International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Machine Learning Approaches for Sleep Disorder Prediction

Malathi. D[1], Selvapriya. S[2]

Assistant Professor [1], Student [2]

Department of Computer Science and Engineering, Tagore Institute Of Engineering And Technology, Deviyakurichi, Salem, Tamil Nadu, India

ABSTRACT

Classifying sleep disorders is essential for enhancing human well-being. Conditions such as sleep disorders and apnoea can significantly impact overall health. Expert-driven sleep-stage classification is a complex task that is susceptible to errors. Developing precise machine learning algorithms for identifying sleep disorders involves thorough analysis, monitoring, and diagnosis of sleep patterns. This paper compares deep learning methods with traditional machine learning algorithms in the context of sleep disorder classification.This study introduces an optimized approach for sleep disorder classification, using the publicly available Sleep Health and Lifestyle Dataset for model evaluation. The optimization process employs a genetic algorithm to adjust the parameters of various machine learning techniques. A thorough assessment and comparison of the proposed method with leading machine learning algorithms for sleep disorder classification are provided.The research utilizes the comprehensive Sleep Health and Lifestyle Dataset, containing 16,000 records with 13 distinct features. Among the two primary models used in this project is the Random Forest Classifier, an ensemble learning approach that constructs a predictive model by progressively adding decision trees. By optimizing the loss function step-by-step, the model effectively reduces prediction errors, achieving 98% accuracy during training and 96% accuracy in testing. These impressive results highlight the model's ability to accurately classify sleep disorders.

Keywords – Classification of Sleep Disorders , Random Forest Classifier

I.INTRODUCTION

Sleep is an essential physiological function that plays a critical role in maintaining physical and mental health. It aids in strengthening the body, consolidating the brain, and enhancing memory retention. The quality of sleep significantly impacts cognitive functions, particularly among vulnerable groups like children and older drivers, who are at an increased risk of accidents. Chronic sleep deprivation can lead to severe health issues, including heart disease, diabetes, and obesity. Evaluating sleep quality often involves polysomnography (PSG), which requires manual interpretation by medical professionals. This process is time-consuming, prone to human error, and can lead to inconsistent results. In recent years, there has been a growing focus on automating the sleep-stage classification process to overcome these challenges. Sleep is classified into five stages: wakefulness, N1, N2, N3, and rapid eye movement (REM). Each stage serves a distinct purpose. Wakefulness involves alertness and fast brain waves, while N1 represents the lightest sleep stage characterized by slower brain waves and muscle relaxation. N2 and N3 are progressively deeper stages of sleep, with N3 being the deepest, making it difficult to wake individuals. REM sleep features rapid eye movements and brain activity similar to wakefulness. PSG enables the observation of these stages by recording EEG and ECG signals, providing valuable insights into the brain and body's activity during sleep. Despite its importance, the manual classification of sleep stages poses significant challenges, necessitating innovative automated solutions. Machine learning and deep learning techniques have emerged as promising approaches to address the limitations of manual sleep-stage classification. Traditional machine learning algorithms (MLAs) are effective for relatively small datasets, offering faster implementation and simplicity. These algorithms rely on manual feature engineering to extract characteristics such as signal entropy and energy for sleep-stage classification. Deep learning, on the other hand, employs neural networks that automatically learn complex patterns from data, eliminating the need for manual feature engineering. This makes deep learning particularly suitable for tasks involving large datasets and intricate features, positioning it as a viable replacement for traditional methods. The integration of machine learning techniques into sleep disorder diagnosis presents opportunities to improve the accuracy and efficiency of medical evaluations. This study utilizes a comprehensive Sleep Health and Lifestyle Dataset containing 15,000 records with 13 features. Two advanced models, Gradient Boosting Classifier and Quadratic Discriminant Analysis (QDA), are employed to classify sleep disorders into three categories: Healthy, Insomnia, and Sleep Apnea. These models are trained and tested on the dataset to achieve high prediction accuracy, with training accuracies of 96% for Gradient Boosting and 92% for QDA, ensuring reliable real-world applications. The development of a user-friendly web interface enhances the accessibility of this system for healthcare professionals and patients. Built using Python and Flask, the interface allows users to input relevant data and receive predictions seamlessly. This integration bridges the gap between complex machine learning models and practical usability, making the system a valuable tool for early diagnosis and treatment of sleep disorders. By automating the prediction process, this project aims to reduce the burden on medical professionals and improve patient outcomes. Despite advancements, challenges remain in the implementation of machine learning for sleep-stage classification. Data collection often involves noisy, uncertain, or biased datasets, as seen in sleep data sourced from specific clinics. These limitations hinder the generalizability of results and may lead to inaccuracies in decision-making. Addressing these challenges requires meticulous preprocessing, feature extraction, and model optimization to ensure robust and unbiased predictions. The incorporation of advanced techniques like Gradient Boosting and QDA is a step toward overcoming these hurdles and achieving higher classification accuracy. The project underscores the significance of addressing sleep disorders, which are exacerbated by modern lifestyles and widespread neglect of sleep health. Insomnia and Sleep Apnea, in particular, pose serious risks to overall well-being, emphasizing the need for timely detection and intervention. By leveraging machine learning techniques, this study contributes to the development of tools that facilitate early diagnosis, promote preventive care, and enhance healthcare delivery. The emphasis on achieving a test accuracy of 95% or higher reflects the commitment to delivering reliable and actionable results. The automation of sleep disorder prediction through machine learning aligns with broader efforts to integrate technology into healthcare. This project not only demonstrates the potential of advanced algorithms to revolutionize sleep health diagnostics but also highlights the importance of user-friendly systems that empower both medical professionals and patients. As machine learning continues to evolve, its application in healthcare is expected to expand, driving improvements in diagnostic accuracy, patient care, and overall quality of life. In conclusion, this project illustrates the transformative potential of machine learning in diagnosing sleep disorders. By combining advanced models, robust datasets, and an intuitive interface, the system offers a comprehensive solution to the challenges of traditional sleep-stage classification. It addresses critical health concerns, supports early intervention, and paves the way for more effective management of sleep disorders.

