Sleeping Stage Classification Using CNN

Lydia Pradeepa S, Dr. S. Jeevitha

Department Of Computer Science and Information Technology, Kalasalingam Academy Of Research And Education

Abstract: Maintaining proper health and mental stability is critical for overall health and well-being. Despite a good deal of research investment, sleep quality continues to be a crucial public challenge. Nowadays, people of all age groups are affected by improper sleep quality. Poor sleep can lead to a variety of neurological disorders. Sleep disorders are common in all subsets of the population, independently of gender. This public health challenge greatly affects quality of life in terms of both physical and mental health. Insomnia, parasomnias, sleep- related breathing difficulties, hypersomnia, bruxism, narcolepsy, and circadian rhythm disorders are some common examples of sleep-related disorders. Some of these disorders can be treated with proper analysis of early symptoms; in such cases, adequate sleep quality is essential for the patient's recovery. Artificial intelligence has several applications in sleep medicine including sleep and respiratory event scoring in the sleep laboratory, diagnosing and managing sleep disorders, and population health. While still in its nascent stage, there are several challenges which preclude AI's generalizability and wide-reaching clinical applications. Artificial intelligence is a powerful tool in healthcare that may improve patient care, enhance diagnostic abilities, and augment the management of sleep disorders. However, there is a need to regulate and standardize existing machine learning algorithms and deep learning algorithm prior to its inclusion in the sleep clinic. In this project, we can develop the framework for sleeping stage classification for both subjective data and ECG data and classify the data using Convolutional neural network algorithm to analyse the multiple sleeping stages.

Index Terms: Sleeping disorder, data preprocessing, neural network classification,

I. INTRODUCTION

Sleep disorders are a prevalent and challenging health issue affecting millions of individuals worldwide. Diagnosing and classifying sleep abnormalities accurately is crucial for effective treatment and management. In recent years, Convolutional Neural Networks (CNNs) have emerged as powerful tools for image recognition tasks. However, their application in sleep abnormality classification is relatively new and holds great promise. CNNs can analyze sleep data in the form of polysomnographic recordings or sleep-related images and extract meaningful features automatically. These features can then be used to classify different types of sleep abnormalities, such as sleep apnea, insomnia, and narcolepsy. By leveraging the hierarchical structure of CNNs, which consists of multiple layers of interconnected neurons, these models can effectively learn complex patterns and variations in sleep data.

This capability enables them to provide accurate and reliable classification results. Moreover, CNNs can handle large volumes of data, making them suitable for processing extensive sleep datasets. Additionally, by leveraging transfer learning, where pre-trained CNN models are fine-tuned on sleep data, the computational cost and training time can be significantly reduced. In summary, the application of CNNs for sleep abnormality classification shows great promise in improving the accuracy and efficiency of diagnosis and treatment of sleep disorders, ultimately leading to better patient outcomes. Another advantage of CNNs is their ability to handle complex and heterogeneous data. Sleep data can include a wide range of modalities, including EEG signals, heart rate measurements, airflow data, and more.

CNNs can handle multi-modal data by incorporating multiple input channels and learning joint representations of the different modalities. This enables them to capture complementary information from various sources and improve the overall classification performance. Moreover, CNNs have the ability to automatically learn and adapt to the features present in the data. Unlike traditional handcrafted feature extraction methods, CNNs learn feature representations directly from the data through the training process.

This eliminates the need for manual feature engineering, which can be time-consuming and may not capture all the relevant information in complex sleep data. By learning features adaptively, CNNs can discover discriminative patterns that may not be apparent to human experts, leading to improved classification performance. In summary, CNNs offer several advantages for sleep abnormality classification, including their ability to handle temporal and spatial data, handle complex and heterogeneous data, and learn adaptive feature representations. By leveraging these capabilities, CNNs can contribute to more accurate and efficient diagnosis and treatment of sleep disorders, ultimately improving the quality of life for individuals suffering from sleep abnormalities.

