**COMPARATIVE ANALYSIS OF MACHINE LEARNING TECHNIQUES**

**IN SLEEP DISORDER SEVERITY CLASSIFICATION**

Babatunde Obisanya

Faculty of Engineering Environment and Computing

Coventry University

Coventry, England

ABSTRACT

**Sleep is very important in human life and has a great effect on physical and mental well-being. Good sleep supports healthy brain function, physical development and overall health. However, lack of sleep can cause many problems, including increased hormone levels, changes in appetite, and an increased risk of obesity and heart disease. Sleep disorders such as sleep apnea and insomnia affect a large part of the world's population and negatively affects wellbeing.**

**This study aims to address sleep disorders using techniques in machine learning, focusing on identifying suitable classification models and evaluating their effectiveness. The literature review describes previous research that utilized machine learning to identify and classify sleep disorders and highlights challenges and advances in the field. The data set used, derived from the National Poll on Healthy Aging (NPHA), facilitates the training and validation of machine learning algorithms. Using experimental designs and methods, including feature selection and model training, this study evaluates the performance of different classifiers.**

**Results show logistic regression as the most accurate among the models that were evaluated, offering practical implications for personalized health interventions and early detection of sleep disorders. By tailoring interventions based on individual needs and facilitating early identification of sever groups among patients, this research will help improve health outcomes and preventive measures for sleep-related health problems.**

***Keywords—* Classification, Logistic regression, Machine learning, Sleep disorder, Random forest, Support vector machine**

I. INTRODUCTION

Sleep is an important natural activity for humans and plays a vital role in overall health[1]. Our bodies support healthy brain function and maintain good physical health during sleep[2]. In addition, sleep is very important for physical development and growth, especially in children and adolescents. Sleep greatly affects the way you think, work, study, work, and many other aspects of your daily life. It also affects our body's circulation, immunity and respiratory system[3].

On the other hand, lack of sleep (sleep sickness) causes many problems and difficulties in everyday life [4]. Sleep disturbances increase hormone levels, control hunger, increase consumption of sweet, salty and fatty foods, decrease levels of physical activity and increase the risk of obesity, stroke and heart disease [5]. It can also cause stress, fatigue and functional impairments [6,7]. In addition, sleep disorders are one of the main causes of sleep apnea. According to recent statistics from the US Census, more than 140 million people (70 million men, 50 million women and 20 million children) snore mainly due to sleep apnea. An estimated 936 million adults worldwide suffer from mild to severe sleep apnea. In addition, many international studies have shown that between 10% and 30% of the world's population suffers from insomnia, and in some countries the rate reaches 60%. In addition, sleep disorders are 7% higher in women than in men. Finally, sleep disorders represent a global epidemic that threatens the health and well-being of nearly 45% of the world's population.

The purpose of this study is to provide additional insights and contribute to solving the sleep disorder problem by using machine learning capabilities in sleep disorder classification [8]. In particular, this study focused on two main objectives.

1. To identify the most suitable classification models for sleeping disorder dataset among three different classification models

2. To evaluate different machine learning models using different evaluation metrics.

This paper is structured as follows: Section 2 reviews existing literature studies that used sequential and parallel ML techniques. Section 3 presents the dataset and experimental design. Section 4 presents the methodology followed in carrying out the experiments. Section 5 shows the results obtained. Section 6 concludes the paper with a conclusion and future directions.

II. LITERATURE REVIEW

Several studies have been published discussing the use of machine learning and deep learning-based techniques to detect and classify sleeping disorders/problems. This section is a summary of recent studies to introduce the existing literature. Summaries are arranged in parallel and sequentially.

1. reviewed 48 articles and discussed the importance and challenges of sleep apnea. In addition, machine learning algorithms (MLA) such as SVM, random forest (RF), and deep learning algorithms can be used to detect sleep apnea based on ECG signals. However, they found some challenges in applying MLA to sleep apnea classification: differences in electrocardiogram signals and limitations in the availability of training data for the models. In their study, SVM and deep learning-based neural networks were the best for detecting sleep apnea from ECG signals.