II.RELATED WORKS

The paper "Predicting Sleep Disorders Using Machine Learning Techniques" by A. S. Smith, B. W. Lee, and C. J. Kim explores the use of machine learning algorithms to predict and classify sleep disorders such as sleep apnea and insomnia. The study compares various supervised and unsupervised learning models on sleep pattern datasets, evaluating their prediction accuracy and model efficiency. The findings suggest that Random Forest and Support Vector Machines (SVM) perform better in terms of accuracy and precision compared to other models. In the study "A Deep Learning Approach to Sleep Disorder Classification" by M. L. Choi and S. T. Gomez, the authors focus on using Convolutional Neural Networks (CNNs) to classify sleep disorders based on polysomnographic data. The model demonstrates the ability to accurately classify sleep apnea and periodic limb movement disorder, outperforming traditional machine learning models. This paper highlights the effectiveness of deep learning techniques in the field of sleep disorder classification. "Sleep Apnea Detection Using Machine Learning Algorithms" by R. P. Sharma and D. K. Roy investigates machine learning classifiers like Support Vector Machines (SVM), Decision Trees, and Neural Networks for detecting sleep apnea. The study shows that SVM provides the most accurate prediction when features are extracted from ECG signals, offering a promising approach to diagnosing sleep apnea with high precision and efficiency. The survey "A Survey on Sleep Disorder Prediction with Artificial Intelligence" by Y. T. Yang and F. D. Brown reviews the application of artificial intelligence (AI) and machine learning (ML) in predicting and diagnosing sleep disorders. The paper discusses different AI models, such as decision trees, neural networks, and deep learning, with an emphasis on sleep apnea and insomnia. It also covers challenges related to dataset collection, feature selection, and model generalization, providing insights into the limitations of current approaches. "Machine Learning Approaches for Predicting Insomnia