II. RELATED WORK

Asra Malik.,,et..al..{1}., proposed that Insomnia is a disorder of sleep in which a person constantly complains about insufficient sleep despite having an adequate amount of time to sleep. It is considered as one of the most common psychological illnesses with an approximate severity of about 10%. Frustration is the key feature of insomnia. The quality and quantity of sleep remain in trouble while falling asleep, sustaining sleep, or rising in the early morning. Insomnia is also associated with several other sleep- wake, emotional, or medical conditions which rises attention for its separate treatment. Increased physical, psychological, and physiological stimulation is considered to be a key underlying in most reports of chronic insomnia along with perpetuating behavioural issues like prolonged time in bed. It is also reported in a research work that insomnia can also combine with cardiovascular diseases. It has a very adverse effect on the job, mental, social, and overall quality of life of an individual. Insomnia is one of the serious sleep disorders. In this disease, a patient feels trouble falling asleep and stay awake for the whole night. Insomnia also co-morbid with other diseases like cancer, cardiovascular disease, asthma, Alzheimer's, Parkinson's disease. This research aims at a method for identifying insomnia using both electrocardiogram (ECG) and electromyogram (EMG) signals. The Cyclic Alternating Pattern (CAP) database from Physionet is used. Preprocessing by empirical mode decomposition (EMD) is carried out in this research. Features with high discriminative ability are extracted from signals to classify via logistic regression (LR), support vector machines (SVM), K-nearest neighbor (KNN), decision tree (DT), ensemble classifier (EC), and Naïve Bayes (NB) classification methods. we have achieved the highest accuracy of 100% on both ECG and EMG signals from our proposed methodology

Bufang Yang..,et..al., [2]., proposed a 1D- CNN model for automatic insomnia identification based on single-channel EEG labelled with sleep stage annotations, and further investigated the identification performance based on different sleep stages. Our experiments demonstrated that our 1D-CNN leveraging the 3 sub datasets composed of REM, LSS and SWS epochs, respectively, achieved higher average accuracy in comparison with baseline methods under both intra-patient and inter-patient paradigms. The experimental results also indicated that amongst all the sleep stages, 1D-CNN leveraging REM and SWS epochs exhibited the best insomnia identification performance in intra-patient paradigm, whereas no statistically significant difference was found in inter-patient paradigm. Overall, for automatic insomnia identification based on single-channel EEG labelled with sleep stages, 1D-CNN model introduced in this paper could achieve superior performance than traditional methods. Further experiment based on larger sleep databases under inter-patient paradigm is still required in future work. Moreover, the comparison of the two experiments demonstrated that the average identification accuracy of our 1D-CNN could achieve 99.16% in intra-patient experiment, whereas it could only reach 87.49% in inter-patient experiment with larger standard deviation. We consider the high accuracy in intra-patient experiment is caused by the similarity of epochs, i.e. epochs from the same patient are utilized both for training and testing. However, inter-patient experiment is more realistic evaluation paradigm which could guarantee the generalizability of the method. Therefore, we suggest future research on automatic insomnia identification based on deep learning should focus on the inter-patient experiment performance

Chih-En Kuo.., et..al.,[3]., Sleep takes approximately one-third of human live. A good sleep can help us getting the body to work right again, improved learning ability, physical development, emotional regulation, and good quality of life in human physiology. However, the prevalence of insomnia symptoms without restrictive criteria is approximately 33% in the general population. In the United States of America, 50-70 million people suffer from sleep disorders: among them, 30% of patients suffer from insomnia and 10% from chronic insomnia. Insomnia is defined as chronic when it has persisted for at least three months at a frequency of at least three times per week. When the disorder meets the symptom criteria but has persisted for less than three months, it is considered short-term insomnia. To diagnose insomnia, the physician may first execute a physical exam to look for signs of medical problems that may be related to insomnia and ask some sleep-related questions, such as sleep-wake pattern and daytime sleepiness. In addition, the physician may ask a subject to keep a sleep diary with the actigraphy for a couple of weeks. If the cause of insomnia is not clear, the subject should spend one or two nights at a sleep centre diagnosing another sleep disorder using polysomnography (PSG), such as sleep apnea. PSG recordings, which including electroencephalogram (EEG), electrooculogram (EOG), electromyogram (EMG), and other physiological signals are usually obtained from patients and scored by a well-trained clinical staff. Moreover, manual sleep scoring and diagnosis is a time consuming and subjective process, and the conventional polysomnography which uses many wires to connect instrument to patient is often a problem that leads to sleep disturbance. In general, people will experience the ''first night effect,'' which often interferes with sleep quality, if they sleep in hospitals or sleep centers.

