TOBB UNIVERSITY OF ECONOMICS AND TECHNOLOGY
BIOMEDICAL ENGINEERING

BME 498 DESIGN PROJECT

Sleep Stage Classification with Machine Learning Algorithms

Feyza ISKAR
Rabia KAŞIKCI

Advisor: Prof. Dr. Osman EROĞUL
Spring Term of 2021

This report was reviewed and found appropriate to meet, as a minimum, the requirements of BME 498 Design Project.

Date: 20 / April / 2021

_____

Prof. Dr. Osman EROĞUL

Project Advisor

_____

Prof. Dr. Osman EROĞUL

Department Chair

ACKNOWLEDGEMENT

We would like to thank our thesis supervisor, Prof. Dr. Osman EROĞUL, for his assistance throughout the study. In addition, we thank Farhad Nassehi, one of our lecturer, for his understanding and endless support.

Finally, we would like thank our family and our friends for their support.

Feyza ISKAR

Rabia KAŞIKCI

ÖZ

Bitirme Tasarım Projesi Tezi

MAKİNE ÖĞRENME ALGORİTMALARI KULLANILARAK UYKU EVRELERİNİN SINIFLANDIRILMASI

Rabia Kaşıkcı, Feyza Iskar

TOBB Ekonomi ve Teknoloji Üniversitesi

Mühendislik Fakültesi

Biyomedikal Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Osman Eroğul

Tarih: Nisan 2021

Uyku, insanların en önemli ihtiyaçlarından biridir ve REM, non-REM1, non-REM2 ve non-REM3 aşamalarından oluşur. Uykunun kalitesi, süresi ve zamanlamasındaki değişiklikleri doğru saptamak, uyku bozukluklarının tanısında önemlidir. Bu nedenle uyku evrelerini gözlemleyebilmek önemlidir. Uyku çalışmaları, bütün gece süren polisomnografi (PSG) kayıtları ile mümkün olmaktadır. Uyku aşaması sınıflandırması, uyku uzmanları tarafından PSG kayıtları üzerinden elle yapılır. Bu süreç zaman alıcı ve zahmetlidir. Bu sebeple otomatik uyku evresi sınıflandırma yöntemleri önem arz eder. Bu çalışmanın amacı, tek kanal EEG sinyali kullanarak yüksek doğrulukta otomatik uyku evresi sınıflandırması yapmaktır. Bu amaçla, Sleep-EDF Veritabanı (Genişletilmiş)'ndan rastgele 10 kişi seçilmiştir. 20 adet zaman ve frekans bölgesi özniteliği kullanılmıştır. Makine öğrenmesi algoritmaları olarak Destek Vektör Makinesi (DVM), Karar Ağacı (KA), K-En Yakın Komşu (K-EYK) ve Yapay Sinir Ağları (YSA) kullanılmıştır ve sırasıyla genel başarı oranları, %84.63, %79.33, %79.30 ve %86.24 olarak elde edilmiştir. Tüm evreler için en yüksek başarı oranına Yapay Sinir Ağları ile ulaşılmıştır ve doğruluk oranları non-REM1, non-REM2, non-REM3, REM ve uyanıklık için sırasıyla %95.37, %90.04, %96.67, %92.44 ve %97.96 olarak elde edilmiştir. Makine öğrenmesi algoritmaları ile uyku uzmanlarının işini kolaylaştırabilecek bir otomatik uyku evresi sınıflandırma yöntemi geliştirilmiştir.

Anahtar Kelimeler: Polisomnografi, Makine öğrenmesi, Destek Vektör Makinesi (DVM), Karar Ağacı (KA), K-En Yakın Komşu (K-EYK), Yapay Sinir Ağları (YSA)

ABSTRACT

Senior Design Project Thesis

SLEEP STAGE CLASSIFICATION WITH MACHINE LEARNING ALGORITHMS

Rabia Kaşıkcı, Feyza Iskar

TOBB University of Economics and Technology

Faculty of Engineering

Biomedical Engineering Science Program

Supervisor: Prof. Dr. Osman Eroğul

Date: April 2021

Sleep is one of the most essential needs of people and consists of REM, non-REM1, non-REM2 and non-REM3 stages. Correctly detecting changes in the quality, duration and timing of sleep is significant in the diagnosis of sleep disorders. For this reason, it is important to be able to observe the sleep stages. Sleep studies are possible with all-night polysomnography (PSG) recordings. Sleep stage classification is made manually by sleep experts over PSG records. This process is time consuming and laborious. Therefore, automatic sleep stage classification methods are substantial. The aim of this study is to make high accuracy automatic sleep stage classification using a single channel EEG signal. For this purpose, 10 people were randomly selected from the Sleep-EDF Database (Extended). 20 time and frequency domain features were used. Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbor (KNN) and Artificial Neural Networks (ANN) were used as machine learning algorithms and the average success rates were obtained as 84.63%, 79.33%, 79.30% and 86.24% respectively. The highest accuracy rates for all stages have been achieved with Artificial Neural Networks algorithm and the accuracy rates are obtained as 95.37%, 90.04%, 96.67%, 92.44% and 97.96% for non-REM1, non-REM2, non-REM3, REM and wakefulness, respectively. With machine learning algorithms, an automated sleep stage classification method has been developed that can facilitate the work of sleep experts.

Keywords: Polysomnography, Machine learning, Support Vector Machine (SVM), Decision Tree (DT), K-Nearest Neighbor (K-NN), Artificial Neural Networks (ANN)

TABLE OF CONTENTS

Pages

LIST OF FIGURES

Pages

LIST OF TABLES

Pages

LIST OF ABBREVIATIONS

PSG	: Polysomnography
EEG	: Electroencephalography
EOG	: Electrooculography
ECG	: Electrocardiography
EMG	: Electromyography
REM	: Rapid Eye Movements
NREM	: Non-Rapid Eye Movements
DWT	: Discrete Wavelet Transform
ZC	: Zero Crossing
PFD	: Petrosian Fractal Dimension
SD	: Standart Deviation
RMS	: Root Mean Square
MAD	: Mean Absolute Deviation
IQR	: Interquartile Range
SVM	: Support Vector Machines
KNN	: K-Nearest Neighbor
DT	: Decision Tree
CART	: Classification and Regression Trees
ANN	: Artificial Neural Network
TN	: True Negative
TP	: True Positive
FN	: False Negative
FP	: False Positive

LIST OF SYMBOLS

Symbols	Definition
α	: Alpha waves
β	: Beta waves
γ	: Gamma waves
θ	: Theta waves
δ	: Delta waves
N	: Signal length
Nσ	: The number of changes in the signal derivative
w	: Weight vector
b	: Bias
a	: Artificial neural network output vector
p_i	: The ratio of states reaching the node

1. INTRODUCTION

1.1. Sleep

Sleep is one of the fundamental need in human life which has an important role on the physical recovery, memory, emotion and health. It is a dynamic and complex process that has been tried to be understood with many studies. [1] [Url-1]

Sleep and wake states are regulated by significant brain structures and chemicals. Natural daily rhythms of the body is called circadian rhythm or internal clock. Circadian rhythm controls biological fluctuations in a day, such as blood pressure, body temperature, metabolism and release of hormones. Moreover, it controls the timing of sleep, induces to sleep at night and to wake in the morning without alarm. Circadian rhythm synchronizes with environmental conditions as light and temperature during the day, however it continues even in the absence of these conditions.