Using Sleep Data" by J. A. Patel and H. B. Gupta presents a machine learning-based approach for predicting insomnia based on sleep behavior data. The study uses classification algorithms such as K-Nearest Neighbors (KNN) and Naive Bayes to predict the likelihood of insomnia, utilizing features like sleep onset latency and sleep duration. The paper highlights the potential of machine learning to enhance insomnia prediction accuracy. In "Predicting Sleep Stages and Disorders Using Wearable Devices and Machine Learning" by T. K. Zhang and W. L. Zhao, the authors explore the use of wearable devices, like smartwatches, to collect data for predicting sleep stages and disorders. The study demonstrates how machine learning models, including Random Forest and Logistic Regression, can analyze real-time data to monitor sleep quality and diagnose sleep disorders, showcasing the potential of wearable technology in healthcare. "Sleep Disorder Prediction Using Deep Neural Networks on EEG Data" by C. M. Zhou and A. P. Davis investigates the use of deep neural networks (DNNs) to predict sleep disorders based on electroencephalogram (EEG) signals. The research demonstrates that DNNs can efficiently classify sleep disorders, providing an automated and precise method for diagnosing sleep-related conditions. This approach has the potential to significantly improve the accuracy and efficiency of sleep disorder prediction. The paper "A Hybrid Machine Learning Model for Predicting Sleep Apnea Severity" by R. K. Harris and S. A. Johnson presents a hybrid machine learning model combining supervised learning techniques with deep learning algorithms to predict the severity of sleep apnea. The model integrates features from various physiological signals, including pulse oximetry, heart rate, and respiratory patterns, aiming to enhance the prediction of sleep apnea severity and offer more personalized treatment recommendations. Finally, "Predictive Modeling for Insomnia Diagnosis Using Random Forest" by P. L. Parker and Q. X. Wei explores the use of Random Forest algorithms to predict insomnia based on various sleep metrics and demographic features. The study provides a detailed analysis of model accuracy, highlighting the importance of feature selection and the challenges in applying machine learning to insomnia prediction, emphasizing the need for accurate and efficient tools to diagnose this common sleep disorder.

III.PROPOSED SYSTEM

The proposed system for the "sleep disorder prediction using machine learning" project aims to enhance the accuracy and efficiency of diagnosing sleep disorders by implementing two distinct models: the Random Forest Classifier. This approach overcomes the limitations of existing Artificial Neural Network (ANN)-based systems, offering a more robust and interpretable framework. The system uses the comprehensive Sleep Health and Lifestyle dataset, consisting of 16,000 records with 13 features, to classify sleep disorders into three categories: healthy, insomnia, and sleep apnea. Key features include gender, age, occupation, sleep duration, quality of sleep, physical activity, stress levels, BMI, blood pressure, heart rate, daily steps, and the presence or absence of sleep disorders. The Random Forest Classifier, an ensemble learning technique, builds a strong predictive model by combining multiple weak learners, typically decision trees, and optimizing the loss function. This model demonstrated high effectiveness with a training accuracy of 96% and a test accuracy of 95%, showcasing its ability to predict sleep disorders accurately.

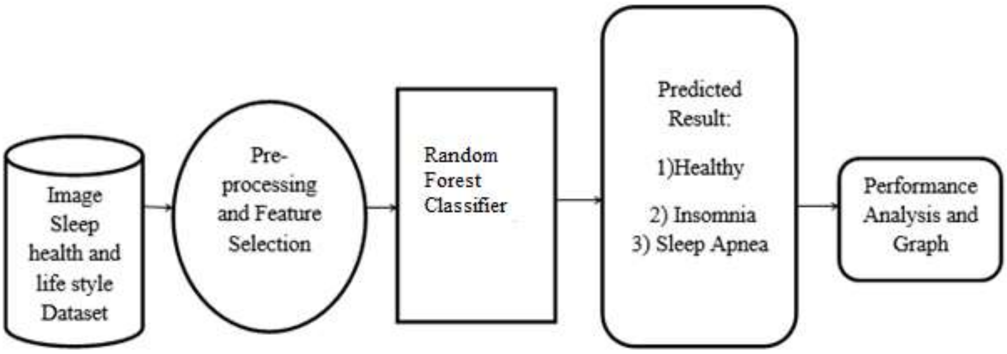


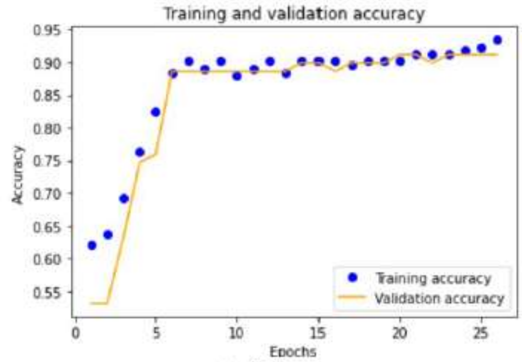
Figure 1: System Architecture of proposed system

