

Article

Employing a Convolutional Neural Network to Classify Sleep Stages from EEG Signals Using Feature Reduction Techniques

Maadh Rajaa Mohammed * and Ali Makki Sagheer

Department of Computer Science, College of Computer Science and Information Technology, Anbar University, Ramadi 31001, Iraq; ali_makki@uoanbar.edu.iq

* Correspondence: moath.raja@gmail.com

Abstract: One of the most essential components of human life is sleep. One of the first steps in spotting abnormalities connected to sleep is classifying sleep stages. Based on the kind and frequency of signals obtained during a polysomnography test, sleep phases can be separated into groups. Accurate classification of sleep stages from electroencephalogram (EEG) signals plays a crucial role in sleep disorder diagnosis and treatment. This study proposes a novel approach that combines feature selection techniques with convolutional neural networks (CNNs) to enhance the classification performance of sleep stages using EEG signals. Firstly, a comprehensive feature selection process was employed to extract discriminative features from raw EEG data, aiming to reduce dimensionality and enhance the efficiency of subsequent classification using mutual information (MI) and analysis of variance (ANOVA) after splitting the dataset into two sets—the training set (70%) and testing set (30%)—then processing it using the standard scalar method. Subsequently, a 1D-CNN architecture was designed to automatically learn hierarchical representations of the selected features, capturing complex patterns indicative of different sleep stages. The proposed method was evaluated on a publicly available EDF-Sleep dataset, demonstrating superior performance compared to traditional approaches. The results highlight the effectiveness of integrating feature selection with CNNs in improving the accuracy and reliability of sleep stage classification from EEG signals, which reached 99.84% with MI-50. This approach not only contributes to advancing the field of sleep disorder diagnosis, but also holds promise for developing more efficient and robust clinical decision support systems.



Citation: Mohammed, M.R.; Sagheer, A.M. Employing a Convolutional Neural Network to Classify Sleep Stages from EEG Signals Using Feature Reduction Techniques. *Algorithms* **2024**, *17*, 229.
<https://doi.org/10.3390/a17060229>

Academic Editor: Maryam Ravan

Received: 15 March 2024

Revised: 1 May 2024

Accepted: 3 May 2024

Published: 24 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: deep learning (DL); convolution neural network (CNN); electroencephalography (EEG); feature selection (FS); sleep stages

1. Introduction

Sleep is an essential aspect of everyone's lives. Good sleep is vital for maintaining one's mental and physical well-being. The rapid development of modern society, high-intensity jobs and education, and irregular lifestyles all have an impact on most people's sleep. Sleep disturbances have recently emerged as a major public health issue [1]. According to the American Academy of Sleep Medicine (AASM) guidelines, sleep is split into five stages: waking (W), non-rapid eye movement (NREM), which includes three stages (N1, N2, and N3), and rapid eye movement (REM), as illustrated in Figure 1 [2].

Humans often go from the W stage to the NREM stage and then the REM state. Sensors connected to different parts of the body capture the electrical activity in the brain linked to each sleep statistic. The brain is active in three different ways: delta, theta, and alpha. The occipital region exhibits alpha activity in the W stage. Shallow slumber, low alpha activity, and theta activity are the hallmarks of the N1 stage [3]. True sleep starts in the N2 stage, during which the sleep spindle forms a unique waveform. The N3 stage, which is a deep sleep phase, occurs when delta waves occur [4]. Lastly, the REM stage is distinguished by theta wave activity that is fast and low-voltage. Table 1 [5] presents the typical frequency of EEG waves for each stage of sleep.

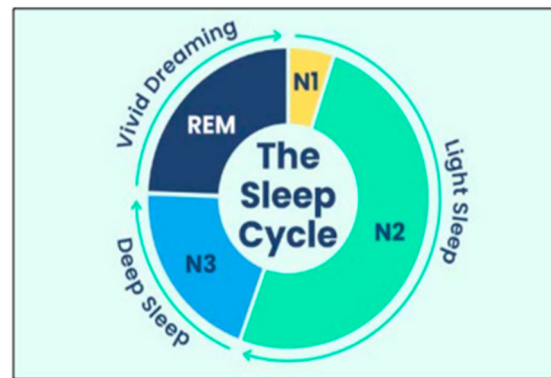


Figure 1. Classification of sleep stages [2].

Table 1. EEG signal characteristic frequency ranges for every sleep stage [5].

Sleep Stage	Characteristic Frequency
W	Alpha (8–12 Hz)
N1	Theta (4–8 Hz)
N2	Spindle (12–15 Hz)
N3	Delta (0.5–4 Hz)
REM	Alpha (8–12 Hz)
	Theta (4–8 Hz)

A normal sleep phase cycle includes the following percentages: 50% to 60% of the total amount of time spent sleeping in light sleep (N1, N2), 15% to 20% in deep sleep (N3), 20% to 25% in REM, and 5% or less in W sleep phases [6]. The electroencephalogram (EEG) signal is widely used in sleep staging because it is a bioelectrical signal that may precisely represent brain activity. The classic manual interpretation based on polysomnography can successfully finish the sleep staging, despite its high subjectivity and partiality to personal experience. Furthermore, the manual calibration method is labor- and time-intensive and takes a long time [7]. Tracking sleep quality and identifying sleep disorders depends on the classification of several sleep phases. Medical experts with clinical competence and training manually classify sleep phases for each sleep data segment. It takes time to manually classify the stages of sleep. When compared to manual classification, the algorithm- or model-based automatic classification approach is a more useful tool in clinical applications [8]. Deep learning- and machine learning-based techniques can be employed to classify sleep stages [9]. On the other hand, in sleep stage classification, deep learning techniques have progressively supplanted conventional machine learning techniques in recent years [10]. Because hand-crafted feature selection is necessary for standard machine learning techniques, it takes time and is unreliable when evaluated on unknown data [11]. Deep learning techniques, on the other hand, can overcome these issues by employing neural networks for sufficient and adaptable feature extraction. Convolutional neural network (CNN) architecture is used by most deep learning-based sleep stage classification methods now in use [12]. Consequently, the goal of this research is to use a one-dimensional-convolution neural network (1D-CNN) to develop a strong deep-learning classifier for sleep stage classification with feature selection techniques. Thus, these experimental results entail the outcomes of earlier experiments using deep learning techniques. The findings from previous deep learning approaches are compared with the output from the suggested Sleep-1D-CNN model. This model has gained sufficient experience in classifying sleep stages through training on a wide range of sleep data. Therefore, this study aims to offer a range of advantages for future sleep medicine research connected to artificial intelligence.