One of the significant hormone that contributes the circadian rhythm is melatonin. Release of melatonin increases at night and drops after dawn. This hormone causes drowsiness. [Url-1] [Url-2]

The hypothalamus is comprised of groups of nerve cells which control sleep and arousal. The suprachiasmatic nucleus (SCN) is placed in the hypothalamus and consists of a large number of cell clusters that receive information from the eyes about light exposure. These cell clusters also controls behavioral rhythm.

The brain stem controls the transition between sleep and wake states by communicating with the hypothalamus. There are cells in the hypothalamus and the brain stem that promotes sleep and produces a brain chemical that called GABA. This chemical is responsible for reducing the activity of arousal centers within the brain stem and the hypothalamus. Besides, the brain stem plays a significant role in REM sleep by relaxing the muscles.

The thalamus transfers information from the senses to the cerebral cortex which interprets and processes information from short-term memory. In REM sleep, the thalamus becomes active and sends images, sounds and other sensations to the cortex in order to animate dreams.

The pineal gland, raises melatonin production by receving signals from the SCN. Another structure promotes sleep and arousal is the basal forebrain, which is part of the midbrain, works as arousal system. Cells in the basal forebrain release adenosine that supports sleep drive. The amygdala involves in emotion processing and becomes active during REM stage. [Url-1]

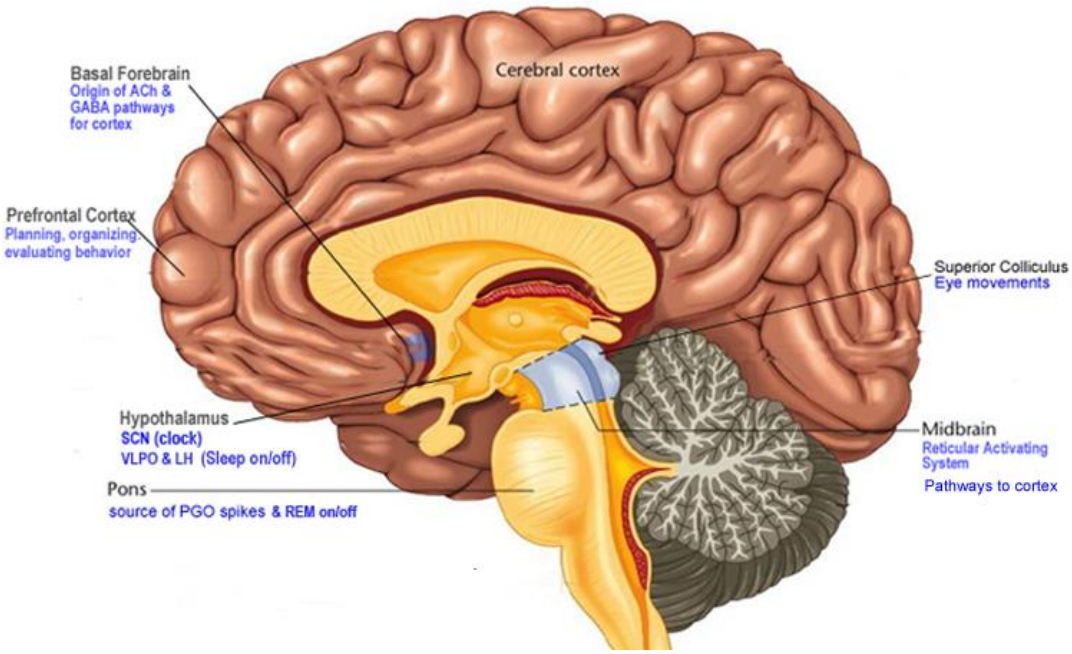

Figure 1.1. Brain Structures Involved in Sleep [Url-3]

1.2. Polysomnography (PSG) and Electroencephalogram (EEG)

Sleep researches are possible with whole night Polysomnography (PSG) recordings. PSG is a type of sleep study that records Electroencephalogram (EEG), Electrocardiogram (ECG), Electrooculography (EOG), Electromyogram (EMG), thoracic and abdominal movements, airflow and oximetry signals from various sensors. In sleep medicine, PSG is used as a diagnostic tool. Sleep specialists use PSG to study sleep disorders.

The EEG recordings can be used to analyse the neural activity dynamics which varies during sleep. [2]

EEG is the electrophysiological recording of the electrical activity on the scalp with the help of metal electrodes. The neurons communicate with each other by way of generating electrical currents regardless of the person is awake or sleeping. EEG is non-invasive, used easily and portable. EEG has low cost, high temporal solution and safe nature. These advantages makes EEG a powerful tool to view brain comparing to other neuroimaging methods. [3]

The electrodes used to obtain the EEG signal have different placement positions on the head that shown in Figure 1.2.

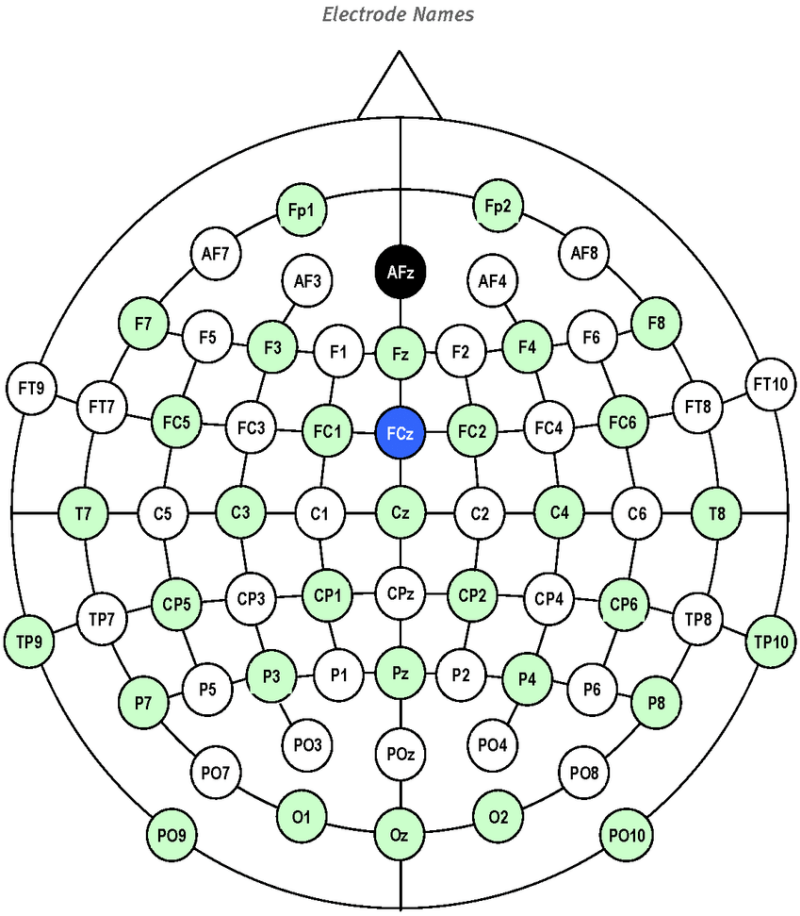

Figure 1.2. EEG Electrode Placement [4]