Giuseppina Andresini., et.,al.,[4]., This paperoutlines a set of open challenges facing modern ML-based intrusion detectors relating to a lack of uniformity in the distribution of traffic data over time. To tackle them, we propose INSOMNIA, a semi- supervised approach that uses active learning to reducelatency in the model updates, label estimation to reduce labelling overhead, and applies explainable AI to describe how the model evolves to fit the shifting distribution. We extend the TESSERACT framework to perform a time-aware evaluation of INSOMNIA on a recently published, revised version of CICIDS2017 and demonstrate that modern intrusion detection systems must address concept drift in order to be effective. We envision that future work may build on INSOMNIA in order to design robust intrusion detection models that can be sustained over time. Theyrely on a subset of ground truth labels to retrain at eachtime step. In contrast to these previously mentioned approaches, INSOMNIA only requires labels at the initial training time, and then sustains itself without requiring manual labelling; instead, it generates pseudo labels based on the nearest centroid neighbour. Updating with the model's predicted labels has been proposed as a solution in the malware domain, but has been shown to lead to self-poisoning; this risk is mitigated in INSOMNIA's use of curtaining by using two distinct algorithms for label estimation andprediction. The results of our evaluation highlight yet another important open problem for NIDS: detection of stealthy, low-prevalence attacks. INSOMNIA struggles to detect the few instances of the Infiltration attack at windows and we observe that maintainingsensitivity to attacks with a very low base rate in the presence of high-volume attacks such as DoS is very challenging—as is generalizing to attack categories of greatly different character. While generalizing across attack types remains a holy grail, future work may consider ensembles that tackle different attack types with separate models

Manish Sharma., et.,al.[5]., Sleep is a fundamental human physiological activity required for adequate working of the human body. Sleep disorders such as sleep movement disorders, nocturnal front lobeepilepsy, insomnia, and narcolepsy are caused due to low sleep quality. Insomnia is one such sleep disorder where a person has difficulty in getting quality sleep. There is no definitive test to identify insomnia; henceit is essential to develop an automated system to identify it accurately. A few automated methods have been proposed to identify insomnia using eitherpolysomnogram (PSG) or electroencephalogram (EEG) signals. To the best of our knowledge, we are the first to automatically detect insomnia using only electrocardiogram (ECG) signals without combiningthem with any other physiological signals. In the proposed study, an optimal anti symmetric bi orthogonal wavelet filter bank (ABWFB) has been used, which is designed to minimize the joint duration-bandwidth localization (JDBL) of the underlyingfilters. We created ten different subsets of ECG signals based on annotations of sleepstages, namely wake (W), S1, S2, S3, S4, rapid eye moment (REM), light sleep stage (LSS), slow-wave sleep (SWS), non-rapid eye movement (NREM) and W+S1+S2+S3+S4+REM for the automated identification of insomnia. Our proposed ECG-based system obtained the highest classification accuracy of 97.87%, F1-score of 97.39%, and Cohen's kappa value of 0.9559 for K- nearest neighbour (KNN) with the ten-fold cross- validation strategy using ECG signals corresponding tothe REM sleep stage. The support vector machine (SVM) yielded the highest value of 0.99 for area underthe curve with the tenfold cross-validation corresponding to REM sleep stage

