



International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 8.699

Volume 14, Issue 3, March 2025

Predictive Analytics for Identifying Rest – Associated Anomaly using -Machine Learning

Dr.K.Sivaraman, Vakkapatla Gokul Krishna, Uppari Lakshmanudu, Vajinapally Vamshi

Associate Professor, Department of CSE, Bharath Institute of Higher Education and Research, Selaiyur, Chennai,
Tamil Nadu, India

B. Tech Students, Bharath Institute of Higher Education and Research, Selaiyur, Chennai, Tamil Nadu, India

ABSTRACT: Sleep disorders, such as sleep apnea and insomnia, affect millions worldwide, leading to serious health complications if left undiagnosed. Traditional diagnostic methods, like polysomnography (PSG), are expensive, time-consuming, and require clinical supervision. Machine learning (ML) provides an efficient alternative by analyzing physiological signals, such as electroencephalogram (EEG), electrocardiogram (ECG), and respiratory patterns, to automatically classify sleep disorders. This study explores various ML techniques, including neural networks, support vector machines (SVM), and decision trees, to improve diagnostic accuracy and efficiency. While ML models demonstrate high potential, challenges like data availability, feature selection, and model interpretability remain. Overcoming these challenges could lead to faster, more accessible, and cost-effective sleep disorder diagnosis, ultimately improving patient outcomes.

KEYWORDS: Genetic Algorithm, Polysomnography, Insomnia, sleep Apnoea, Artificial Neural Networks (ANN).

I. INTRODUCTION

Sleep is a vital physiological function necessary for physical and mental health. Sleep helps strengthen the body and consolidate the brain and memories. Sleep quality affects cognitive functions, particularly in children and older drivers at increased risk of accidents. Sleep deprivation can affect the human body and cause health problems like heart disease, diabetes and obesity. Physicians, doctors, medical professionals and experts must manually evaluate polysomnography (PSG) records, which can lead to different assessments of sleep stages. Manual classification is prone to human error and is time-consuming for sleep-stage classification.

Philips conducts an annual World Sleep Day survey on sleep-related attitudes and behaviours. In 2021, the survey polled more than 13,000 adults in 13 countries. Only 55% of adults were satisfied with their sleep, and the rest were dissatisfied with their sleep quality. They suffered from sleep quality because of such factors as the coronavirus disease 2019 (COVID-19) pandemic, sleep apnoea and insomnia. The statistics revealed that 37% said the pandemic negatively influenced their ability to sleep well. Moreover, 37% of participants reported suffering insomnia, while 29% snore, 22% have a shift-work sleep disorder, and 12% experience sleep apnoea. Medical professionals and sleep experts evaluate the quality of sleep by analysing the sleep system classified for various sleep stages. There are five stages of sleep: wakefulness, N1, N2, N3, and rapid-eye

II. LITERATURE REVIEW

A Review on Machine Learning Algorithms for Sleep Disorder Detection" -O. Tsinalis, P. M. Matthews, and Y. Guo, 2024, Academic Strive Exploration of machine learning algorithms, particularly deep learning. Discussed recent developments in machine learning applications in sleep medicine, emphasizing the potential of AI in improving detection and diagnosis of sleep disorders. Provided a comprehensive review of sleep stage classification, analyzing 38 papers and highlighting the effectiveness of combined signals and machine learning technique.

III. METHODOLOGY

This section focuses on implementing deep learning algorithms and conventional MLAs to classify sleep disorders. The following sections describe the datasets to assess the proposed algorithms, the performance metrics to evaluate the models, and the feature importance technique to calculate a score for inputted features. In addition, the classification algorithm used in this research is briefly explained. The data set used in this study is the Sleep Health and Life style Dataset downloaded from the Kaggle website [22]. The original dataset includes 400 observations and 13 columns of various data types. Each observation represents the actual sleep state. These data can be categorised into 13 variables relevant to sleep and daily habits, such as gender, age, occupation, sleep duration and sleep quality. Column 13 presents the sleep disorder for each person. This dataset groups the data into three sleep disorder categories, none, sleep apnoea and insomnia, pre-processing step was performed to replace the labels namely: None, Sleep Apnoea and Insomnia .