The authors of [10] used MLA methods for sleep phase classification using an electroencephalogram spectrogram (EEG). Classification of sleep stages takes more time. It is error-sensitive and uses MLAs (EEG) along with the signals for classification. In addition, accuracy is low because the data is unbalanced. They used four public datasets to evaluate their models. The results showed that the proposed algorithms achieved classification accuracies of 94.17%, 86.82%, 83.02% and 85.12% for the four datasets. They used deep learning algorithms to classify sleeping floors and built a deep learning model. Convolutional Neural Networks (CNN) were used to extract frequency features and time from the EEG spectrogram. The model contains several hidden layers directional long-short-term memory (LSTM) that recognizes predictive sequences, which is an important method for phase classification of EEG spectrograms.

Researchers in [11] used MLA to predict the severity of obstructive sleep apnea (OSA) using real-time data collected from 4,014 patients that were not publicly available. The authors performed supervised and unsupervised learning techniques such as gradient boosting, RF, and K-means. Their proposed method achieved a classification accuracy of 88%, 89%, and 91%, respectively. However, their study has limitations. Data was collected from a single center which may be biased and some values are missing from the data. They developed an MLA model that can be effectively used to predict the severity and duration of OSA.

In addition, researchers [12] proposed a model using traditional MLAs such as DT, KNN, RF and deep algorithms to detect sleep apnea using single-lead ECG. The authors used the PhysioNet ECG Sleep Apnea v1.0.0 dataset containing 70 records and hybrid convolutional recurrent CNNs for feature extraction and deep recurrent neural networks (DRRN) to capture the temporal pattern of the data. They used principal component analysis for dimensionality reduction. The accuracy of the hybrid CNN-DRNN architecture was better than other algorithms. They proposed hybrid deep neural networks for sleep apnea detection from ECG.

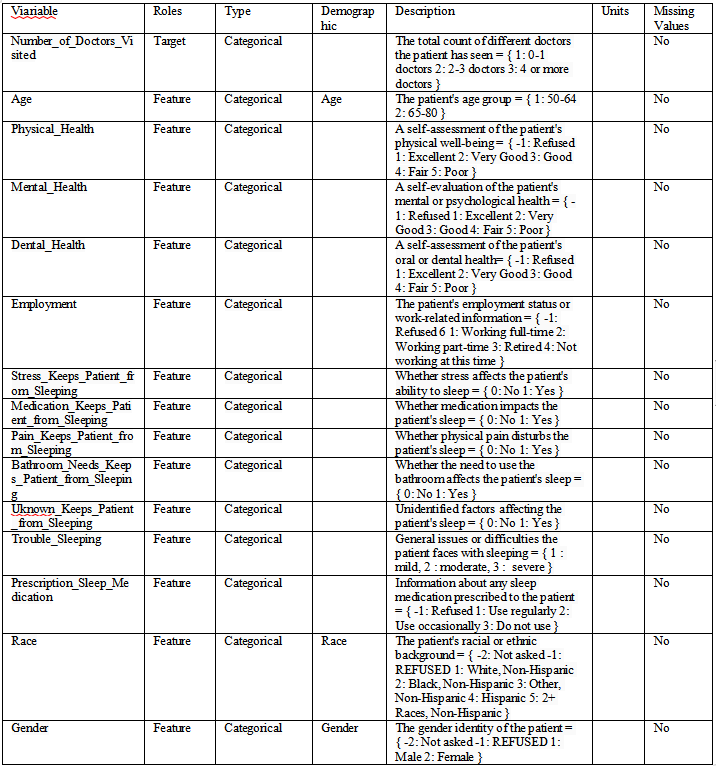
III. DATASET

The dataset used for this research is National Poll on Healthy Aging (NPHA) DataSet, gotten from UCI Machine Learning Repository [13]. It is a subset of the NPHA dataset filtered down to develop and validate ML-algorithms. The authors shed light on the key features that have been proven solid and efficient in predicting the stages of sleeping disorder stages. Each row represents a survey respondent. It contains sensitive data about race/ethnicity, gender and age. The dataset with 715 rows and 15 columns is divided into 2 parts, training set and test set. The training set makes up 70% of the data and test set is the remaining 30%.