The sleep stage classification rules are standardized by Rechtschaffen and Kales. These rules depend on modifications of EEG and four non-REM stages. [5]

1.3. Sleep Stages and Related Brain Waves

Sleep is a nonstationary process that have time-varying statistics. Nevertheless, it is considered stationary within a short period of time called sleep stage. [2]

Basicly, sleep consists of two stages which are called REM and non-REM. Non-REM includes 4 stages (N1,N2,N3,N4). However, according to the AASM standardization rules, N3 and N4 were counted one stage as N3, also known as slow wave sleep. [6]

Brain waves change in different stages of sleep due to changing brain activity. Frequency range of the waves are shown in Table 1.1.

Table 1.1. Frequency Range of the Brain Waves [7]

Brain Waves	Frequency Range (Hz)
Delta (δ)	0.5-4
Theta (θ)	4-8
Alpha (α)	8-13
Beta (β)	13-35
Gamma (γ)	>35

Brain waves synchronize according to electrical activity and become dominant. Beta and gamma waves have high frequency and low amplitudes and are associated with wakefulness. Alpha waves have lower frequency, high amplitude patterns of brain activity. Theta waves have lower frequency and higher amplitude than alpha waves. Delta waves are brain waves that have very low frequency and high amplitude. Brain waves are shown in Figure 1.3.

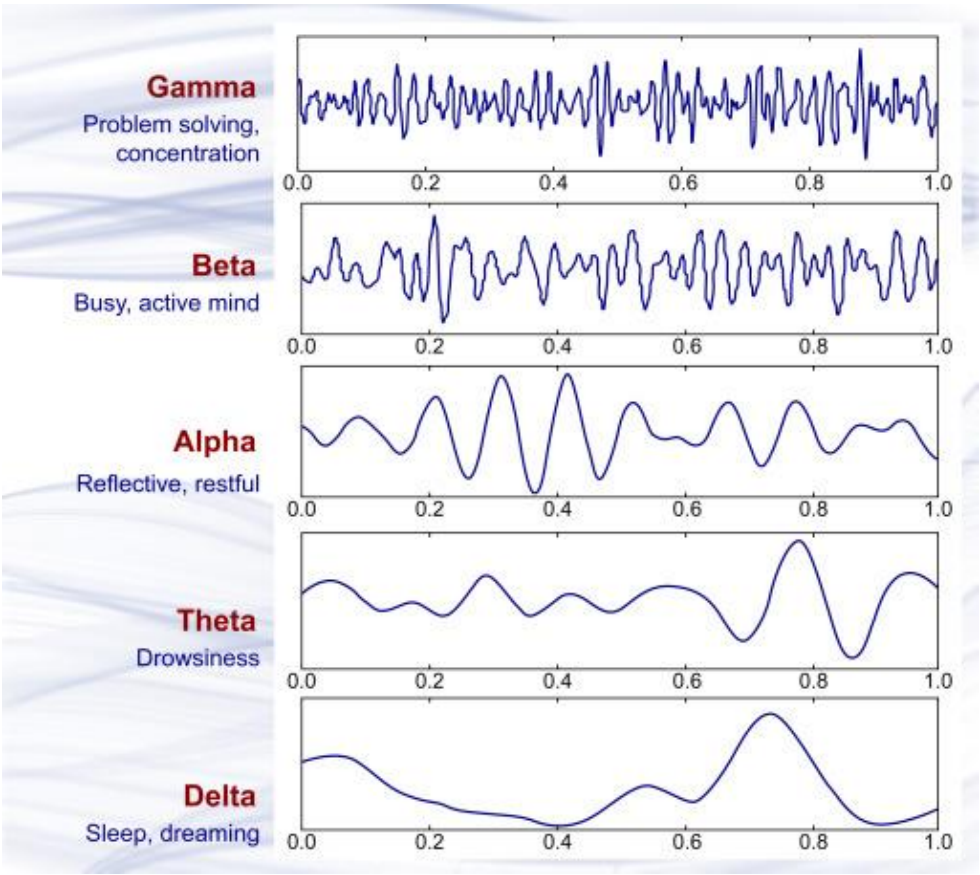

Figure 1.3. Representation of Brain Waves [8]

1.3.1. Stages of Sleep

1.3.1.1. Non-REM 1

The first stage of non-REM sleep is stage 1 (N1). N1 is occurs between sleep and wakefulness. During this period, respiration and heartbeat rates diminish, body temperature starts to decrease, muscles relax. In stage 1 sleep, people lose their awareness.

With regards to brain waves, alpha and theta waves are dominant in N1. The early part of the stage 1 includes alpha waves that mean a person who is relaxed but still awake. In the later part of stage 1, alpha waves are replaced by theta waves. It is easier to wake in stage 1 than waking in other stages.

1.3.1.2. Non-REM 2

In the second stage of non-REM sleep, deeper relaxation occurs. The breathing and heart rate become more slower. Electrical activity of brain becomes irregular.

In N2, theta waves are dominant and interrupted by sleep spindles which are high frequency sections continue in short time. It is thought that sleep spindles might have an important role in memory and learning. Likewise, N2 is associated with K-complexes. K-complexes are high amplitude patterns in response to the stimuli from surroundings. Scientists consider k-complexes as a waking system when needed. N2 maintains nearly half of the night. [Url-2] [Url-4]

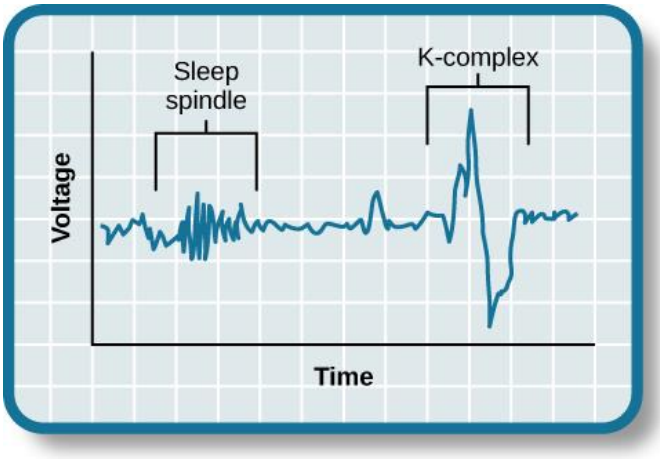

Figure 1.4. Representation of Sleep Spindle and K-complex [Url-4]

1.3.1.3. Non-REM 3

The third stage of non-REM sleep is referred to as slow-wave sleep or deep sleep. The heart rate and respiration decreases dramatically and breathing becomes regular. Blood pressure drops.