EEG signals include a large amount of information and not all of aspects may be equally important for classification tasks. Feature selection approaches like ANOVA and

MI can help in identifying the most discriminative EEG features related to sleep stages, thereby enhancing the model's capacity to discriminate between them. In addition, EEG signals often have high dimensionality due to the large number of electrodes and time points. ANOVA and MI will help to reduce the dimensionality by selecting the most effective features, which can lead to a more efficient and accurate CNN model. Selecting the most relevant features using these FS techniques can help prevent overfitting and improve the generalization ability of the proposed CNN model on EEG signals. Finally, FS techniques can reveal insights into the underlying patterns and features of EEG signals related to different sleep stages, advancing our understanding of brain activity through sleep.

The main contribution of deep learning-based models for EEG (electroencephalogram) signal classification, coupled with ANOVA (analysis of variance) and mutual information feature selection methods, lies in their ability to effectively extract relevant features from EEG data and improve classification performance. Combining deep learning models with ANOVA and mutual information feature selection techniques leverages the strengths of both approaches. Deep learning models provide the capacity to learn complex representations directly from raw EEG data, while ANOVA and mutual information feature selection help identify relevant features and reduce the dimensionality of the feature space, thereby enhancing classification accuracy and interpretability. Overall, this integrated approach can lead to more robust and effective EEG signal classification systems, with applications in various domains such as brain-computer interfaces, healthcare, and neuroscience research.

This research is presented in different sections. A brief survey of related studies is provided in Section 2. The proposed architecture is shown in Section 3. Section 4 discusses the specifics of the suggested methodology and algorithmic reasoning. A description of the result analysis is provided in Section 5. Finally, Section 6 discusses our study's concluding ideas.

Research Questions

To Assist Scientists Investigating EEG Analysis Using FS and DL:

1. With what deep learning model should researchers start with EEG analysis? Furthermore, improving the feature selection methods that work well can improve sleep stage classification accuracy. By providing researchers with the necessary resources and information, our work helps them effectively navigate the subject of deep learning and establish a strong foundation in it.
2. Which major EEG datasets are the best places for academics to begin their EEG-DL research? This paper offers a perceptive summary of recent developments, highlighting well-known datasets and the DL techniques that go along with them. By responding to these research questions, researchers may remain ahead of the curve in their area and uncover interesting possibilities for their studies by staying up to speed with cutting-edge research directions.

2. Literature Survey

Since polysomnography was established as the primary diagnostic method for measuring sleep duration and evaluating the quality of sleep in the 1980s, researchers from all over the world have carried out many studies to advance sleep classification approaches. The medical field continues to be quite interesting. Considerable effort has been made in this regard by evaluating different classifiers over a range of input data to increase the accuracy attained by employing such artificially constructed models. Table 2 shows related studies using deep learning for sleep stage classification. Many studies have been limited by the availability of high-quality and sufficient datasets. Limited access to annotated sleep data specifically for sleep disorders or diverse populations can hinder the development and evaluation of classification algorithms. Some studies may lack external validation or independent datasets, limiting the generalization of the proposed algorithms. Models trained and tested on specific datasets may not perform well on different datasets collected using different protocols or from different populations. For instance, a CNN model was

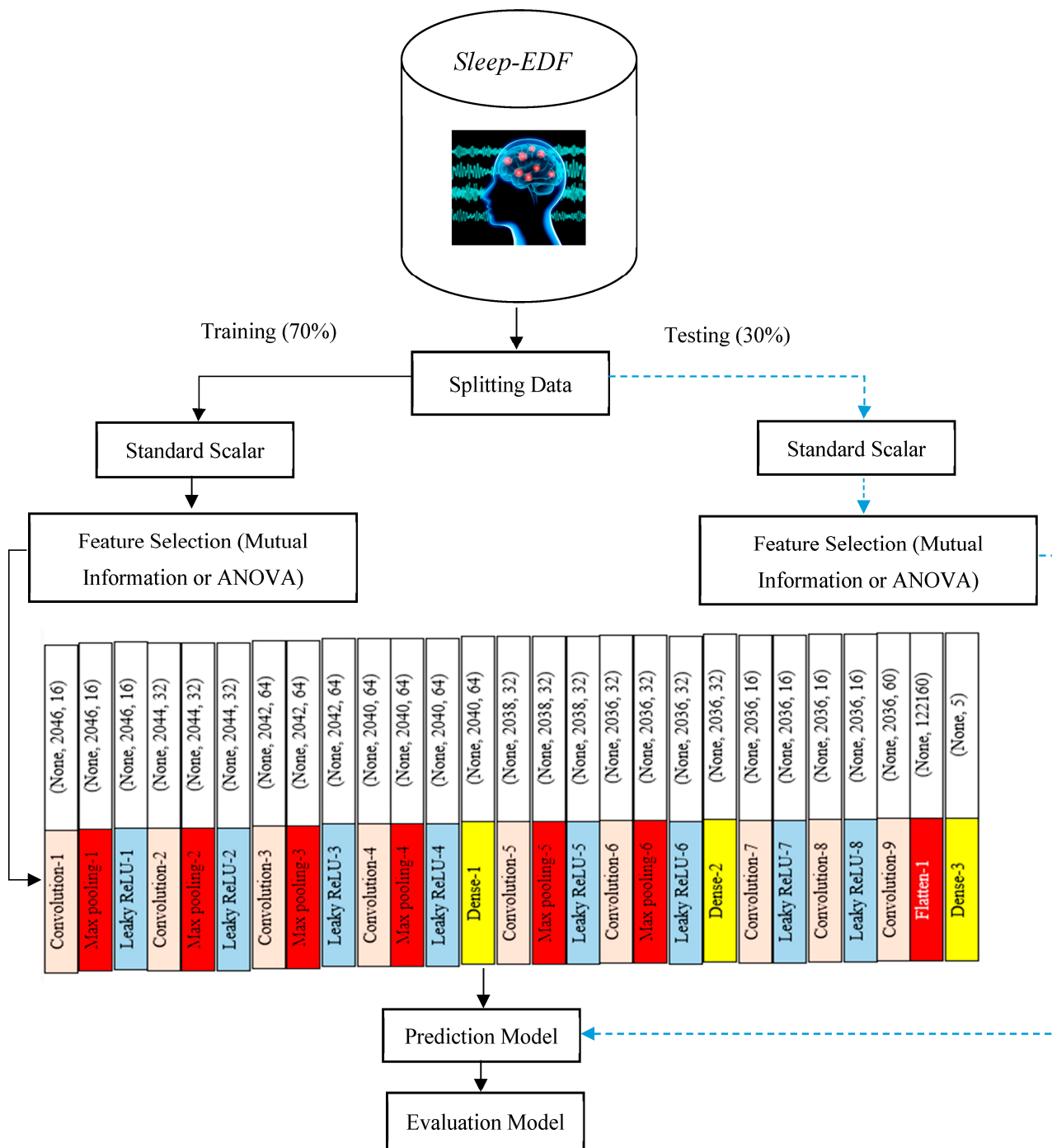
suggested by Yildirim et al. [13] to remove the need for feature engineering by extracting features from EEG and electrooculogram (EOG) segments. Ten 1D-CNN layers and a fully connected layer were used. The accuracy of their model was 94.24% in the three-sleep-stage classification and 90.98% in the five-sleep-stage classification. A CNN method was used by Nguyen et al. [14] to categorize the five phases of sleep. Their design comprised three 1D-CNN layers: max pooling and dropout layers follow the first layer, max pooling and batch normalization follow the second layer, and max pooling and two fully connected layers follow the final layer. With an accuracy rate of 87.67%, their method successfully identified the five phases of sleep. Z. Mousavi et al. [15] introduced an innovative deep learning-based method for automatically classifying different phases of sleep. The proposed method uses a single-channel EEG signal to divide classes. During the preprocessing stage, data augmentation is used to balance the number of epochs that correspond to different sleep phases. The proposed method for classifying two to six classes achieved an accuracy of more than 90%, which is superior to that of other existing methods. T. Zhu et al. [16] suggested a deep learning model using the attention mechanism and a 1D-CNN. It was found that the accuracy of classification for the five stages of sleep was 82.80% for Sleep-EdfX and 93.7% for Sleep-Edf datasets. S. Santaji and V. Desai, 2020 [17], proposed an effective approach for sleep stage classification utilizing electroencephalogram (EEG) data processing and machine learning classifiers spanning 10s epochs. EEG signals are considered to have an important role in automated sleep stage classification. A band-pass filter partitions EEG data into frequency sub-bands. Statistical features are extracted and trained using decision tree (DT), support vector machine (SVM), and random forest (RF) techniques, with various testing dataset percentages. The findings based on the Sleep-Edf dataset show that the random forest algorithm achieved 97.8% accuracy. Z. Cui et al. [18] suggested a CNN model that uses no feature engineering techniques to extract features from EEG, EMG, and EOG segments. They employed two 2D-CNN layers, max pooling, and a fully connected layer as the final layer. Their model's accuracy at classifying the five stages of sleep was 92% on a subject-wise test set. H. Phan et al. [19] proposed a 2D-CNN model and the fast Fourier transform (FFT) technique for classifying sleep stages into five stages. Their model's accuracy on a subject-specific test set was 83.6%. S. Satapathy et al. [20] suggested a deep learning and machine learning-based sleep staging paradigm. Utilizing 80:20 percentages for training and testing data, the proposed approach was trained on sleep-related diseases and healthy control patients. CNN+ Long Short-Term Memory (LSTM) and other deep learning algorithms were utilized to test the proposed model. The accuracy of the proposed CNN + LSTM was 87.4%. M. Lee et al. [21] suggested SeriesSleepNet, a novel automatic sleep stage classification model that combines bi-LSTM and CNN. This model works well with a time series model and is advantageous for learning temporal information, both intra- and inter-epoch. For the five-class classification, the total accuracy of the results was 87%. I. Masad et al. [22] suggested a novel approach that focuses on high accuracy while classifying sleep stages using EEG channels. By offering a new light CNN model for classification, the suggested method created a novel algorithm for identifying and categorizing sleep stages. With 99.39% accuracy, the suggested methodology can generate a precise and accurate sleep stage classification.