III.BACKGROUND OF THE WORK

Insomnia is a common sleep disorder that can lead to difficulty falling asleep, staying asleep, or waking up too early. There are various existing systems for insomnia sleep stage prediction that can provide valuable information about sleep patterns. Polysomnography (PSG) is considered the gold standard for sleep stage classification and involves monitoring various physiological signals to identify different stages of sleep. There are also various wearable devices available in the market that claim to track sleep stages, and machine learning models can be used to predict sleep stages based on various physiological signals. Support Vector Machine (SVM)is another machine learning algorithm that can be used for sleep stage classification. SVM is a supervised learning algorithm that can classify data into different categories by finding the best separating hyperplane.In sleep stage classification, SVMs are trained using features extracted from PSG recordings. The feature extraction process involves analysing the PSG recordings to extract relevant information about thedifferent sleep stages. The extracted features are then used to train the SVM, which can then classify new epochs of PSG recordings into different sleep stages. Some common features used for sleep stageclassification using SVMs include spectral features,such as power in different frequency bands, and statistical features, such as the mean and variance ofthe EEG signal. SVMs can also be combined withother machine learning algorithms, such as k-nearest neighbours (k-NN) or decision trees, to improve classification performance.

IV.SLEEPING STAGE CLASSIFICATION

Sleep is the brain's primary function and plays a fundamental role in individual performance, learning ability, and physical movement. One of the essential physiological processes of humans is sleep vital for physical and cognitive well-being and resurgence. Sleep is a reversible state in which the eyes are closed, and several nerve centres are disabled.

Sleep creates partial or unique or full anaesthesia for the individual, in which case the brain becomes a less complicated network. The conventional solution is to detect the artefacts and denoise the signal by removing corresponding epochs from the sleep signal. However, this way, the EEG signal will be manipulated and may lose important information. One of the motivations of this thesis is to develop and improve noise cancelation method that does not manipulate the signal and protect its originality. The proposed system for sleep stage classification using CNNs from an EDF and CSV file involves several steps. The first step is data pre-processing, where the raw data is filtered, resampled, and features are extracted. This step helps to remove noise and ensure that the data is consistent across all recordings. The next step is data splitting, where the data is split into training, validation, and testing sets. The training set is used to train the CNN model, the validation set is used to tune the hyperparameters and prevent overfitting, and the testing set is used to evaluate the performance of the model. Once the data is pre-processed and split, the next step is to design theCNN architecture. The architecture should be designedto learn relevant features from the pre-processed data and classify the sleep stages accurately. A common approach is to use 1D or 2D CNNs with multiple convolutional layers and pooling layers. The output of the final layer is fed to a fully connected layer and softmax activation to predict the sleep stage. And also predict the sleeping disorder with new datasets, then provide the precaution details.
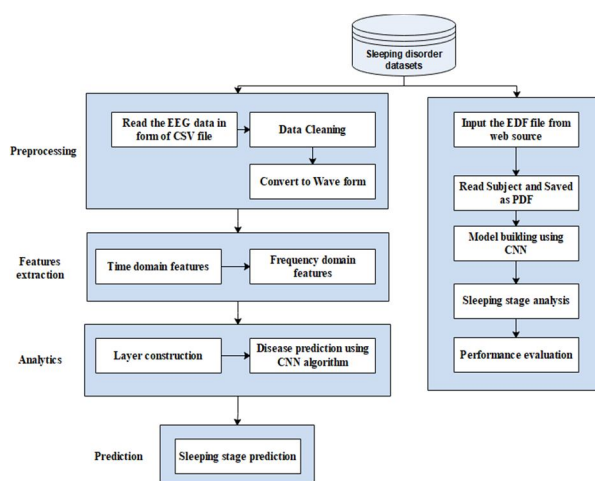