Table 1 presents the NPHA dataset and the criteria for each feature of the data. It’s characteristics for the sleep disorder data severity.

1. Experimental Design

Different Machine Learning (ML) and libraries, such as Scikit Learn [14], Numpy [15], and Pandas [16] was used for implementation.

Three machine learning (ML) algorithms, SVM [17], random forest [18], and logistic regression [19], as well as a correlation matrix-based feature importance metric was used.

IV. METHODOLOGY

For this study, an experimental set-up was created to compare the performance of different classifiers in classifying the severity of sleep disorders. Recursive classification experiments were conducted to evaluate different model settings and configurations, including accuracy, recall, precision, F1 score, and confusion matrix.

Specifying the feature selection process, recursive Feature Elimination (RFE) was used to remove the six smallest features based on their gain ratio. The dataset was then split into 70% and 30% training and test sets, respectively, before standard 10-fold cross-validation train/test trials. Ensuring fair representation of all categories in the test data was prioritized. The classifiers were parametrized empirically to optimize performance based on basic experimental results.

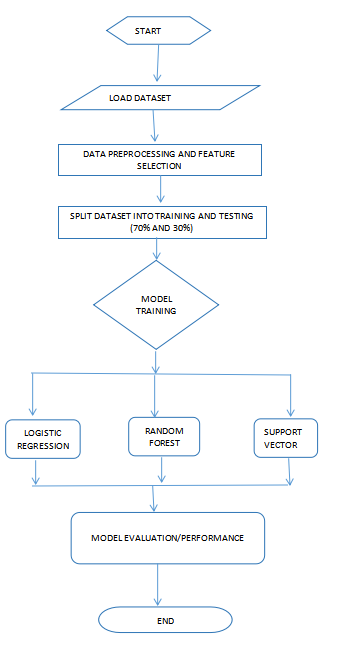
The system flow diagram is shown in Figure 1, which shows the sequential steps of the method.

Step 1 (The process starts by loading dataset): The NPHA dataset is read using pandas python library for systematic analysis.

Step 2 (Data preprocessing): Before modelling, the dataset goes through preprocessing steps such as coding categorical variables, handling missing values, and scaling numerical functions.

Step 3 (Feature selection): Feature selection techniques such as recursive feature elimination (RFE) are used to identify the most important features that contribute to sleep disorder severity classification.

Step 4 (Split the dataset into training and test set): The dataset is divided into two subsets: training set (70% of the data) to train the machine learning models and a test set (30% of the data) used to evaluate the performance of the models.



Step 5 (Model training) : Three machine learning models - random forest, logistic regression and support vector machine - are trained on the training data to learn the features in the data patterns and classify the severity of sleep disorders.

Step 6 (Evaluation of models): Trained models are evaluated using various metrics such as accuracy. , Recall, Precision, F1 Score, and Confusion Matrix to evaluate their effectiveness in predicting sleep disorder severity in a test dataset.

Step 7 (Classifier Performance Comparison): The performance of different classifiers is compared based on evaluation measures. which model works best in classifying severe sleep disorders.

Step 8 (END): The process ends with the classifiers after comparing their performance, which gives an idea of ​​the effectiveness of each model in dealing with the sleep disorder classification task..