At this stage, body repairs itself. The pituitary gland releases growth hormone that stimulates muscle repair and tissue growth at the beginning of N3. Researchers have found that blood levels of substances which activate immune system increased at this stage.

Delta waves are dominated in this stage. It is the hardest stage to awaken a person due to reduced sensitivity of the brain to external stimuli. N3 constitutes 20% of the sleep time with the periods of up to 30 minutes. At first half of the night, longer periods of N3 occur. Most people over the age of 65 have lack of deep sleep.

1.3.1.4. Rapid Eye Movement (REM)

In REM sleep, body temperature and blood pressure rise, heart rate and respiration accelerate. Brain activity rises and the brain waves show similar patterns with the wakefulness. REM sleep is associated with rapid movements of the eyes behind closed lids.

At this stage, muscle systems are paralyzed except eye muscles and muscles that maintain circulation and respiration. Voluntary muscles movement cannot occur during REM stage. REM sleep is known as paradoxical sleep in consequence of absence of muscle movements and high brain activity.

The most vivid dreams occurs in REM stage which may be the result of increased brain activity. In any sleep stage, dreams can occur. However, in the non-REM stages, dreams are less intense and common.

The first REM stage lasts a few minutes, yet later REM stages might last about an hour at the second half of the night. REM sleep generates 25% of sleep in healthy adults. REM sleep is thought to play a specific role in creativity, learning and memory. [Url-1] [Url-2] [Url-4]

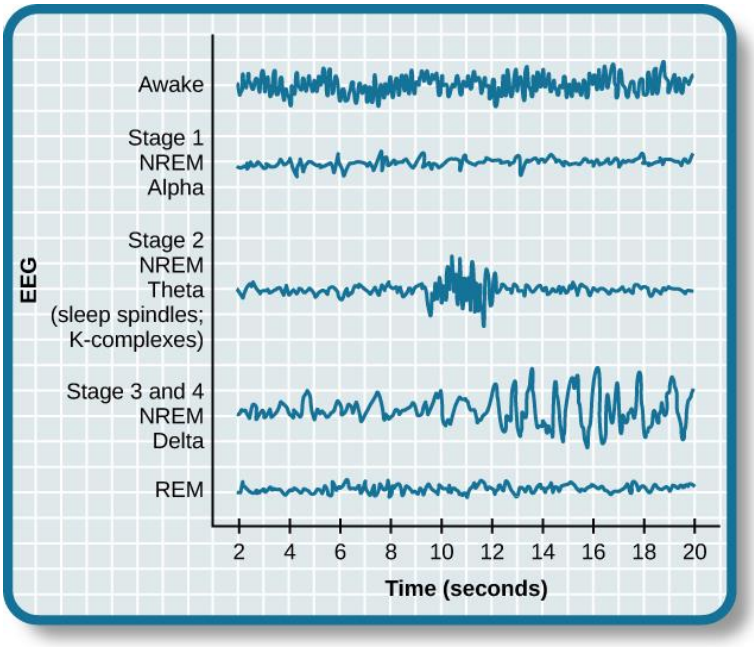

Figure 1.5. Changing Brain Wave Activity During Different Stages of Sleep [Url-4]

1.3.2. Sleep Architecture

Sleep is a cyclic process. Stages of sleep differs between a predictible range. Hypnogram is a diagram that represents sleep stages across hours of sleep. Sleep specialists entitle the alternating stages of sleep as sleep architecture. [Url-2]

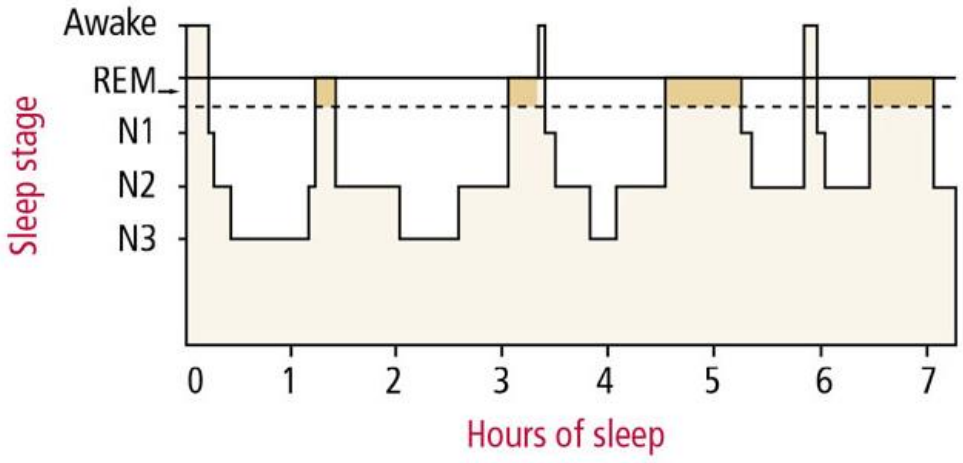

Figure 1.6. Sleep Architecture [Url-2]

Table 1.2. Duration of Sleep Stages [Url-5]

Sleep Stage	Lenght of the Stage (minutes)
N1	1-5
N2	10-60
N3	20-40
REM	10-60

1.4. Automated Sleep Stage Classification

Sleep scoring is an important diagnostic method in neurology and psychiatry. [5] Classification of sleep stages are made by a sleep specialist manually. Sleep expert marks and scores the PSG recordings that includes multichannel signals. This process is time-consuming, difficult and human error might occur during scoring.

Hence, there have been a various researches on the purpose of generating an automatic sleep stage classification method. An automatic sleep stage classification method generally consists of three sections as preprocessing, feature extraction and classification. [9]

1.5. Aim of the Study

The aim of this study is to determine the stages of sleep in a computer environment by examining a person's sleep signals with the help of machine learning. The Sleep-EDF Database Expanded containing EEG, EOG EMG signals of 197 patients was obtained from the internet. Only the EEG signal was used for this study. Ten patients were selected randomly among 19 patients and PSG.edf files were extracted as a signal with the help of Matlab. HPG.edf files, the equivalent of PSG.edf files scored by knowledgeable people, are also obtained from this database. Since only the sleep stages were concerned in this study, the sleep times of the patients to be examined in the study were determined and the signals in this time interval were used.