ID	Gen	Age	Occu	Sle Dur	Q of Sle	Phys Act	Str Lev	BMI Cat	Blood Pr	HR	DS	Sleep Dis- or- der
1	M	27	SW	6.1	6	42	6	Overw	126/83	77	4200	None
2	M	28	DR	6.2	6	60	8	Normal	125/80	75	10000	None
3	M	28	DR	6.2	6	60	8	Normal	125/80	75	10000	None
4	M	28	Sal	5.9	4	30	8	Obese	140/90	85	3000	Apnoea
5	M	28	Sal	5.9	4	30	8	Obese	140/90	85	3000	Apnoea
6	M	28	SW	5.9	4	30	8	Obese	140/90	85	3000	Insomnia
7	M	29	Teac	6.3	6	40	7	Obese	140/90	82	3500	Insomnia
8	M	29	DRr	7.8	7	75	6	Normal	120/80	82	8000	None

Table: Detailed information about the sleep health and lifestyle database records in this study.

This section proposes an assessment algorithm for the classification of sleep disorders. The methodology comprises of two approaches. The first approach, the model learns from data that are not scored on model performance, and the remaining 30% of the dataset is the testing set. Model performance is evaluated on unseen testing sets, where the model learns from the data without tuning and optimising the parameters.

IV. IMPLEMENTATION

An SVM is a supervised learning algorithm that can be used for classification or regression . The SVM aims to make the best line, called a decision boundary, and is based on the class of hyperplanes that is the line with the highest margin between two classes. The margin is the vertical distance between the decision boundary and closest data points. Moreover, the SVM is efficient when the number of samples is less than the number of dimensions in the dataset. In addition, SVMs can be used with various kernel functions, such as the RBF and linear functions, allowing the model to learn complicated decision functions. The KNN is a nonparametric supervised learning algorithm that can be used for classification and regression to classify a data point associated with its closest neighbor. The KNN algorithms are based on feature similarity. The value of k is a process called parameter tuning that refers to the number of nearest neighbor data points to include in the majority voting process. There are various types of distance metrics, such as Euclidean, Manhattan and Minkowski .