Figure 2 shows a full visualization of the correlation matrix for all features in the dataset. The heatmap was used for visual representation

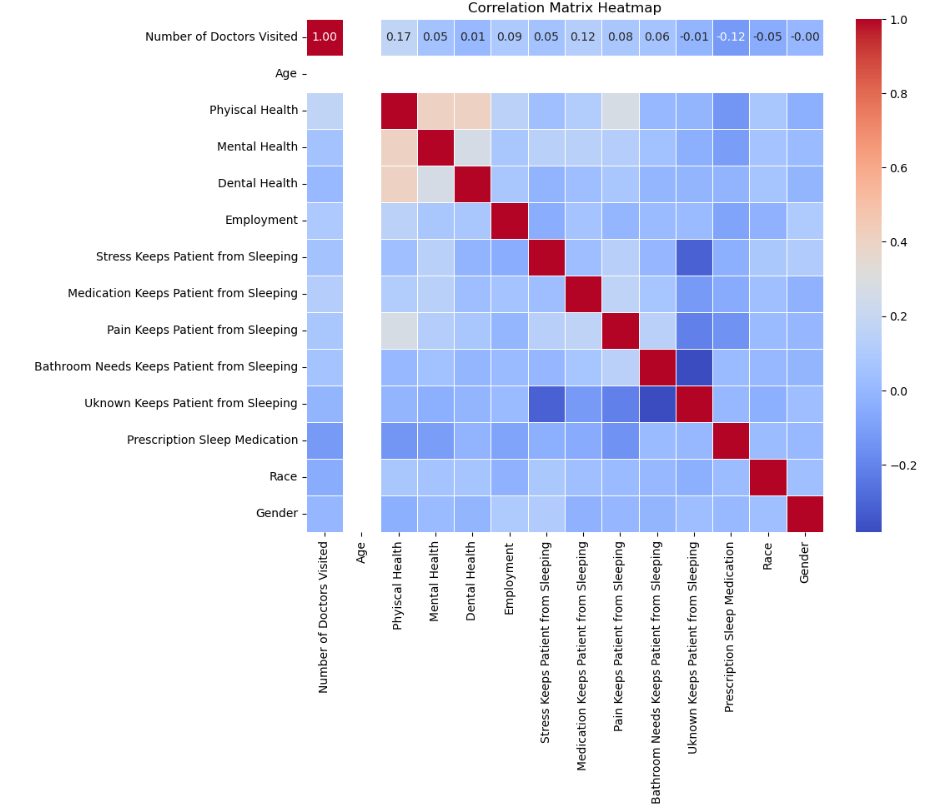


Fig 2 Correlation matrix with respect to all

Features

Figure 3 shows the correlation of the selected features to the Trouble Sleeping column class consisting of 14 features from th. The gender column was dropped due to having low correlation coefficient with multiple features.

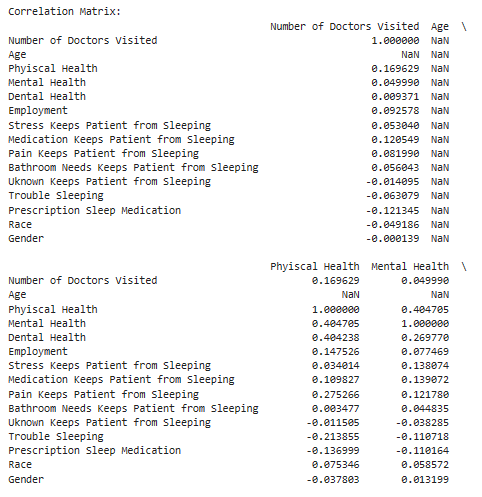


Fig 3. Selected Features

V. RESULTS

In this section, we evaluate various machine learning models using precision, accuracy, f1 score, recall and confusion matrix as evaluation metrics. Accuracy is the ratio of the number of correct predictions put to the total number of predicted samples. Accuracy is the ratio of true positives to predicted positives. Note that the proportion of true positives is correctly classified. The F1 score is a number between 0 and 1 and is a harmonic mean of precision and recall. F1 scores maintain a balance between precision and recall. All experiments were done in jupyter notebook. Figure 4 shows the distribution of classes in the Trouble Sleeping column of the level using a histogram.