Table 2. Summary of literature survey.

Architecture	Year	Dataset	Input Signal	Accuracy
1D-CNN	2019	Sleep-Edfx	Electroencephalogram (EEG), Electrooculogram (EOG)	94.24% for 3 stages 90.98% for 5 stages
1D-CNN	2021	Sleep-Edf Sleep-EdfX ISRUC-Sleep	Electroencephalogram (EEG)	87.67%
CNN	2019	Sleep-Edf	Electroencephalogram (EEG)	90%
1D-CNN	2020	Sleep-Edf Sleep-EdfX	Polysomnography (PSG)	93.7% 82.80%
DT	2020	Sleep-Edf	Electroencephalogram (EEG)	DT = 93.80%
SVM				SVM = 94.14%
RF				RF = 97.8%
2D-CNN	2018	ISRUC-Sleep	Electroencephalogram (EEG)	92%
2D-CNN	2019	Sleep-Edf	Fast Fourier Transform (FFT)	83.6%
CNN + LSTM	2023	Sleep-Edf	Electroencephalogram (EEG)	87.4%
CNN + LSTM	2023	Sleep-Edf SHHS	Electroencephalogram (EEG)	87%
CNN	2024	CWT	Electroencephalogram (EEG)	99.39%

3. The Proposed System Architecture

Based on the principles of deep learning (DL), the suggested automated sleep stage categorization system makes use of feature selection strategies to boost the accuracy of the proposed model. As is customary in manual sleep stage scoring, the convolutional neural network (CNN) provides an accurate method of autonomously extracting significant features from the data supplied to the model, leading to a model that resembles the conventional approach quite closely while also significantly improving the procedure's overall reliability. The proposed system consists of three stages, as shown in Figure 2. The first stage is the preprocessing stage (data splitting and standard scaler). The second stage is feature selection (FS); in this stage, two feature selection methods are used to achieve the aim of this work. Finally, the classification stage is based on the proposed Sleep-1D-CNN model to classify the sleep stages. A notable highlight of our model is using feature selection techniques with a CNN model to increase the classification accuracy and decrease the execution time. The following steps briefly describe the work process involved. The steps given below are executed in order:



white males and females were captured for the EEG dataset without the use of medication. The subjects were between the ages of 21 and 35. Generally speaking, the sleep datasets were kept in the EDF format and were collected from several subjects between 1989 and 1994. Two EEG channels (Fpz-Cz and Pz-Oz) and one horizontal channel electrooculogram (EOG) were included in each database recording. A frequency of 100 Hz was used to sample the datasets. Two subsets with the names *sc** and *st** exist in this database. Four recordings make up each subset. Resporonasal, an event marker, and a submental-EMG envelope that indicates rectal body temperature, are included in the first subset, *sc**. For each of the subsets, the sampling rate was 1 Hz. The submental EMG in recording *st**, the second subset, has a 100 Hz sampling rate. Moreover, a 1 Hz sample of the even marker was taken. Every recording was scored according to the R&K criterion. The EEG recordings were split into 30 s epochs based on these criteria. Three thousand data points were present in each segment. These segments were classified as follows: AWA, S1, S2, S3, S4, REM movement time, and UNS (unscored). Six sleep stages were covered by the total number of segments (14,963) employed in this study. This study also made use of the Pz-Oz channel since it performed better in classification than the Fpz-Cz channel [23].

4.2. Data Preprocessing Phase

Data can be found in many different formats, including images, audio files, videos, organized, and unstructured tables. The free text, video, or images cannot be directly understood by a machine; the provided data must be converted into 1 s and 0 s. Therefore, it is not possible to feed raw data into a deep-learning model and expect it to learn [24]. Data preprocessing is the first step in machine learning in which data is transformed/encoded so that the machine can quickly read or comprehend it. In other terms, it might be understood as the model algorithm's ability to quickly assess data features. The most important aspect affecting how well machine and deep learning algorithms perform in terms of generalization is data preprocessing [25]. The amount of training data grows exponentially with respect to the dimension of the input space. Preprocessing may take between 50 and 80 percent of the entire classification time, based on estimations, indicating its importance in model construction. Enhancing the quality of the data is also necessary for improved performance [26].