Normalization, outlier removal and bandpass processes were applied as signal preprocessing steps, since it is known that the signal needs to go through some processes in order to provide machine learning. After these procedures, feature extraction steps were applied from the clean EEG signal obtained.

The differences in the features of the EEG signal were observed and the features that were thought to give the best performance were obtained from the EEG signal. The resulting feature matrices was separated to be used as training and test matrices for the machine learning algorithm.

As a result of the researches, it was deemed appropriate to use k nearest neighbor (KNN), support vector machines (SVM), decision trees algorithm (DT) and artificial neural networks (ANN) structures for the classification algorithm. It has been tried to compare and observe the algorithm systems used, and it has been tried to determine the algorithms that can give good results with their accuracy in the classification of sleep stages in a system using EEG (single channel). Finally, in addition to the purpose of the study, the factors affecting the classification were evaluated.

2. MATERIAL AND METHOD

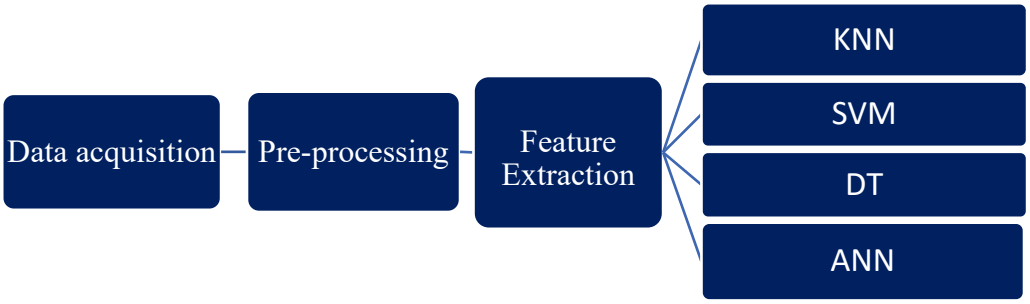

Figure 2.1. Flowchart of This Study

2.1. Collecting Data

In this study, the data set named Sleep-EDF Database (Expanded) was used. This database includes 197 patients' recording of EEG EOG EMG signals throughout the day. Some data in these records include body temperature and respiration. Sleep stages were also scored manually by knowledgeable researchers according to the Rechtschaffen and Kales guideline and transferred into the data. [10]

While PSG.edf files in the data set represent the biological signals received from patients as shown in Figure 2.2. Hypnogram.edf files contain sleep scoring corresponding to the PSG.edf file information as shown in Figure 2.3. These scores represent the sleep stages and are indicated as W, R, 1, 2, 3, 4 ve ? (not scored). [Url-6]

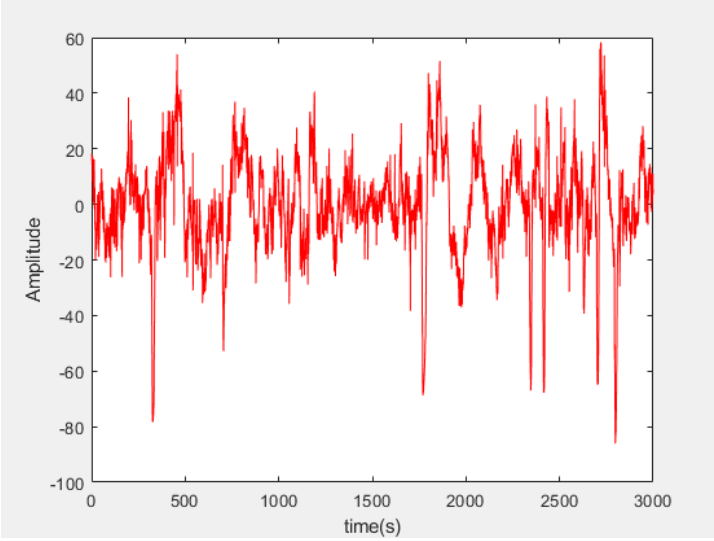

Figure 2.2. Signal Output for 1 Epoch (30s) Obtained from the PSG.edf File

12060 sec	"Sleep stage W"	19740 sec
31800 sec	"Sleep stage 1"	240 sec
32040 sec	"Sleep stage 2"	870 sec
32910 sec	"Sleep stage 3"	30 sec
32940 sec	"Sleep stage 2"	120 sec
33060 sec	"Sleep stage 3"	30 sec
33090 sec	"Sleep stage 2"	30 sec
33120 sec	"Sleep stage 3"	120 sec
33240 sec	"Sleep stage 2"	30 sec
33270 sec	"Sleep stage 3"	90 sec
33360 sec	"Sleep stage 2"	30 sec
33390 sec	"Sleep stage 3"	90 sec
33480 sec	"Sleep stage 4"	210 sec
33690 sec	"Sleep stage 3"	300 sec
33990 sec	"Sleep stage 4"	30 sec
34020 sec	"Sleep stage 3"	30 sec
34050 sec	"Sleep stage 4"	30 sec

Figure 2.3. An Example of a Scored Matrix Obtained from the HPG.edf File

In addition, all biological signals in the data set are divided into epochs for 30 seconds. Considering that a day is 24 hours, there are 2880 epochs for a patient. Among the 197 patients included in this data set, 10 patients were randomly selected for this study and their information was shown in Table 2.1. EEG signals of the selected patients were studied throughout the study. [Url-6]

Table 2.1. Information of Selected Patients

Patients Label	Gender	Age
1.SC4001E0-PSG	Female	33
2.SC4002E0-PSG	Female	33
3.SC4011E0-PSG	Female	33
4.SC4012E0-PSG	Female	33
5.SC4021E0-PSG	Female	26
6.SC4031E0-PSG	Female	26
7.SC4032E0-PSG	Female	26
8.SC4251E0-PSG	Female	56
9.SC4152E0-PSG	Male	31
10.SC4332F0-PSG	Male	60

2.2. Preprocessing

All procedures and methods described in this section have been applied to the signals obtained with the parietal electrode.

Before extracting features from the signal to be examined, it must go through some preprocessing. These preprocesses help to obtain the necessary regions while preparing the signals for the next steps, thereby reducing the work of future processes. In this study, 3 processes were applied as preprocessing.