Fig 3: Proposed Framework

A. Datasets Acquisition

In this module we can input the CSV file aboutEEG data and also EDF file. The Physionet dataset is a dataset of EEG recordings for multiple persons PhysioNet is a public repository of biomedical signals and time series data, which can be used for research and development in the field of sleep stage prediction. It contains large datasets of sleep and physiological data, including EEG, ECG, and respiration signals, which can be used to train and evaluate deep learning algorithms for sleep stage prediction. In this module, we can input the CSV file and EDF file. CSV file contains the frequency values. EDF (European Data Format) is a common file format for storing physiological signals, including those recorded during sleep. EDF files typically contain several signals, including EEG, EOG, EMG, and ECG. These signals are recorded using electrodes placed on the scalp, face, and body, and provide information about the electrical activity of the brain and muscles during sleep. To use an EDF file for sleeping stage prediction, the first step is to load the file into a software program or Python library that can read EDF files. Examples of such libraries include MNE-Python and pyEDFlib. Once theEDF file is loaded, the next step is to extract features from the signals that are relevant to sleep stage prediction. Common features include the power spectrum, the amplitude of specific frequency bands, and the correlation between different signals

B. Preprocessing

Preprocessing is an important step in the analysis of EEG data, as it helps to remove artifactsand improve the quality of the data. In the pre- processing stage, continuous EEG recordings arefirstly segmented without overlapping by a sliding time window. In this module, perform preprocessing steps to eliminate the irrelevant and missing datasets from uploaded CSV file. The first step is to load the EDF file into memory using a library such as MNE- Python or pyEDFlib. The CSV file containing the sleep stage annotations can also be loaded into a Pandas DataFrame. Sleep signals are often contaminated by noise and artifacts. Signal filtering can be used to remove these unwanted components and improve the quality of the signal. Common filtering techniques include bandpass filtering, notch filtering, and high-pass filtering

C. Features Extraction

Feature extraction is the process of transforming raw data into a set of features that can be used to represent the data in a more meaningful and useful way. Relevant features need to be extracted from the signal to represent the underlying physiological processes. Common features for sleepstage classification include power spectral density, amplitude of specific frequency bands, and correlation between different signals. In the context of a Convolutional Neural Network (CNN) for EEG data, feature extraction involves transforming the raw EEG signals into a set of features that can be used as inputto the CNN. In this module, we can calculate the mean and standard deviation for uploaded CSV file. Normalizing the data is important to ensure that all features have the same scale and range. This helps to improve the performance of the machine learning algorithm. In CSV file, train the Y_values with stage and X_train values contain the frequency values. Based on features model file can be created using CNN algorithm for future verification

D. Classification

The features are divided into training,validation, and test sets, and are organized into a format that can be used as input to the CNN. A CNN isdesigned and configured, including the choice of architecture, number of layers, and activation functions. The CNN is trained on the training set, using a loss function and an optimization algorithm. The model is validated on the validation set to prevent overfitting and monitor the performance of the model. The trained CNN is evaluated on the test set to assess its performance in classifying the sleep stages. A CSV (Comma-Separated Values) file can be used for sleep stage classification in EEG datasets by providing annotations for each epoch of data. An epoch is afixed-length segment of EEG data, typically 30 seconds in length. The annotations in the CSV file indicate the sleep stage that each epoch belongs to, as determined by a sleep expert. To classify EEG datasetsusing a CSV file, the first step is to load the EEG data and the corresponding CSV file into memory. The EEG data can be loaded using a library such as MNE- Python or pyEDFlib, while the CSV file can be loaded into a Pandas DataFrame. Once the data is loaded, the next step is to preprocess the EEG data to extract features that are relevant for sleep stage classification. This can involve techniques such as filtering the data, resampling the data, and extracting features such as power spectral density, amplitude of specific frequency bands, and correlation between different EEG channels. A common approach for EEG classification is to use a convolutional neural network (CNN). The CNN is trained on the labeled epochs to learn the relationship between the EEG features and the sleep stages. The performance of the CNN is evaluated using metrics such as accuracy, precision, recall, and F1 score