4.2.1. Dataset Splitting

The process of dividing data into groups for independent model training and evaluation is called data splitting, or train:test splitting. The procedure comprises dividing the dataset into test and train sets, with equal percentages of development, testing, and cross-validation sets, and the biggest training set [27]. While the training dataset is used to build the model, the testing dataset assesses its predictive potential. Nevertheless, the train:test split ratios and dataset sizes might have a significant influence on the model's output and classification performance. Therefore, dividing data is a common technique used in deep learning [28]. The dataset may be split using a number of ratios, such as 10:90, 20:80, 30:70, 40:60, 50:50, 60:40, 70:30, 80:20, and 90:10 train:test splits. The testing dataset is used to evaluate the model's predictive capacity, whereas the training dataset is used to develop the model [29]. The Sleep-EDF dataset is split into two subsets for this work: the training dataset contains 70% of the total observations in the original dataset, and the test dataset contains 30% of the total observations.

4.2.2. Dataset Scaling

The raw data found in datasets coming from the real world frequently have problems that make it impossible to use them for machine and deep learning tasks. The multiple scales used to portray the different dataset features are one of the most noticeable problems. As a result of this problem, the algorithms in deep learning and predictive machine learning learn less accurate models [30]. This is because algorithms tend to rely on features that fluctuate within a larger range of the dominant features, even if they may not necessarily

be the most enlightening or significant elements for the appropriate categorization of data examples. To solve this problem, scaling techniques are used for the relevant data such that each feature fluctuates within the same range. Scaling is, therefore, one of several preprocessing steps that are usually necessary prior to applying a deep learning model to a dataset [31]. The Z-score normalization is implemented using the *Standard Scaler* approach, which standardizes features by taking the “mean” of each value and dividing the result by the standard deviation of the attribute. This yields a distribution with unit variance and zero mean. Let $\bar{x}$ be the mean of the “ x ” variable; Equation (1) transforms (scales) a value x_i into $\bar{x}_i$.

$$\bar{x}_i = \frac{x_i - \bar{x}}{S} \quad (1)$$

The sample mean of the property in this case serves as the translational term, while the standard deviation serves as the scaling factor. This approach has the potential to transform attributes with positive and negative values into a distribution that is substantially similar. When compared to a characteristic without any outliers, the final distribution of inliers narrows excessively in the presence of outliers [32].

4.3. Feature Selection Phase

Datasets can contain irrelevant and redundant features that impede the machine and deep learning processes. As a result, features that contribute to the machine learning process should be identified utilizing feature selection methods. Using domain knowledge and common sense, and seeking assistance from domain experts, may also aid in the detection and elimination of redundant features [33]. However, a more methodical approach is required for high-dimensional datasets. Because redundant and irrelevant characteristics have no substantial impact on machine learning, eliminating them from the learning process will aid in increasing learning speed, reducing overfitting, avoiding the curse of dimensionality, and creating simple models [34]. The standard approach for supervised feature selection consists of four steps, as shown in Figure 3.

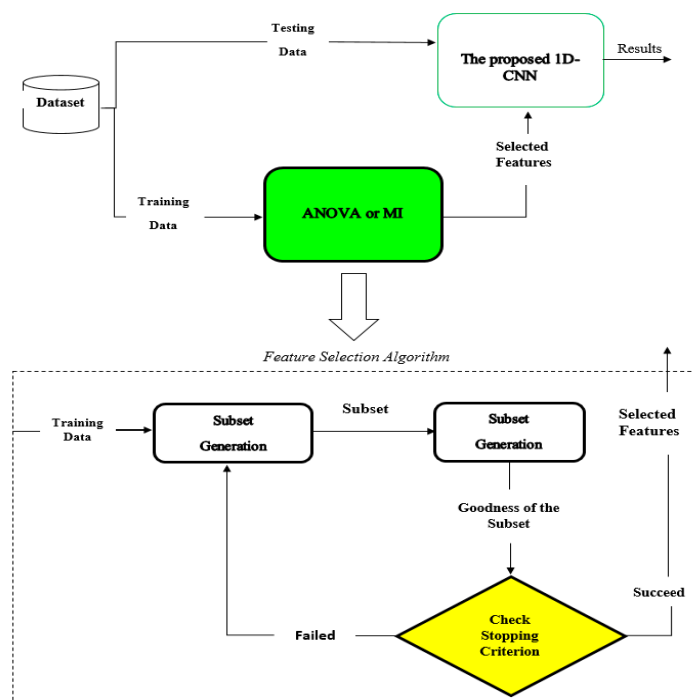


Figure 3. Steps of feature selection.

Two of the most popular feature selection methods (ANOVA and mutual information) were used in the suggested Sleep-1D-CNN model. The total number of input variables is

reduced by employing feature selection techniques to eliminate superfluous or redundant features. Then, only the features that are most relevant to the deep learning system remain in the set. The goal of feature selection in deep learning is to identify the most valuable collection of features that can be applied to build the proposed Sleep-1D-CNN model for sleep stage classification from EEG signals.

4.3.1. Select Features Based on Mutual Information (MI)

Mutual information is calculated between two variables to determine their mutual dependency. Mutual information always equals or exceeds zero. The higher the degree of mutual information, the stronger the association between the two variables. When mutual information equals 0, it indicates that the two variables are independent [35]. It is therefore connected to the entropy of a random feature, which is established by the quantity of data the feature includes. The mutual information between two discrete random variables, X and Y , is defined by Equation (2).

$$I(X; Y) = \sum_{x \in X} \sum_{y \in Y} P(x, y) \log_2 \left(\frac{P(x, y)}{p(x)p(y)} \right) \quad (2)$$

where $H(X)$ is the entropy of X , $H(X|Y)$ is the conditional entropy of X given Y , $p(x)$ and $p(y)$ are the marginal distributions of x and y , and $P(x, y)$ is the joint distribution of x and y [36].