2.2.1. Normalization

Normalization, which is one of the signal preprocessing steps, ensures that the features have an equal dynamic range and at the same time provides convenience in calculations. For this reason, as in many studies, the normalization process was applied in this study in order to provide simplicity in the future operations. [11]

There are many methods for the normalization process. In this study, Z-score normalization (ZS), which is generally preferred for feature extraction and normalizes to a standard deviation of 1 and mean of 0, was used. [12]

$$f_i^{normalize} = \frac{f_i - mean}{\mathrm{STD}} \tag{2.1}$$

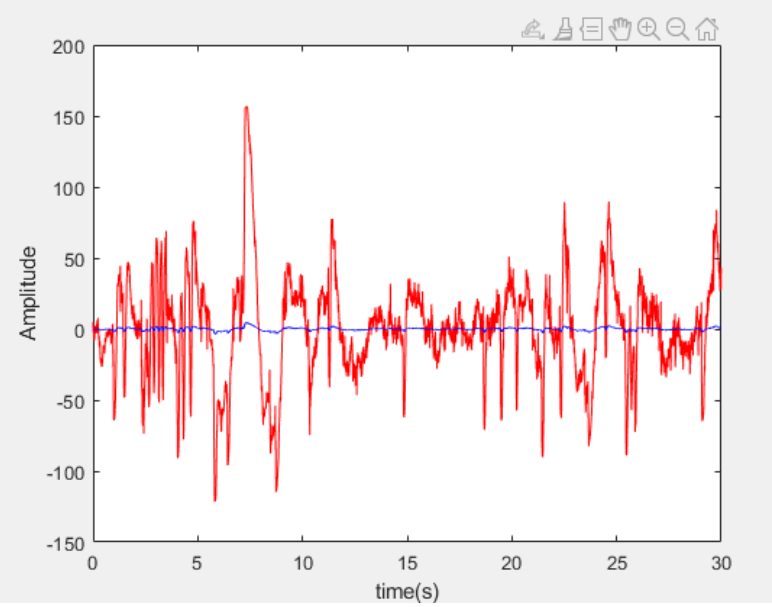

Figure 2.4. The Signal Before Normalization (red) and After Normalization (blue)

2.2.2. Outlier Removal

Outliers may occur in the data due to noise that may occur during data recording or errors. Such outliers can lead to false estimates and determinations. [13] [14] Therefore, in this study, these outliers were determined in order not to cause problem in further processing. By reducing the obtained determinations to the average value of the signal, errors that may occur were tried to be prevented.

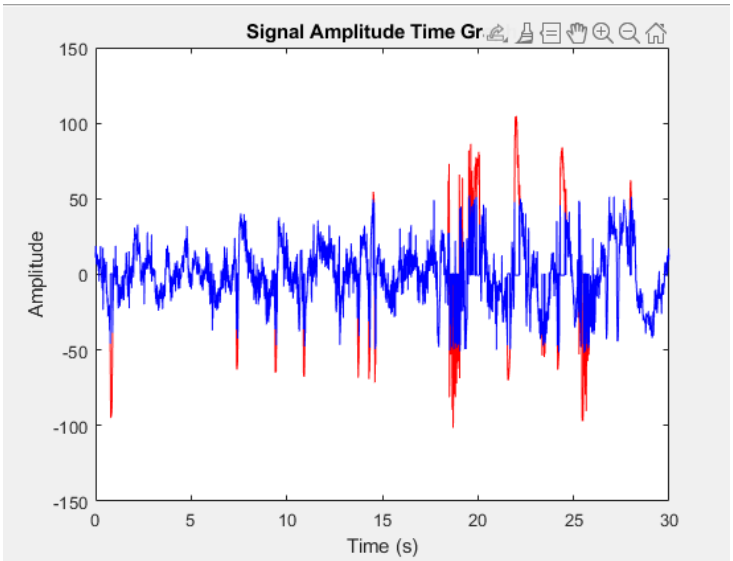

Figure 2.5. The Signal Before Outlier Removal (red) and After Outlier Removal (blue)

2.2.3. Bandpass

According to the spectral representation theory, any large class in the time series can be divided into different frequency divisions. According to this theory, this process can be done with a tool called band pass. Through this tool, a certain frequency range is not touched and other ranges are removed from the data. [15] [16]

Since the sleep stages include alpha, beta and tetha waves, the band pass filter to be used in this study is designed between 0.4-13 Hz.

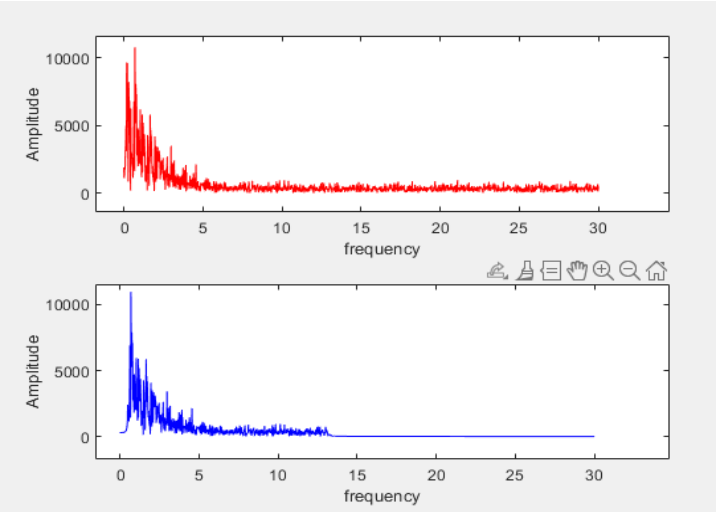

Figure 2.6. The Signal Without Bandpass Filter (red) and With Bandpass Filter (blue)

2.3. Feature Extraction

In this study, 20 different features of the time domain and frequency domain were extracted from the signal. First of all, 6 different time domain features were extracted from the signal. These features are zero crossing, Petrosian fractal dimension, distance between maximum minimum value, activity, complexity and mobility.

Then 7 frequency domain features were extracted to increase accuracy. Name of these features are median frequency, band power of delta, band power of theta, band power of alpha, approximate entropy of delta, approximate entropy of theta and approximate entropy of alpha.

7 more time domain features were extracted from the signal in order to increase accuracy, as the accuracy rate was lower than required. These features are as follows: kurtosis, skewness, average, standard deviation, root mean square, mean absolute deviation and interquatril interval.