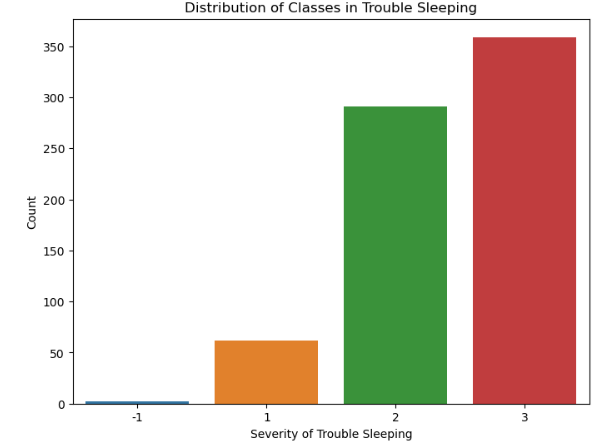


Fig 4. Distribution of classes

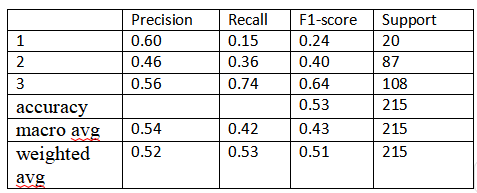
The values in the Trouble Sleeping column was plotted in the histogram in figure 4. the values of 1, 2, and 3 represents the mild, moderate and severe categories of sleeping abnormalities. The histogram shows that the severe category is more common among older men. The chart shows that a large amount of cases of about 650 fall under the moderate and severe categories. Where a little above 50 falls under the mild category. Table 2, 3, and 4 presents the classifiers outcomes of the evaluation of the ML models.

Random Forest Accuracy: 0.530

Random Forest Classification Report:

TABLE II.

Performance of each model

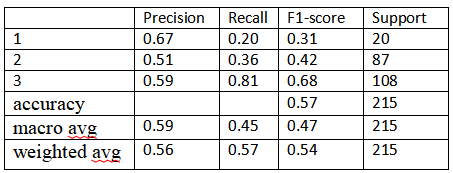


Logistic Regression Accuracy: 0.567

Logistic Regression Classification Report:

TABLE III.

Performance of each model

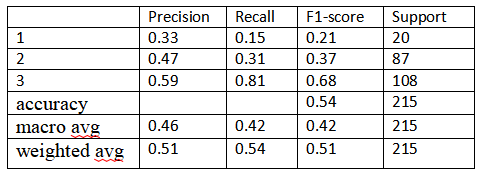


Support Vector Machine Accuracy: 0.544

Support Vector Machine Classification Report:

TABLE IV.

Performance of each model



The NPHA dataset was analysed and the accuracy of each machine learning model was evaluated. The result show that Logistic Regression is the most accurate and Random forest showed the lowest accuracy with regards to the dataset.

VI. CONCLUSION

In this work, based on the NPHA dataset, the features have been analyzed to discover more patterns of Sleeping problems. The selected features can reduce the training time and memory usage while maintaining high detecting accuracy. By using the feature selection approach integrated with the UCI NPHA dataset, the Logistic Regression algorithm is the most accurate with 56.7% accuracy compared to the remaining two of 54.4% and 53.0% (suffer from a minimal accuracy deterioration of 0.03%) using only 13 features removing the gender feature with lowest correlation coefficient of the original number of features.

Practical implications of this study include (a) personalized health interventions. By accurately classifying people into various degrees of sleep disturbance, healthcare providers can tailor interventions and treatment plans to the specific needs of each patient. Such a personalized approach can lead to more effective treatments and better health outcomes (b) Early detection and prevention: early identification of individuals at risk of serious sleep disorders allows timely intervention and preventive measures to reduce the potential health risks associated with sleep disorders. It can help prevent serious health problems such as obesity, cardiovascular disease and mental disorders.

