboost comes in second with a little lower recall of 94.7% accuracy and a 93.3% F1-score. Because of its lesser precision (84.8%), which results in more false positives, LSTM has the lowest accuracy (89.3%) and F1-score (87.5%). Adaboost is a good substitute for Gradient Boosting, which is the best model overall. Because of its poorer performance, LSTM might not be the best option. To get the most accurate and dependable predictions, Gradient Boosting should be used in accordance with these measures.

Model Performance Metrics

	Accuracy	Precision	Recall	F1 Score
Lstm	89.3%	84.8%	90.3%	87.5%
Gradient Boosting	96.0%	96.7%	93.5%	95.1%
Adaboost	94.7%	96.6%	90.3%	93.3%

Metrics Explanation:

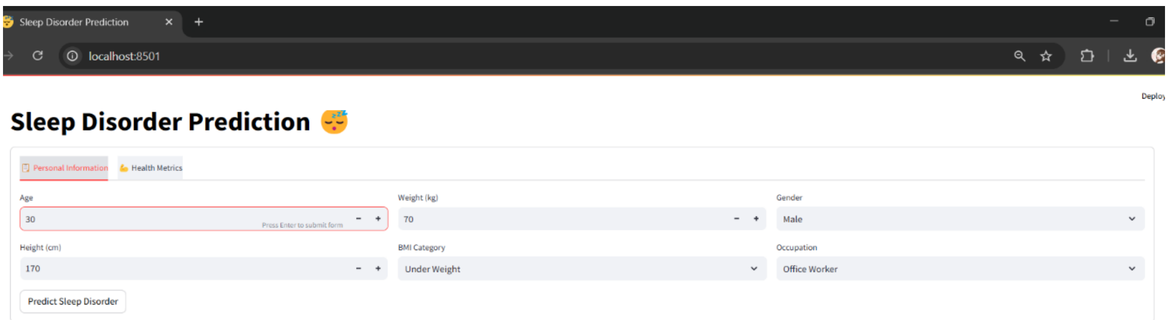
- **Precision:** Percentage of correct positive predictions
- **Recall:** Percentage of actual positives correctly identified
- **F1 Score:** Balance between precision and recall
- **Accuracy:** Overall correct predictions

3. **Advantages of proposed system**

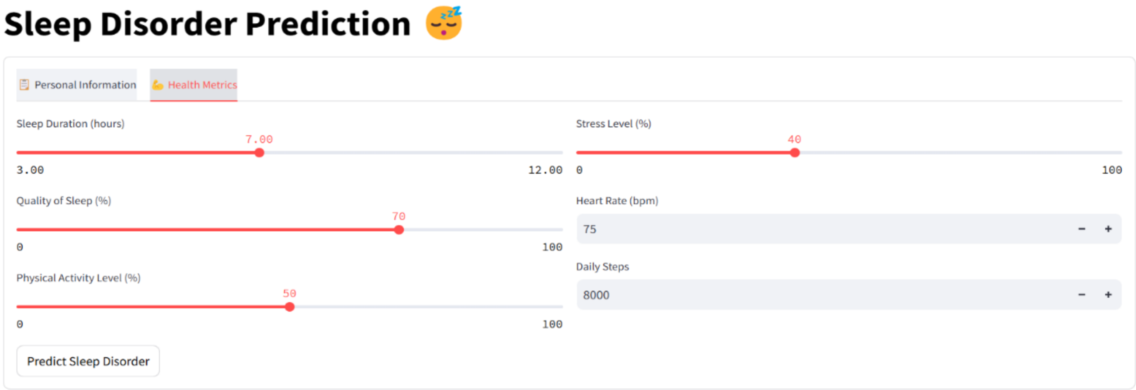
- High Accuracy: Makes accurate predictions by combining ensemble techniques with deep learning.
- Efficient Sequential Data Processing: Long-term sleep patterns are captured by LSTM for improved analysis.
- Generalization & Robustness: AdaBoost and Gradient Boosting enhance dependability by minimizing overfitting.
- Cost-effective & Non-Invasive: Removes the need for pricey sleep lab testing.Real-time sleep disorder predictions are made possible by automated and quick detection.
- Scalable and User-Friendly: Capable of integrating with mobile apps and wearable technology. Preventive care and early diagnosis assist users in taking prompt action to avoid serious health problems.

Output screens

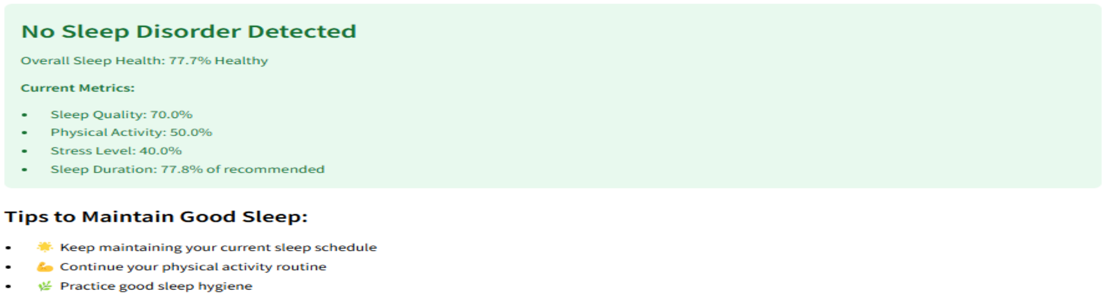
1.