4.3.2. Select Features Based on ANOVA

This approach is predicated on the F-test, which is limited to quantitative data. Highly correlated features receive a high score from the ANOVA F-test, which calculates the degree of linear dependency between an input variable and the target variable. Less linked features receive a low score. The F-value scores determine whether, after grouping the numerical characteristic by the goal vector, the means for each group differ significantly from one another [37]. Applying the following notation:

$$s_j^2 = \sum_{i=1}^{N_j} (X_{ij} - \bar{X}_j)^2 / (N_j - 1) \quad (3)$$

$\bar{x}$: The grand mean of predictor X :

$$\bar{x} = \sum_{j=1}^J N_j \bar{X}_j / N \quad (4)$$

The notations given above rely on pairings of (X, Y) . Next, p -value = Prob $\{F(J-1, N-J) > F\}$ yields the p -value using the F statistic, where:

$$F = \frac{\sum_{j=1}^J N_j (\bar{X}_j - \bar{X})^2 / (J - 1)}{\sum_{j=1}^J (N_j - 1) s_j^2 / (N - 1)} \quad (5)$$

4.4. Classifying Features Using Convolution Neural Network (CNN)

Convolutional neural networks (CNNs) are feedforward deep neural networks that process EEG data through three layers: convolutional, pooling, and fully connected. The convolutional layer [38] combines the tensor with the shape, whereas the pooling layer [39] simplifies calculations to minimize data dimensions. The fully connected layer [40] connects neurons from previous layers to predict or classify target goals. This stacked-layer DL technique reduces input data while capturing the distinctive spatial dependency of EEG signals. In this work, a 1D-CNN is used; for this reason, we explain it in some detail. Containing millions of parameters and several hidden layers, deep 2D-CNNs can learn complex objects and patterns if they are trained on a large visual database containing ground-truth labels. When paired with appropriate training, this unique ability makes them the primary tool for a variety of technical applications utilizing 2D signals, including

pictures and video frames [41]. In contrast to 1D signals, this might not be a feasible option in many applications, especially if the training data is constrained or application-specific [42]. 1D-CNNs were recently proposed as a solution to this issue, and they quickly achieved state-of-the-art performance levels in a number of applications, such as motor fault detection, structural health monitoring, early diagnosis and personalized biomedical data classification, anomaly detection and identification in power electronics, and structural health monitoring [43]. Another significant benefit is that while 1D-CNNs only conduct 1D convolutions (scalar multiplications and additions), their compact and straightforward design makes real-time, low-cost hardware implementation possible [44]. In this section, a one-dimensional convolutional neural network model (Sleep-1D-CNN) is developed using the Sleep-Edf dataset. The proposed Sleep-1D-CNN showed state-of-the-art performance levels in applications involving the classification of sleep stages using EEG signals, such as identifying sleep quality. The DL-EEG classification system displayed here is based on a one-dimensional convolutional neural network (EEG-1D-CNN), which is unique in that the CNN kernels can save time by simply swiping over the elements of the input pattern's single dimension during convolution. The suggested Sleep-1D-CNN model classifies EEG signals to classify sleep stages; Table 3 displays the suggested Sleep-1D-CNN layers. Following the loading, splitting, processing, and feature selection of the dataset, the data are immediately classified using the suggested Sleep-1D-CNN model. The suggested CNN model has 26 layers, which are as follows:

Table 3. The proposed Sleep-1D-CNN layers.

NO.	Layer Type	Filters	Size/Stride	Activation Function	#Param
1	Convolutional	16	3/1	—	64
2	Max Pooling	—	—	—	0
3	Leaky ReLU	—	—	—	0
4	Convolutional	32	3/1	—	1568
5	Max Pooling	—	—	—	0
6	Leaky ReLU	—	—	—	0
7	Convolutional	64	3/1	—	6208
8	Max Pooling	—	—	—	0
9	Leaky ReLU	—	—	—	0
10	Convolutional	64	3/1	—	12,352
11	Max Pooling	—	—	—	0
12	Leaky ReLU	—	—	—	0
13	Dense	32	—	Linear	4160
14	Convolutional	32	3/1	—	6176
15	Max Pooling	—	—	—	0
16	Leaky ReLU	—	—	—	0
17	Convolutional	32	3/1	—	3104

Table 3. Cont.

NO.	Layer Type	Filters	Size/Stride	Activation Function	#Param
18	Max Pooling	–	–	–	0
19	Leaky ReLU	–	–	–	0
20	Dense	32	–	Linear	1056
21	Convolutional	16	3/1	–	1552
22	Max Pooling	–	–	–	0
23	Leaky ReLU	–	–	–	0
24	Convolutional	60	3/1	–	2940
25	Flatten	–	–	–	0
26	Dense	32	–	Softmax	610,805

The following details how the 1D-CNN works with EEG signals:

1. **Input layer:** The input of the 1D-CNN is a sequence of EEG signal samples recorded from a single electrode channel. Each sample represents the electric activity measured at a specific point in time.
2. **Convolution layer:** The convolutional layers of the network apply convolutional filters to the input EEG signal. These filters move over the input signal, extracting local patterns and characteristics at varying temporal scales. Each filter looks for specific patterns in the EEG signal, like oscillatory components or transient events.
3. **Pooling Layers:** Optional pooling layers can be added after convolutional layers to downscale the feature maps generated by the convolutional layers. Pooling reduces the dimensionality of feature maps while keeping crucial information, resulting in a more computationally efficient network and less overfitting.
4. **Activation Functions:** Non-linear activation functions, such as Leaky ReLU (Rectified Linear Unit), are added to the output of convolutional and pooling layers to bring nonlinearity into the network and allow it to understand complicated data correlations.
5. **Fully Connected Layers:** After processing the EEG data with convolutional and pooling layers, the feature maps are flattened and sent through one or more fully connected layers. These layers conduct high-level feature collection and manipulation, allowing the network to acquire discriminative representations for the given task.

5. Results and Discussion

A computer with an Intel (R) Core (TM) i5-1135G7 CPU running at 2.40 GHz was used for the research (RAM: 4.00 GB; x64-based CPU; 64-bit operating system). Evaluation metrics show the classifier and how well it performed on the entire test dataset. This is an overview of the predictions made for various classification tasks according to the evaluation metrics in Equations (6)–(9) [45].

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FB} + \text{FN}} \quad (6)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (7)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (8)$$

$$F_1 = 2 * \frac{\text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (9)$$

5.1. Evaluation Model

The proposed Sleep-1D-CNN model achieved the best accuracy of 99.11% when the ANOVA-20 feature selection technique was used and 99.84% when the mutual information MI-50 feature selection technique was used. The graphical presentation shown in Figures 4–7 presents the reported accuracy, precision, recall, and F-measure for the proposed classification Sleep-1D-CNN model considered in this research work.

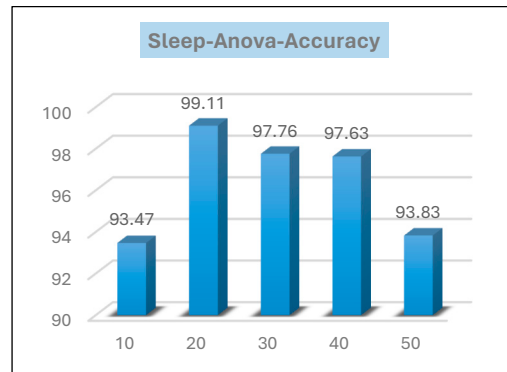


Figure 4. Accuracy results.