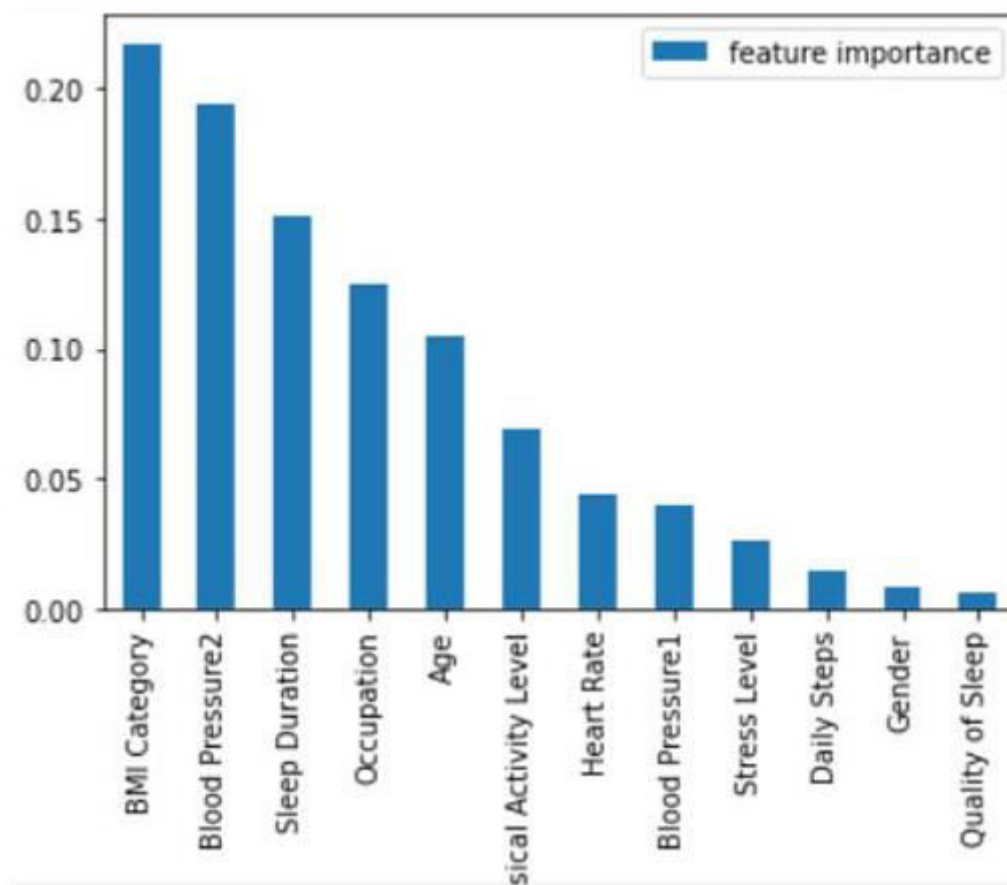


FIGURE . Feature importance

V. RESULTS AND CONCLUSION

In this paper, an optimised model for sleep disorder classification is proposed that implements MLAs with a genetic algorithm to explore optimal hyperparameter values for each model and obtain good results. This paper analysed the performance of MLAs for sleep disorder classification and evaluated many state-of-the-art MLAs on the real-world Sleep Health and Lifestyle Dataset. In addition, MLAs can learn from high-dimensional sleep data and attempt to classify sleep disorders without depending on expert-defined features. The proposed optimised ANN with GA achieved the highest accuracy over the other MLAs at 92.92%. The precision, recall, and F1-score values on the testing data were 92.01%, 93.80% and 91.93%, respectively. Even with 36120 a limitation in the amount of data. This study addressed the challenges in implementing MLAs for classification sleep disordering. However, large datasets are still needed for training and evaluating models in this field. The MLAs with GA can significantly improve the accuracy of sleep disorder classification. Future work will focus on developing MLAs using unsupervised learning in addition to assessing the dataset on a new model and comparing its performance against existing state-of-the-art models.

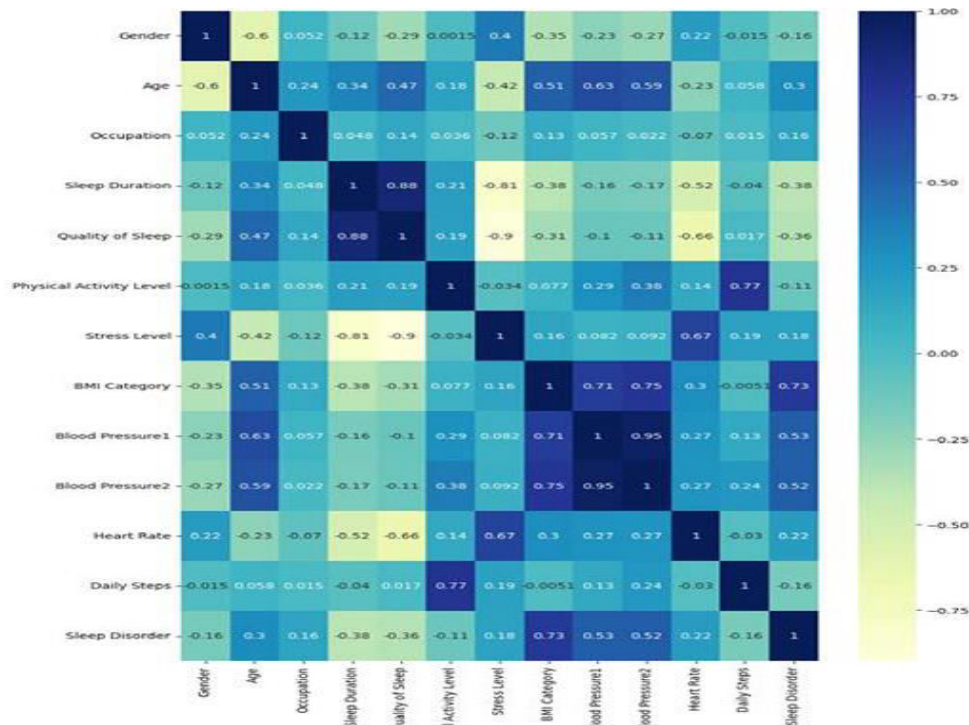


Fig : Correlation coefficient.

present a performance plot during training and validation. These plots were generated while attempting to classify the experimental values. Despite demonstrating similar loss curves, the points might not all be identical due to model weight changes. However, this training and validation loss provides a good comprehension of the learning performance changes over the number of epochs. This method assists in determining problems for the model during the learning phase to prevent overfitting the model and identify whether adding more training patterns improves the validation score. shows the results of the performance of all evaluated MLAs by training phase and presents the results were obtained using a 5-fold cross-validation.