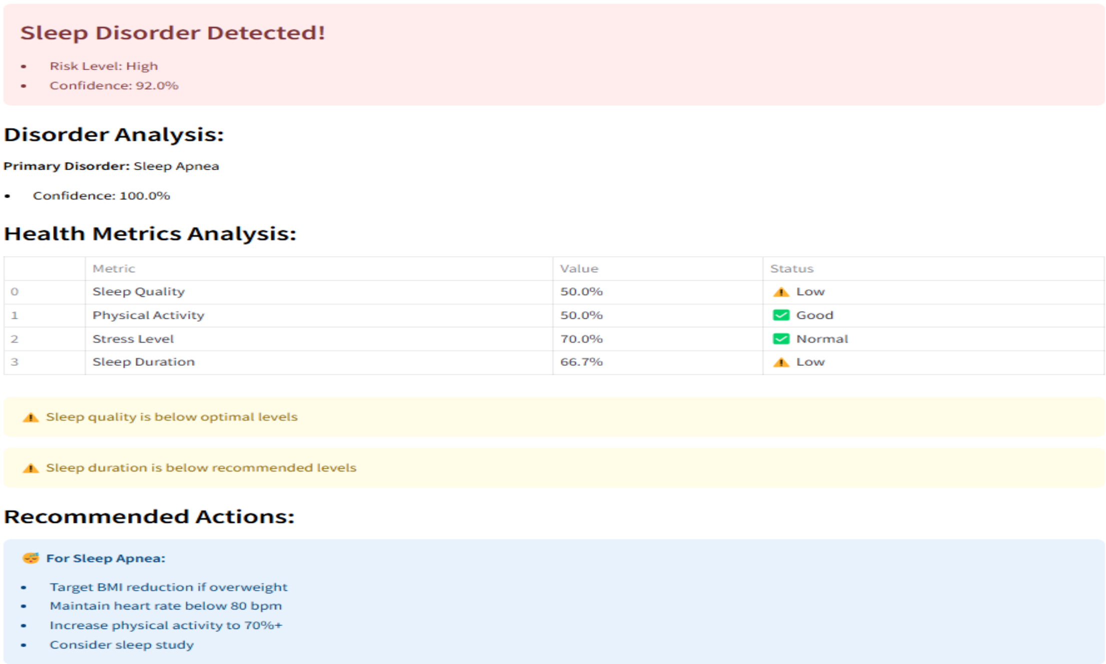
2.



3.



4.



3. Future scope

AI-driven customized sleep schedules, cloud-based smartphone apps, and realtime monitoring using wearable technology are all part of the future of sleep problem prediction. While multi-modal data fusion (EEG, ECG, lifestyle factors) will boost insights, advanced deep learning models (CNN + LSTM, Transformers) will improve accuracy. Explainable AI (XAI), telemedicine integration, and automated alarms for serious illnesses would increase the system's dependability for medical professionals. Sleep health monitoring will become more broadly accessible and efficient by extending detection to additional sleep disorders supporting different languages, and guaranteeing worldwide accessibility.

4. Conclusion

In order to identify sleep problems, the Sleep Disorder Prediction project examines user health metrics. Accuracy, precision, recall, and F1-score were used to assess three models: LSTM, Gradient Boosting, and Adaboost. With an F1-score of 95.1% and an accuracy of 96.0%, gradient boosting was the most dependable model. While LSTM had the lowest accuracy (89.3%) and F1-score (87.5%), primarily because of its lower precision (84.8%), which resulted in more false positives, Adaboost came in second (94.7% accuracy, 93.3%F1-score). Gradient Boosting is the optimum option for deployment based on these measures, guaranteeing precise and effective sleep problem identification. The technology helps users enhance the quality of their sleep by offering insightful analysis and tailored suggestions.

References

1. Xplore," Insomnia, Sleep Apnea, Serenity: Unveiling Sleep Disorders with Machine Learning", https://ieeexplore.ieee.org/document/10468035
2. SPRINGER," Neurobehavioral, health, and safety consequences associated with shift work in safety-sensitive professions", https://link.springer.com/article/10.1007/s11910-009-0024-7
3. MDPI, "Utilizing Multi-Class Classification Methods for Automated Sleep Disorder Prediction", https://www.mdpi.com/2078-2489/15/8/426
4. IEEE Xplore, " Detection and Prediction of Sleep Disorders by Covert Bed-Integrated RF Sensors", https://ieeexplore.ieee.org/document/9913659
5. SPRINGER, "An intelligent deep feature based metabolism syndrome prediction system for sleep disorder diseases", https://link.springer.com/article/10.1007/s11042-023-17296-4
6. IEEE Xplore," Development of methods for sleep disordered breathing to identify phenotypes", https://ieeexplore.ieee.org/document/8037185
7. Hajak G, Rodenbeck A, Voderholzer U, et al. Doxepin in the treatment of primary insomnia - a placebo-controlled, double-blind, polysomnographic study. J Clin Psychiatry. 2001;62(6):453–63. [PubMed]
8. IEEE Xplore," Performance Evaluation of the Circadia Contactless Breathing Monitor and Sleep Analysis Algorithm for Sleep Stage Classification", https://ieeexplore.ieee.org/document/9175419
9. ATS Journals, "The Epidemiology of Adult Obstructive Sleep Apnea", https://www.atsjournals.org/doi/full/10.1513/pats.200709-155MG
10. Biomedical and Pharmacology Journal, "Relationship between sleep and hypertension: A Narrative Review",
11. SPRINGER, "Predicting Sleep Disorders for Improved Health Care—A Comprehensive Study", https://link.springer.com/chapter/10.1007/978-981-97-3588-4_23