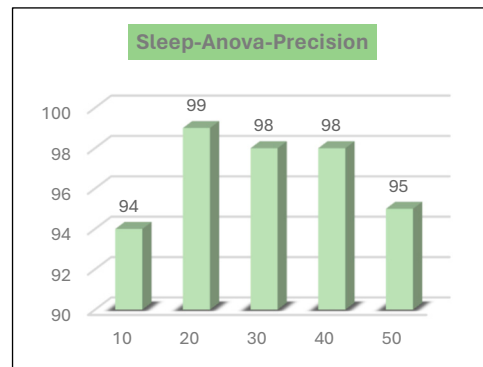


Figure 5. Precision results.

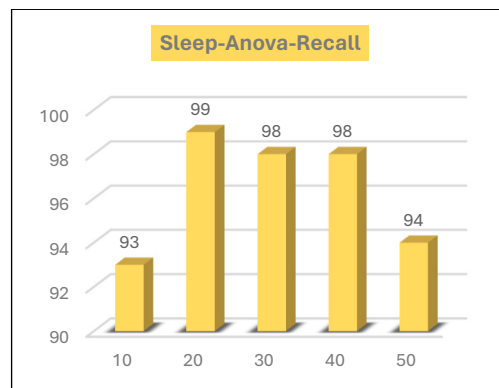


Figure 6. Recall results.

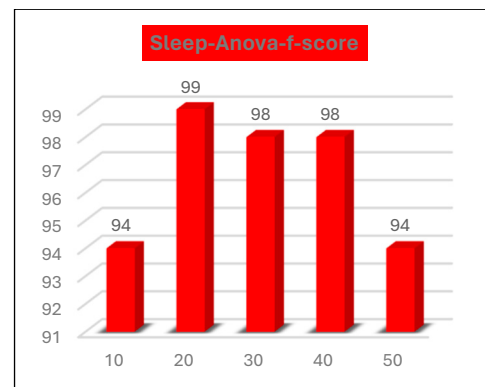


Figure 7. F-score.

Figure 8 shows the confusion matrix, which is explained in detail, and shows exactly all the results according to the five phases of sleep. In addition, Figures 9 and 10 show the accuracy curve and the loss function.

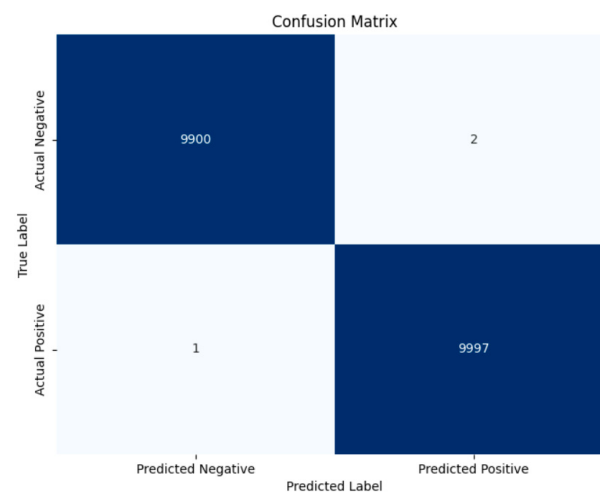


Figure 8. Confusion matrix of results (TP = 9900, TN = 9997, FN = 1, TN = 2).

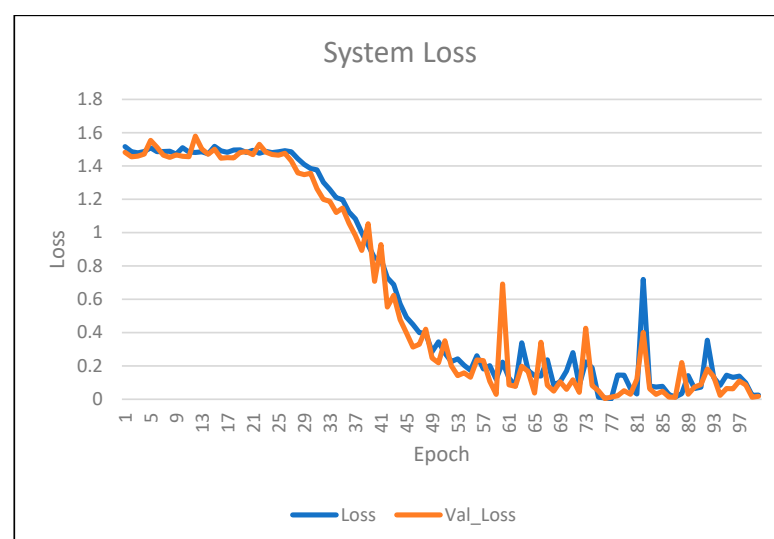


Figure 9. Loss function curve.

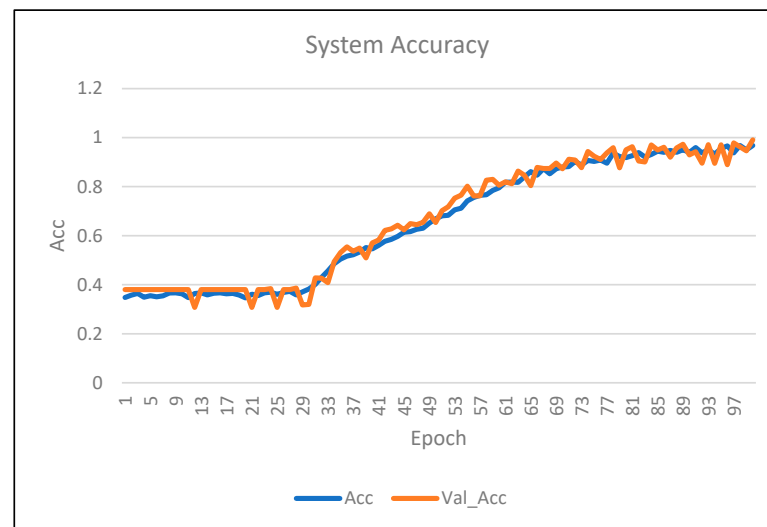


Figure 10. Accuracy curve.

5.2. Comparative Analysis

We examined the outcomes of the suggested Sleep-1D-CNN model and other contributing approaches that were tested on the same dataset to validate our suggested methodology. Table 4 display the full analysis.

Table 4. Comparison results.