REFERENCES

[1]. Zhang, M.-M.; Ma, Y.; Du, L.-T.; Wang, K.; Li, Z.; Zhu, W.; Sun, Y.-H.; Lu, L.; Bao, Y.-P.; Li, S.-X. Sleep disorders and non-sleep circadian disorders predict depression: A systematic review and meta-analysis of longitudinal studies. Neurosci. Biobehav. Rev. 2022, 134, 104532. [CrossRef] [PubMed]

[2]. Greenlund, I.M.; Carter, J.R. Sympathetic neural responses to sleep disorders and insufficiencies. Am. J. Physiol.-Heart Circ. Physiol. 2022, 322, H337–H349. [CrossRef] [PubMed]

[3]. Hu, X.; Li, J.; Wang, X.; Liu, H.; Wang, T.; Lin, Z.; Xiong, N. Neuroprotective Effect of Melatonin on Sleep Disorders Associated with Parkinson’s Disease. Antioxidants 2023, 12, 396. [CrossRef] [PubMed]

[4]. Sheta, A.; Thaher, T.; Surani, S.R.; Turabieh, H.; Braik, M.; Too, J.; Abu-El-Rub, N.; Mafarjah, M.; Chantar, H.; Subramanian, S.Diagnosis of Obstructive Sleep Apnea Using Feature Selection, Classification Methods, and Data Grouping Based Age, Sex, and Race. Diagnostics 2023, 13, 2417. [CrossRef] [PubMed]

[5]. Controne, I.; Scoditti, E.; Buja, A.; Pacifico, A.; Kridin, K.; Del Fabbro, M.; Garbarino, S.; Damiani, G. Do Sleep Disorders and Western Diet Influence Psoriasis? A Scoping Review. Nutrients 2022, 14, 4324. [CrossRef] [PubMed]

[6]. Alzyoud, M.; Alazaidah, R.; Aljaidi, M.; Samara, G.; Qasem, M.; Khalid, M.; Al-Shanableh, N. Diagnosing diabetes mellitus using machine learning techniques. Int. J. Data Netw. Sci. 2024, 8, 179–188. [CrossRef]

[7]. Aiyer, I.; Shaik, L.; Sheta, A.; Surani, S. Review of Application of Machine Learning as a Screening Tool for Diagnosis of Obstructive Sleep Apnea. Medicina 2022, 58, 1574. [CrossRef]

[8]. Sheta, A.; Turabieh, H.; Thaher, T.; Too, J.; Mafarja, M.; Hossain, S.; Surani, S.R. Diagnosis of obstructive sleep apnea from ECG signals using machine learning and deep learning classifiers. Appl. Sci. 2021, 11, 6622. [CrossRef]

[9] 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.

[10] 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.

[11] 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.

[12] 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.

1. Malani, Preeti N., Kullgren, Jeffrey, and Solway, Erica. National Poll on Healthy Aging (NPHA), [United States], April 2017. Inter-university Consortium for Political and Social Research [distributor], 2019-05-29. <https://archive.ics.uci.edu/dataset/936/national+poll+on+healthy+aging+(npha)>

[14] Scikit-learn: machine learning in Python — scikit-learn 0.22.1 documentation, Scikit learn.org. [Online]. Available: <https://scikitlearn.org/stable/>

[15] NumPy. [Online]. Available: <https://numpy.org/>

[16] Python Data Analysis Library, Pandas.pydata.org. [Online]. Available: <https://pandas.pydata.org/>

[17] Support Vector Machines — scikit learn 0.22.1 documentation, Scikit learn.org. [Online]. Available: <https://scikitlearn.org/stable/modules/svm.html>. [Accessed: 10- Jun-2019]

[18] Ensemble methods — scikit-learn 0.22.1 documentation, Scikit-learn.org. [Online]. Available: <https://scikitlearn.org/stable/modules/ensemble.html>

[19] Logistic-Regression <https://scikitlearn.org/stable/modules/generated/sklearn.linear_model.LogisticRegression.html>