Constructing the CNN Model

function INITCNNMODEL $(\theta, [n-5])$

layerType = [convolution, max-pooling, fully-connected, fully-connected];

layerActivation = [tanh(), max(), tanh(), softmax()]model = new Model();

for i=1 to 4 do

layer = new Layer(); layer.type = layerType[i];layer.inputSize = ni

layer.neurons = new Neuron [ni+1];layer.params = θi; model.addLayer(layer);

end for return model;end function

Training the CNN Model

Initialize learning rate, maximum iterations, minimum errors, and training batches. training BATCHES, batchsize SIZE, etc.;

Compute $n2$, $n3$, $n4$, $k1$, $k2$, according to $n1$ and $n5$;Generate random weights θ of the CNN;

cnnModel = InitCNNModel(θ, [$n1$–5]);iter = 0; err = +inf;

while err >ERRmin and iter<ITERmax doerr = 0;

for bach = 1 to BATCHES$_{training}$ do

$[\nabla\theta J(\theta), J(\theta)]$ = cnnModel.train (TrainingDatas,TrainingLabels),

Update θ

err = err + mean($J(\theta)$);

end for err = err/BATCHES$_{training}$;iter++;

end while

Save parameters θ of the CNN

E. Prediction

Finally, the trained CNN can be used to classify the sleep stages in new EEG datasets. TheEEG data is divided into epochs, and the CNN predicts the sleep stage for each epoch. This information can be used to analyze the sleep patterns of individuals and can provide valuable insights for clinical and research purposes. In this module, input the EEG readings in terms of numeric format. And match the reading with CNN model file. Then classify the stages of sleeping disorders and provide the precautions to users

V. EXPERIMENTAL RESULTS

Different performance measures such as accuracy, sensitivity, specificity, error rate and precision can be derived for analyzing the performanceof the system.

True positive (TP): number of true positives - perfectpositive prediction

False positive (FP): number of false positives -imperfect positive prediction

True negative (TN): number of true negatives - perfectnegative prediction

False negative (FN): number of true negatives -imperfect negative prediction

A. Error Rate

Error rate (ERR) is computed as the fraction oftotal number of imperfect predictions to the total number of test data. The finest possible error rate is 0.0, whereas the very worst is 1.0. Minimization of thiserror rate will be the prime objective for any classifier.

$$ERR = \frac{FP + FN}{TP + TN + FN + FP}$$

ALGORITHM	ERROR RATE
RANDOM FOREST	0.75
SUPPORT VECTOR MACHINE	0.5
CONVOLUTIONAL NEURAL NETWORK	0.4

Fig 4: Error Rate for algorithms

According to the aforementioned graph, theproposed CNN method offers a lower failure rate than the current technique.

B. Accuracy

The percentage of overall flawless forecasts to the complete test data is known as accuracy (ACC). Additionally, it can be written as 1 - ERR. The maximum accuracy is 1.0, and the minimum accuracy is 0.0.

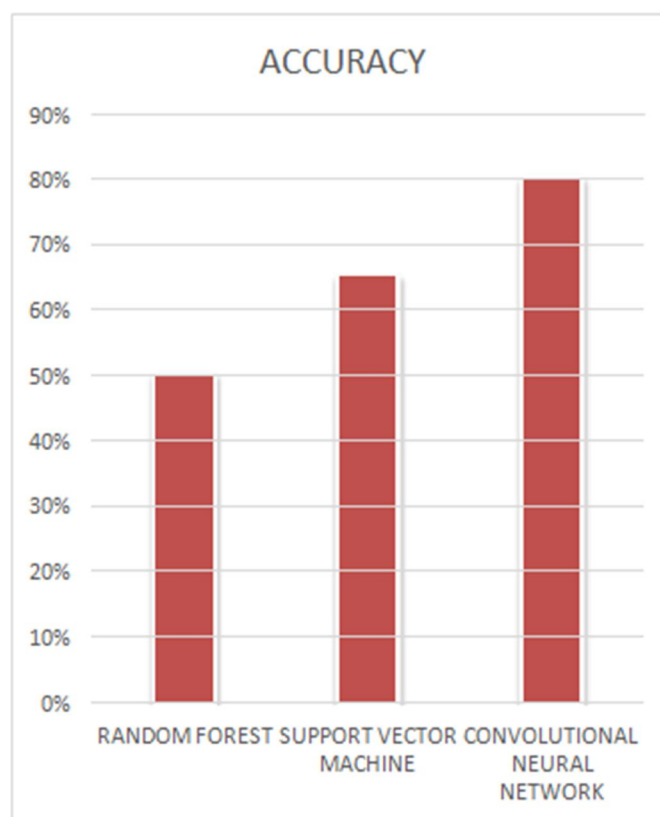| ALGORITHM	ACCURACY
RANDOM FOREST	50%
SUPPORT VECTOR MACHINE	65%
CONVOLUTIONAL NEURAL NETWORK	80%