Ref.	Best Accuracy
Yildirim et al. [13]	94.24
Nguyen et al. [14]	87.67
Z. Mousavi et al. [15]	90
T. Zhu et al. [16]	93.7
S. Santaji and V. Desai [17]	97.8
Z. Cui et al. [18]	83.6
H. Phan et al. [19]	87.4
S. Satapathy et al. [20]	87
The proposed Sleep-1D-CNN-MI-50	99.84

6. Conclusions

In this paper, we presented a deep learning-based classification model for sleep staging. Seventy-three percent of training and testing data were used to train the suggested model utilizing sleep-related diseases and healthy control patients. In five scenarios (10, 20, 30, 40, and 50), the suggested model was evaluated utilizing two feature selection strategies, ANOVA and mutual information, using a proposed deep learning model called Sleep-1D-CNN. The efficacy of the model was evaluated using several performance metrics, including F-score, recall, accuracy, and precision. With MI-50 having the best accuracy, the suggested Sleep-1D-CNN achieved 99.84% accuracy. 1D-CNN-MI-50 achieved the maximum accuracy of 99.84%, precision of 100%, recall of 100%, and F-score of 100% when compared to state-of-the-art works.

Our proposed system achieved impressive results; however, there is always room for improvement and further research. For future work, some processing methods can be applied, such as in experimentation with different CNN architectures, including variations in the number of layers, type of layers (e.g., convolution, pooling, and dropout), and connectivity patterns (e.g., ResNet, Inception, DenseNet). Exploring deeper or wider

architectures to capture more complex patterns in data can be considered. The hyperparameters of the model, such as learning rate, batch size, optimizer choice, dropout rate, and regularization strength, can be further fine-tuned. Finally, techniques like grid search or random search can be used to systematically explore the hyperparameter space and identify optimal configurations.

Author Contributions: Conceptualization, M.R.M. and A.M.S.; methodology, M.R.M. and A.M.S.; software, M.R.M.; validation, M.R.M. and A.M.S.; formal analysis, M.R.M.; investigation, M.R.M. and A.M.S.; data curation, M.R.M.; writing—original draft preparation, M.R.M. and A.M.S.; writing—review and editing, M.R.M.; supervision, M.R.M.; project administration, M.R.M.; funding acquisition, M.R.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: <https://www.physionet.org/content/sleep-edf/1.0.0/> (accessed on 1 May 2024).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Phan, H.; Mikkelsen, K.; Chén, O.Y.; Koch, P.; Mertins, A.; De Vos, M. Sleep Transformer: Automatic Sleep Staging with Interpretability and Uncertainty Quantification. *IEEE Trans. Biomed. Eng.* **2022**, *69*, 2456–2467. [CrossRef] [PubMed]
2. Kumar, R.R.; Rithesh, A.; Josh, P.; Raj, B.K.; John, V.; Prasad, D.S. Sleep Track: Automated Detection and Classification of Sleep Stages. *E3S Web Conf.* **2023**, *430*, 01020. [CrossRef]
3. Almutairi, H.; Hassan, G.M.; Datta, A. Classification of Obstructive Sleep Apnoea from Single-Lead ECG Signals Using Convolutional Neural and Long Short Term Memory Networks. *Biomed. Signal Process. Control* **2021**, *69*, 102906. [CrossRef]
4. Dauvilliers, Y.; Schenck, C.H.; Postuma, R.B.; Iranzo, A.; Luppi, P.H.; Plazzi, G.; Montplaisir, J.; Boeve, B. REM sleep behaviour disorder. *Nat. Rev. Dis. Primers* **2018**, *4*, 19. [CrossRef]
5. Almutairi, H.; Hassan, G.; Datta, A. Classification of sleep stages from EEG, EOG and EMG signals by SSNet. *arXiv* **2023**, arXiv:2307.05373v1.
6. Peever, J.; Fuller, P.M. The Biology of REM Sleep. *Curr. Biol.* **2017**, *27*, R1237–R1248. [CrossRef] [PubMed]
7. Timplalexis, C.; Diamantaras, K.; Chouvarda, I. Classification of Sleep Stages for Healthy Subjects and Patients with Minor Sleep Disorders. In Proceedings of the IEEE 19th International Conference on Bioinformatics and Bioengineering, Athens, Greece, 28–30 October 2019.
8. Faust, O.; Razaghi, H.; Barika, R.; Ciaccio, E.J.; Acharya, U.R. A review of automated sleep stage scoring based on physiological signals for the new millennia. *Comput. Methods Programs Biomed.* **2019**, *176*, 81–91. [CrossRef] [PubMed]
9. Sri, T.R.; Madala, A.J.; Duddukuru, S.L.; Reddipalli, R.; Polasi, P.K. A Systematic Review on Deep Learning Models for Sleep Stage Classification. In Proceedings of the 6th International Conference on Trends in Electronics and Informatics, Tirunelveli, India, 28–30 April 2022.
10. Fernandez-Blanco, E.; Rivero, D.; Pazos, A. EEG Signal Processing with Separable Convolutional Neural Network for Automatic Scoring of Sleeping Stage. *Neurocomputing* **2020**, *410*, 220–228. [CrossRef]
11. Yang, B.; Zhu, X.; Liu, Y.; Liu, H. A single-channel EEG based automatic sleep stage classification method leveraging deep one-dimensional convolutional neural network and hidden Markov model. *Biomed. Signal Process. Control* **2021**, *68*, 102581. [CrossRef]
12. Jiang, X.; Zhao, J.; Du, B.; Yuan, Z. Self-supervised contrastive learning for EEG-based sleep staging. In Proceedings of the 2021 International Joint Conference on Neural Networks (IJCNN), Virtual, 18–22 July 2021; IEEE: New York, NY, USA, 2021; pp. 1–8.
13. Yildirim, O.; Baloglu, U.B.; Acharya, U.R. A deep learning model for automated sleep stages classification using PSG signals. *Int. J. Environ. Res. Public Health* **2019**, *16*, 599. [CrossRef]
14. Rajbhandari, E.; Alsadoon, A.; Prasad, P.; Seher, I.; Nguyen, T.Q.V.; Pham, D.T.H. A novel solution of enhanced loss function using deep learning in sleep stage classification: Predict and diagnose patients with sleep disorders. *Multimed. Tools Appl.* **2021**, *80*, 11607–11630. [CrossRef]
15. Mousavi, Z.; Rezaei, T.Y.; Sheykhivand, S.; Farzamnia, A.; Razavi, S. Deep convolutional neural network for classification of sleep stages from single-channel EEG signals. *J. Neurosci. Methods* **2019**, *324*, 108312. [CrossRef] [PubMed]
16. Zhu, T.; Luo, W.; Yu, F. Convolution-and attention-based neural network for automated sleep stage classification. *Int. J. Environ. Res. Public Health* **2020**, *17*, 4152. [CrossRef] [PubMed]
17. Santaji, S.; Desai, V. Analysis of EEG Signal to Classify Sleep Stages Using Machine Learning. *Sleep Vigil.* **2020**, *4*, 145–152. [CrossRef]
18. Cui, Z.; Zheng, X.; Shao, X.; Cui, L. Automatic sleep stage classification based on convolutional neural network and fine-grained segments. *Complexity* **2018**, *2018*, 9248410. [CrossRef]
19. Phan, H.; Andreotti, F.; Cooray, N.; Chen, O.Y.; De Vos, M. Joint classification and prediction CNN framework for automatic sleep stage classification. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 1285–1296. [CrossRef] [PubMed]