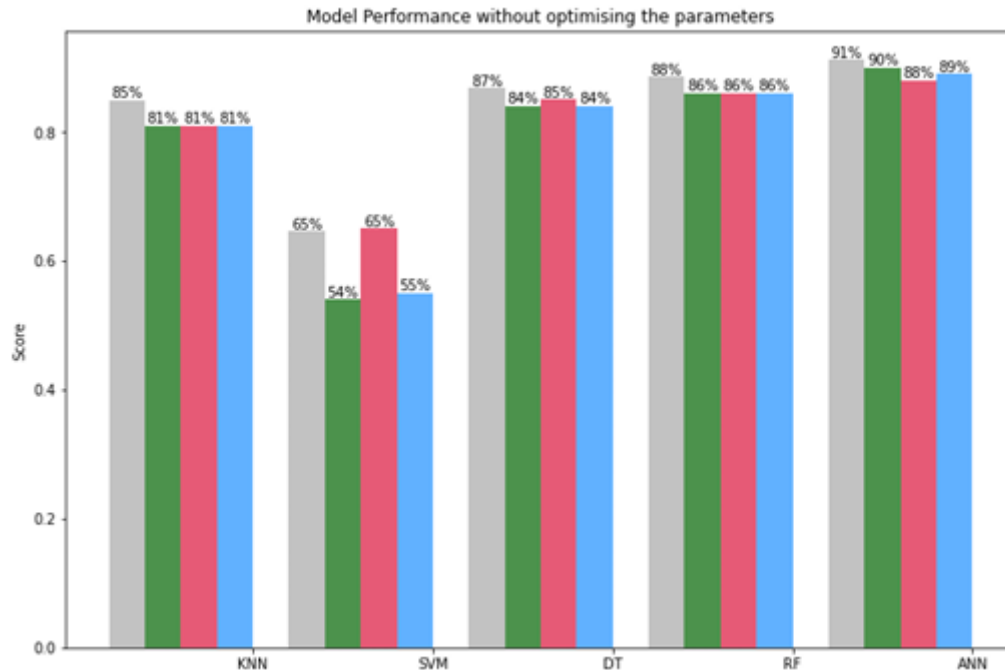


Fig : Results of the performance of all evaluated MLAs (As default parameters).

VI. LIMITATIONS AND FUTURE WORK

Integration of More Data: Combine various data sources (EEG, ECG, actigraphy, and lifestyle factors) for a more accurate classification of sleep disorders.

Advanced Machine Learning Techniques: Use advanced models (like ensemble methods, RNNs, and transformers) to improve classification accuracy, especially for time-based data.

Real-time Monitoring: Develop wearable devices with ML for real-time tracking of sleep patterns, enabling early detection and intervention.

Personalized Classification Models: Customize models based on factors like age, gender, and medical history for more accurate diagnoses across different populations.

Optimization of Algorithms: Enhance genetic algorithms and use techniques like particle swarm optimization to improve feature selection and model efficiency.

REFERENCES

- [1] F. Mendonça, S. S. Mostafa, F. Morgado-Dias, and A. G. Ravelo-García, "A portable wireless device for cyclic alternating pattern estimation from an EEG monopolar derivation," *Entropy*, vol. 21, no. 12, p. 1203, Dec. 2019.
- [2] Y. Li, C. Peng, Y. Zhang, Y. Zhang, and B. Lo, "Adversarial learning for semi-supervised pediatric sleep staging with single-EEG channel," *Methods*, vol. 204, pp. 84–91, Aug. 2022.
- [3] E. Alickovic and A. Subasi, "Ensemble SVM method for automatic sleep stage classification," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 6, pp. 1258–1265, Jun. 2018.
- [4] D. Shrivastava, S. Jung, M. Saadat, R. Sirohi, and K. Crewson, "How to interpret the results of a sleep study," *J. Community Hospital Internal Med. Perspect.*, vol. 4, no. 5, p. 24983, Jan. 2014.
- [5] V. Singh, V. K. Asari, and R. Rajasekaran, "A deep neural network for early detection and prediction of chronic kidney disease," *Diagnostics*, vol. 12, no. 1, p. 116, Jan. 2022.

- [6] J. Van Der Donckt, J. Van Der Donckt, E. Deprost, N. Vandenbussche, M. Rademaker, G. Vandewiele, and S. Van Hoecke, “Do not sleep on traditional machine learning: Simple and interpretable techniques are competitive to deep learning for sleep scoring,” *Biomed. Signal Process. Control*, vol. 81, Mar. 2023, Art. no. 104429.
- [7] H. O. Ilhan, “Sleep stage classification via ensemble and conventional machine learning methods using single channel EEG signals,” *Int. J. Intell. Syst. Appl. Eng.*, vol. 4, no. 5, pp. 174–184, Dec. 2017.
- [8] Y. Yang, Z. Gao, Y. Li, and H. Wang, “A CNN identified by reinforcement learning-based optimization framework for EEG-based state evaluation,” *J. Neural Eng.*, vol. 18, no. 4, Aug. 2021, Art. no. 046059.
- [9] Y. J. Kim, J. S. Jeon, S.-E. Cho, K. G. Kim, and S.-G. Kang, “Prediction models for obstructive sleep apnea in Korean adults using machine learning techniques,” *Diagnostics*, vol. 11, no. 4, p. 612, Mar. 2021.
- [10] 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.



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details