Fig 5: Accuracy rate

According to the graph above, the proposedCNN algorithm has a higher accuracy rate than the current approach.

VI. CONCLUSION

In conclusion, using a CSV file for sleep stage classification in EEG datasets can provide valuableinformation about an individual's sleep patterns. By loading the EEG data and CSV file into memory and pre-processing the data to extract relevant features, a labelled dataset of epochs can be created that can be used to train and evaluate machine learning algorithms. A CNN is a common machine learningalgorithm used for EEG classification, which can learn the relationship between the EEG features and the sleep stages.

The performance of the CNN can be evaluated using metrics such as accuracy, precision, recall, and F1 score. The trained CNN can then be usedto predict the sleep stage in new EEG datasets by dividing the data into epochs and classifying each epoch based on its EEG features. Overall, CSV file classification for EEG datasets provides a powerful tool for sleep researchers and clinicians to analyze large amounts of sleep data quickly and accurately. It can help to identify sleep disorders, monitor treatment progress, and provide insights into the mechanisms of sleep.

REFERENCES

[1] Malik, Asra, et al. "Detection of Insomnia using Electrocardiography and Electromyography." 2021 International Conference on Artificial Intelligence(ICAI). IEEE, 2021.

[2] Yang, Bufang, and Hongxing Liu. "Automatic identification of insomnia based on single-channel EEG labelled with sleep stage annotations." IEEE Access 8 (2020): 104281-104291.

[3] Kuo, Chih-En, and Guan-Ting Chen. "A short-time insomnia detection system based on sleep EOG with RCMSE analysis." IEEE Access 8 (2020): 69763-69773.

[4] Andresini, Giuseppina, et al. "Insomnia: Towards concept-drift robustness in network intrusion detection." Proceedings of the 14th ACM Workshop on Artificial Intelligence and Security. 2021.

[5] Sharma, Manish, Harsh S. Dhiman, and U.Rajendra Acharya. "Automatic identification of insomnia using optimal antisymmetric biorthogonal wavelet filter bank with ECG signals." Computers in Biology and Medicine 131 (2021): 104246.

[6] Lee, Mi Hyun, et al. "Multitask fMRI and machine learning approach improve prediction of differential brain activity pattern in patients with insomnia disorder." Scientific reports 11.1 (2021): 1-13.

[7] Afshani, Mortaza, et al. "Discriminating paradoxical and psychophysiological insomnia based on structural and functional brain images: a preliminary machine learning study." (2022).

[8] Islam, Md Muhaiminul, et al. "Prediction of chronic insomnia using machine learning techniques." 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT). IEEE, 2020.

[9] Ma, Xiaofen, et al. "Functional connectome fingerprint of sleep quality in insomnia patients: Individualized out-of-sample prediction using machine learning." NeuroImage: Clinical 28 (2020): 102439.

[10] Wen, Ze-ying, et al. "Identification of discriminative neuroimaging markers for patients on hemodialysis with insomnia: a fractional amplitude of low frequency fluctuation-based machine learning analysis." BMC Psychiatry 23.1 (2023): 1-13.

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)