20. Satapathy, S.K.; Kondaveeti, H.K.; Sreeja, S.R.; Madhani, H.; Rajput, N.; Swain, D. A Deep Learning Approach to Automated Sleep Stages Classification Using Multi-Modal Signals. *Procedia Comput. Sci.* **2023**, *218*, 867–876. [\[CrossRef\]](#)
21. Lee, M.; Kwak, H.G.; Kim, H.J.; Won, D.O.; Lee, S.W. SeriesSleepNet: An EEG time series model with partial data augmentation for automatic sleep stage scoring. *Front. Physiol.* **2023**, *14*, 1188678. [\[CrossRef\]](#)
22. Masad, I.; Alqudah, A.; Qazan, S. Automatic classification of sleep stages using EEG signals and convolutional neural networks. *PLoS ONE* **2024**, *19*, e0297582. [\[CrossRef\]](#)
23. Al-Salman, W.; Li, Y.; Oudah, A.Y.; Almaged, S. Sleep stage classification in EEG signals using the clustering approach based probability distribution features coupled with classification algorithms. *Neurosci. Res.* **2023**, *188*, 51–67. [\[CrossRef\]](#)
24. Kadhim, A. An evaluation of preprocessing techniques for text classification. *Int. J. Comput. Sci. Inf. Secur.* **2018**, *16*, 1947–5500.
25. Kubinski, R.; Djamen-Kepaou, J.Y.; Zhanabaev, T.; Hernandez-Garcia, A.; Bauer, S.; Hildebrand, F.; Korcsmaros, T.; Karam, S.; Jantchou, P.; Kafi, K.; et al. Benchmark of data processing methods and machine learning models for gut microbiome-based diagnosis of inflammatory bowel disease. *Front. Genet.* **2022**, *13*, 784397. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Amato, A.; Di Lecce, V. Data preprocessing impact on machine learning algorithm performance. *Open Comput. Sci.* **2023**, *13*, 20220278. [\[CrossRef\]](#)
27. Vrigazova, B. The Proportion for Splitting Data into Training and Test Set for the Bootstrap in Classification Problems. *Bus. Syst. Res.* **2021**, *12*, 228–242. [\[CrossRef\]](#)
28. Nguyen, Q.H.; Ly, H.-B.; Ho, L.S.; Al-Ansari, N.; Van Le, H.; Tran, V.Q.; Prakash, I.; Pham, B.T. Influence of Data Splitting on Performance of Machine Learning Models in Prediction of Shear Strength of Soil. *Math. Probl. Eng.* **2021**, *2021*, 4832864. [\[CrossRef\]](#)
29. Muraina, I. Ideal Dataset Splitting Ratios in Machine Learning Algorithms: General Concerns for Data Scientists and Data Analysts. In Proceedings of the 7th International Mardin Artuklu Scientific Researches Conference, Mardin, Turkey, 13–15 December 2022.
30. Dzierżak, R. Comparison of the Influence of Standardization and Normalization of Data on the Effectiveness of Spongy Tissue Texture Classification. *Inform. Autom. Pomiar W Gospod. I Ochr. Środowiska* **2019**, *9*, 66–69. [\[CrossRef\]](#)
31. Raju, V.N.G.; Lakshmi, K.P.; Jain, V.M.; Kalidindi, A.; Padma, V. Study the Influence of Normalization/Transformation process on the Accuracy of Supervised Classification. In Proceedings of the 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 20–22 August 2020; IEEE: New York, NY, USA, 2020; pp. 729–735.
32. de Amorima, L.; Cavalcantia, G.; Cruz, R. The choice of scaling technique matters for classification performance. *arXiv* **2022**, arXiv:2212.12343v1. [\[CrossRef\]](#)
33. da Costa, N.L.; de Lima, M.D.; Barbosa, R. Analysis and improvements on feature selection methods based on artificial neural network weights. *Appl. Soft Comput.* **2022**, *127*, 109395. [\[CrossRef\]](#)
34. Hossain, M.; Islam, M. A novel hybrid feature selection and ensemble-based machine learning approach for botnet detection. *Sci. Rep.* **2023**, *13*, 21207. [\[CrossRef\]](#)
35. Barraza, N.; Moro, S.; Ferreyra, M.; de la Peña, A. Mutual information and sensitivity analysis for feature selection in customer targeting: A comparative study. *J. Inf. Sci.* **2019**, *45*, 53–67. [\[CrossRef\]](#)
36. Macedo, F.; Valadas, R.; Carrasquinha, E.; Oliveira, M. Feature selection using Decomposed Mutual Information Maximization. *Neurocomputing* **2022**, *513*, 215–232. [\[CrossRef\]](#)
37. Pereira, G.; dos Santos, M.; Carvalho, A. Evaluating Meta-Feature Selection for the Algorithm Recommendation Problem. *arXiv* **2021**, arXiv:2106.03954v2.
38. Sarker, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Comput. Sci.* **2021**, *2*, 160. [\[CrossRef\]](#)
39. Janiesch, C.; Zschech, P.; Heinrich, K. Machine learning and deep learning. *Electron. Mark.* **2021**, *31*, 685–695. [\[CrossRef\]](#)
40. Ahmad, J.; Farman, H.; Jan, Z. Deep learning methods and applications. In *Deep Learning: Convergence to Big Data Analytics*; Springer Briefs in Computer Science; Springer: Singapore, 2019; pp. 31–42.
41. Mattioli, F.; Porcaro, C.; Baldassarre, G. A 1D CNN for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface. *J. Neural Eng.* **2022**, *18*, 066053. [\[CrossRef\]](#)
42. Azizjon, M.; Jumabek, A.; Kim, W. 1D CNN based network intrusion detection with normalization on imbalanced data. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIIC), Fukuoka, Japan, 19–21 February 2020; pp. 218–224.
43. Qazi, E.; Almorjan, A.; Zia, T. A One-Dimensional Convolutional Neural Network (1D-CNN) Based Deep Learning System for Network Intrusion Detection. *Appl. Sci.* **2022**, *12*, 7986. [\[CrossRef\]](#)
44. Singh, K. 1D-CNN based Model for Classification and Analysis of Network Attacks. *Int. J. Adv. Comput. Sci. Appl.* **2021**, *12*, 604–613. [\[CrossRef\]](#)
45. Goodwin, M.; Halvorsen, K.T.; Jiao, L.; Knausgård, K.M.; Martin, A.H.; Moyano, M.; Oomen, R.A.; Rasmussen, J.H.; Sørvalen, T.K.; Thorbjørnsen, S.H. Unlocking the potential of deep learning for marine ecology: Overview, applications, and outlook. *ICES J. Mar. Sci.* **2022**, *79*, 319–336. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.