2.3.1. Zero Crossing

In previous studies, the period analysis of the EEG was made by classifying the intervals between zero transitions into various frequency bands. [17]

In this study, taking into account that each part of the EEG signal will pass through the x-axis at different degrees, zero crossing values of the EEG signal were calculated for all epochs according to the formula [18] below.

$$\mathrm{ZC} = \sum_{n=1}^{n-1} \mathbf{sgn}(\boldsymbol{X_n} * \boldsymbol{X_{n+1}}) \cap |\boldsymbol{X_n} - \boldsymbol{X_{n+1}}| \geq \mathbf{0} \quad (2.2)$$

2.3.2. Petrosian Fractal Dimension

Fractal is the name given to the structures formed by the magnification and reduction of the same shape itself. Calculating the fractal size of any wave also provides information for transient detection. Many algorithms have been found for fractal analysis, which is also important for the analysis of biomedical signals. [19]

Petrosian fractal dimension calculation algorithm was used in this study. Petrosian Fractal Dimension is expressed as:

$$\boldsymbol{PFD} = \frac{\boldsymbol{log_{10}N}}{\boldsymbol{log_{10}N} + \boldsymbol{log_{10}}\left(\frac{\boldsymbol{N}}{\boldsymbol{N+0.4N\sigma}}\right)} \quad (2.3)$$

In this formula, N gives the length of the signal, while Nσ represents the number of changes in the signal derivative. [20] [21]

2.3.3. Distance Between Maximum Value and Minimum Value

While examining the signals, it was taken into consideration that each band would have different maximum and minimum values and the difference between these two values was calculated mathematically. In this mathematical calculation, the Pythagorean theorem is applied.

$$\boldsymbol{Distance\ between} \max \min \boldsymbol{value} =$$

$$\sqrt{(\boldsymbol{Xmaximum\ y\_Xminumum\ y})^2 + (\boldsymbol{Ymax} - \boldsymbol{Ymin})^2} \quad (2.4)$$

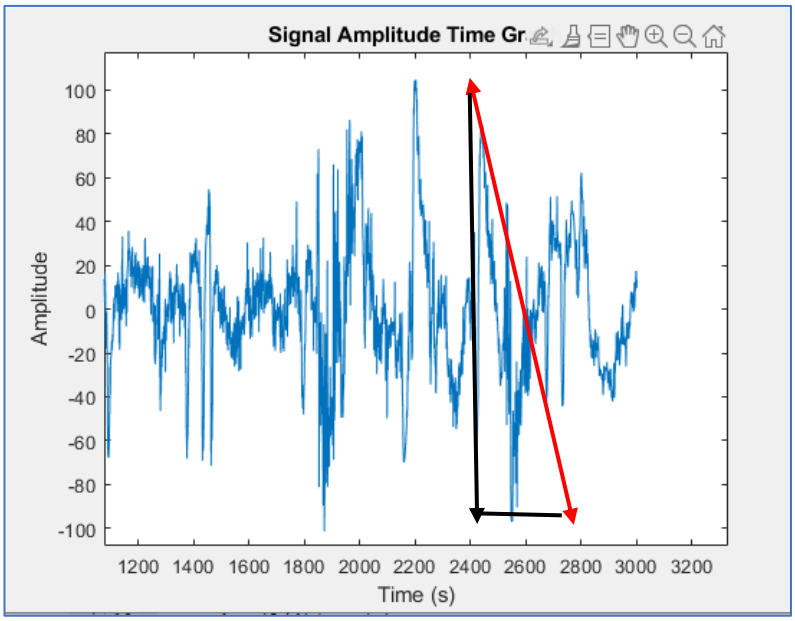

Figure 2.7. The Distance Between the Maximum and Minimum Value

2.3.4. Activity

Activity indicates the variance of a signal and is also an expression of the average power of the signal.

$$Activity = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{X}(\boldsymbol{t_i}) - Mean)^2 \quad (2.5)$$

2.3.5. Mobility

Mobility is an estimate of the average frequency and is formulated as:

$$Mobility = \sqrt{\frac{var\left(\frac{dx(t)}{dt}\right)}{var(x(t))}} \quad (2.6)$$

In short, the formula is the square root of the ratio of the activity of the first derivative of the signal to the activity of the signal.

2.3.6. Complexity

Complexity gives an estimate of the bandwidth of the signal. It is formulated as follows.

$$\boldsymbol{Complexity} = \sqrt{\frac{\boldsymbol{Mobility}\left(\frac{dx(t)}{dt}\right)}{\boldsymbol{Mobility}(x(t))}} \quad (2.7)$$

Briefly, the formula can be expressed as the ratio of the mobility of the first signal derivative to the signal mobility. [22]

2.3.7. Median Frequency

Median frequency represents the midpoint of the power distribution and is expressed as the frequency of a power spectrum where 50% of the total power is at lower frequencies and 50% of the power is at higher frequencies. [23] [24]

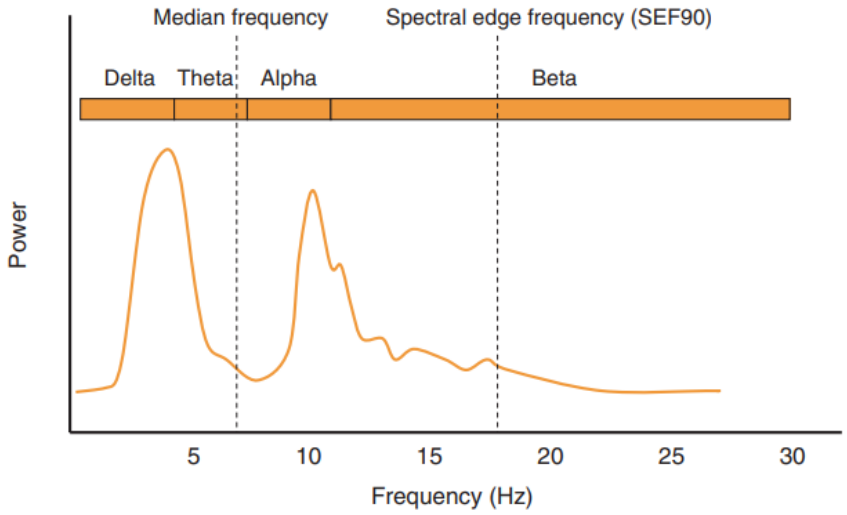

Figure 2.8. Representation of Median Frequency [24]

2.3.8. Bandpower

Bandpower is a parameter that gives the band power for each frequency division within a band. [25] [26] In this study, as stated before, the band powers of the alpha delta tetha divisions were calculated. Since it is known that the sleep signal contains alpha delta theta waves, the band power of 3 waves was calculated in this study. In order to calculate bandpower and approximate entropy for different brain waves, it is necessary to obtain the desired bands from the signal. In this study, Discrete Wavelet Transform (DWT) was used to obtain the desired bands.

a) Discrete Wavelet Transform

Discrete Wavelet Transform (DWT) is a structure that separates the received signal together with some operations and converts it into a smaller wave and is preferred to find the moment of sudden change. The main purpose of this structure is that the signal given to the system is decomposed into wavelet coefficients and these wavelets come together to give a linear structure of the signal.