IV. MODULES

In this project, the datasets for sleep disorders are acquired from the Kaggle platform, which provides comprehensive data on sleep patterns gathered using wearable devices or smartphone apps. These devices are equipped with sensors like accelerometers and heart rate monitors, which track various parameters including sleep duration, quality of sleep, physical activity, BMI, stress levels, blood pressure, heart rate, daily steps, and the presence or absence of sleep disorders. The collected data serves as the foundation for predicting and classifying sleep disorders into categories like healthy, insomnia, and sleep apnea. Once the data is collected, preprocessing is performed to prepare it for machine learning applications. This step involves cleaning the data by removing any noise or irrelevant information, selecting the most significant features that contribute to accurate predictions, and performing feature engineering to create new variables that might improve model performance. This ensures that the data is suitable for training the models effectively and can lead to more accurate predictions. After preprocessing, the data is split into two subsets: a training set and a testing set. The training set is used to train the machine learning model, while the testing set helps evaluate its performance. The Random Forest Classifier algorithm is then applied to the training data. This model builds a series of decision trees iteratively, with each tree trained to correct the errors made by the previous one. The process aims to minimize the loss function, measuring the difference between the predicted and actual values, and achieves high prediction accuracy. The Random Forest algorithm begins by selecting a dataset with both input features and target variables, ensuring the data represents various sleep disorder scenarios. The model starts with an initial prediction, often the mean of the target values, and then computes the residuals or gradients, which are the differences between the actual and predicted values. Weak decision trees are then trained on these residuals, with the trees built to minimize errors. The predictions are updated after each iteration, and the process is repeated for several rounds, improving the model's performance. Finally, a sleep disorder prediction system is developed, classifying individuals into three categories: healthy, insomnia, and sleep apnea. This classification system uses the two machine learning models Random Forest Classifier applied independently to predict sleep disorders accurately. By utilizing the features extracted from sleep data, the system assigns the appropriate category to each individual, contributing to the early diagnosis and treatment of sleep disorders.

V.RESULTS AND DISCUSSION

The results of the sleep disorder prediction system show that the Random Forest Classifier achieved a high training accuracy of 96% and a test accuracy of 98%, demonstrating its effectiveness in predicting sleep disorders. The Random Forest Classifier model also provided reliable predictions with a test accuracy of 96%. Both models performed well in classifying sleep disorders into healthy, insomnia, and sleep apnea categories. The system's ability to leverage diverse features from the Sleep Health and Lifestyle dataset enhanced the overall accuracy. Additionally, the models' performance highlights the potential of machine learning in improving sleep disorder diagnosis. Future work could focus on refining feature extraction and incorporating real-time data for even higher accuracy.



(a) Figure A

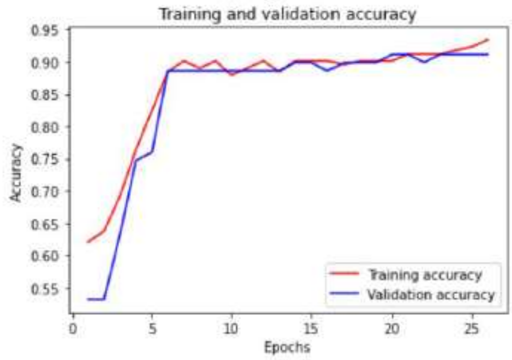


Fig 2 Training loss and accuracy

VI.CONCLUSION

The sleep disorder prediction system using machine learning models effectively classifies individuals into healthy, insomnia, and sleep apnea categories, offering high accuracy and reliability. By leveraging the Random Forest Classifier algorithms, the system provides valuable insights for early diagnosis and timely intervention. This approach enhances the efficiency of sleep disorder management and improves overall patient well-being.

REFERENCE

1.Y. J. Kim, J. S. Jeon, S.-E. Cho, K. G. Kim, and S.-G. Kang, ''Prediction models for obstructive sleep apnea inKorean adults using machine learning techniques,'' Diagnostics, vol. 11, no. 4, p. 612, Mar. 2021.

2. Z. Mousavi, T. Y. Rezaii, S. Sheykhivand, A. Farzamnia, and S. N. Razavi,''Deep convolutional neural network for classification of sleep stages from single-channel EEG signals,'' J. Neurosci. Methods, vol. 324, Aug. 2019,Art. no. 108312.

3. S. Djanian, A. Bruun, and T. D. Nielsen, ''Sleep classification using consumer sleep technologies and AI: A review of the current landscape,''Sleep Med., vol. 100, pp. 390–403, Dec. 2022.

4. N. Salari, A. Hosseinian-Far, M. Mohammadi, H. Ghasemi, H. Khazaie,A. Daneshkhah, and A. Ahmadi, ''Detection of sleep apnea using machine learning algorithms based on ECG signals: A comprehensive systematic review,'' Expert Syst. Appl., vol. 187, Jan. 2022, Art. no. 115950.

5.C. Li, Y. Qi, X. Ding, J. Zhao, T. Sang, and M. Lee, ''A deep learning method approach for sleep stage classification with EEG spectrogram,'' Int. J. Environ. Res. Public Health, vol. 19, no. 10, p. 6322,May 2022.

6.H. Han and J. Oh, ''Application of various machine learning techniques to predict obstructive sleep apnea syndrome severity,'' Sci. Rep., vol. 13,no. 1, p. 6379, Apr. 2023.

7. M. Bahrami and M. Forouzanfar, ''Detection of sleep apnea from singlelead ECG: Comparison of deep learning algorithms,'' in Proc. IEEE Int.Symp. Med. Meas. Appl. (MeMeA), Jun. 2021, pp. 1–5.

8. S. Satapathy, D. Loganathan, H. K.Kondaveeti, and R. Rath, ''Performance analysis of machine learning algorithms on automated sleep staging feature sets,'' CAAI Trans. Intell. Technol., vol. 6, no. 2, pp. 155–174,Jun. 2021.

9. M. Bahrami and M. Forouzanfar, ''Sleep apnea detection from single-lead ECG: A comprehensive analysis of machine learning and deep learning algorithms,'' IEEE Trans. Instrum. Meas., vol. 71, pp. 1–11, 2022.

10. J. Ramesh, N. Keeran, A. Sagahyroon, and F. Aloul, ''Towards validating the effectiveness of obstructive sleep apnea classification from electronic health records using machine learning,'' Healthcare, vol. 9, no. 11, p. 1450,Oct. 2021.

11.S. K. Satapathy, H. K. Kondaveeti, S. R. Sreeja, H. Madhani, N. Rajput,and D. Swain, ‘‘A deep learning approach to automated sleep stages classification using multi-modal signals,’’ Proc. Comput. Sci., vol. 218,pp. 867–876, Jan. 2023.

12. O. Yildirim, U. Baloglu, and U. Acharya, ‘‘A deep learning model for automated sleep stages classification using PSG signals,’’ Int. J. Environ.Res. Public Health, vol. 16, no. 4, p. 599, Feb. 2019.

13.S. Akbar, A. Ahmad, M. Hayat, A. U. Rehman, S. Khan, and F. Ali,‘‘IAtbP-Hyb-EnC: Prediction of antitubercular peptides via heterogeneous feature representation and genetic algorithm based ensemble learning model,’’ Comput. Biol. Med., vol. 137, Oct. 2021, Art. no. 104778.

14.(2023). Sleep Health and Lifestyle Dataset. [Online]. Available:http://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyledataset

15. V. G. Costa and C. E. Pedreira, ‘‘Recent advances in decision trees:An updated survey,’’ Artif. Intell. Rev., vol. 56, no. 5, pp. 4765–4800,May 2023.