Wavelet transform separates the signal into different frequency bands, and again, using the same wavelet transform, a decomposition is made to obtain smaller bands. These processes are obtained by passing the signal through low and high pass filters as shown below. [27] [28]

$$\text{High Pass Filter} = D_{1j} = \sum_{k=0}^{n} S(k)h(2j-k) \tag{2.8}$$

$$\text{Low Pass Filter} = A_{1j} = \sum_{k=0}^{n} S(k)g(2j-k) \tag{2.9}$$

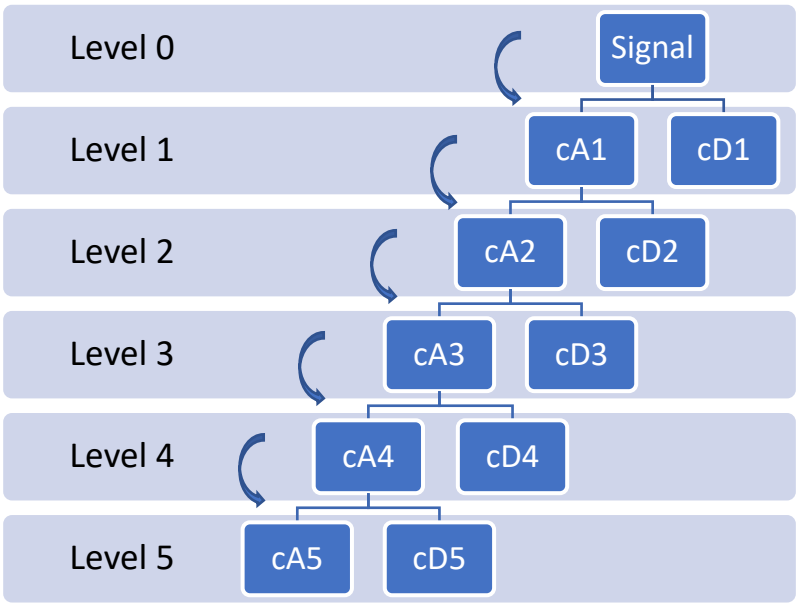

Figure 2.9. Representation of a Structure With 5 Segregation Levels [27]

In addition, as the level of wavelet separation increases, bandwidth narrowing increases, so the desired band gap separation levels can be increased or decreased. [29]

Wavelet transform is a highly preferred method in signal processing and is effective to extract different properties of signals when other processes are ineffective. [30]

2.3.9. Approximate Entropy

Approximate entropy is a statistical parameter that characterizes the amount of irregularity and predictability of subsequent amplitude values based on previous amplitude values. [31] [32]

Approximate entropy account schematized in Figure 2.10. Since it is known that the sleep signal contains alpha, delta and theta waves; the approximate entropy of 3 waves were calculated in this study.

The previously mentioned Discrete Wavelet Transform was used to obtain the desired alpha, delta and theta waves from the signal. The entropy was calculated for the obtained waves.

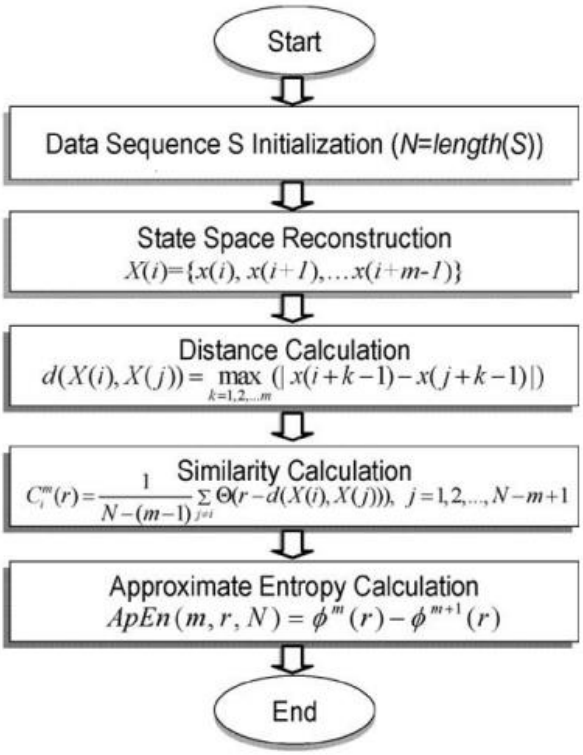

Figure 2.10. Scheme of Approximate Entropy Calculation [33]

2.3.10. Kurtosis

Kurtosis is a property that measures the relative uniformity of a distribution.[34] Kurtosis is formulated as follows.

$$\boldsymbol{K = \frac{1}{N}\frac{\sum_{i=1}^{N}(x(t_i)-mean)^4}{\mathrm{STD}^4}} \tag{2.10}$$

2.3.11. Skewness

Skewness is a feature that aims to verify and calculate symmetry in a data. [35]

$$\boldsymbol{S = \frac{1}{N}\frac{\sum_{i=1}^{N}(x(t_i)-mean)^3}{\mathrm{STD}^3}} \tag{2.11}$$

2.3.12. Mean

The mean property, which is preferred to extract signal features, is a parameter that gives the average value of the signal.

$$Mean = \frac{1}{N} \sum_{i=1}^{N} \boldsymbol{X(t_i)} \tag{2.12}$$

2.3.13. Standart Deviation

Standard deviation is a parameter used for the distribution of data values.

$$\boldsymbol{STD} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{X(t_i)} - \boldsymbol{mean})^2} \tag{2.13}$$

2.3.14. Root Mean Square

RMS is a statistical measure of the magnitude of a changing value, and it is a method of measuring the strength of biosignals. [36]

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{X(t_i)})^2} \tag{2.14}$$

2.3.15. Mean Absolute Deviation

MAD is a parameter that expresses the distance of each point of the signal to the signal mean. [37]

$$\boldsymbol{MAD} = \frac{1}{N} \sum_{i=1}^{N} (\boldsymbol{X(t_i)} - \boldsymbol{mean}) \tag{2.15}$$

2.3.16. Interquartile Range

It is a parameter that shows the difference between 25% and 75% of the signal. [38]

$$\boldsymbol{IQR} = \frac{\boldsymbol{3y(t)}}{\boldsymbol{4}} - \frac{\boldsymbol{y(t)}}{\boldsymbol{4}} \tag{2.16}$$

2.4. Classification

The main purpose of this study is to classify sleep stages using sleep signals. K-Nearest Neighbor Algorithm (KNN), Support Vector Machine (SVM), Decision Tree, Artificial Neural Networks were used for classification.

2.4.1. Classification Algoritms

2.4.1.1. Support Vector Machine (SVM)

Support vector machines, one of the machine learning algorithm types, is a remarkable algorithm in diagnosing disease and classifying biological signals. When given a labeled data, the SVM algorithm learns this and creates a model to predict which group the given test data belongs to. The model that makes up the working mechanism of the algorithm can be explained as follows.

Among the classes, 2 points are selected as shown in the Figure 2.11. These two points are called Support vectors. Two lines are drawn through these support vectors and these lines are called boundary lines. The zone [green zone] between these 2 lines is called Margin. The wider the margin, the better the classification is done and represents 2/w numerically. [Url-7] Another line is drawn to divide these 2 lines at an equal distance and is called the decision line (hyperplane). The SVM classes divided into two are denoted as -1, +1, and the line below the decision line is expressed as wx + b = -1 and the area above it as wx + b = 1. [39]

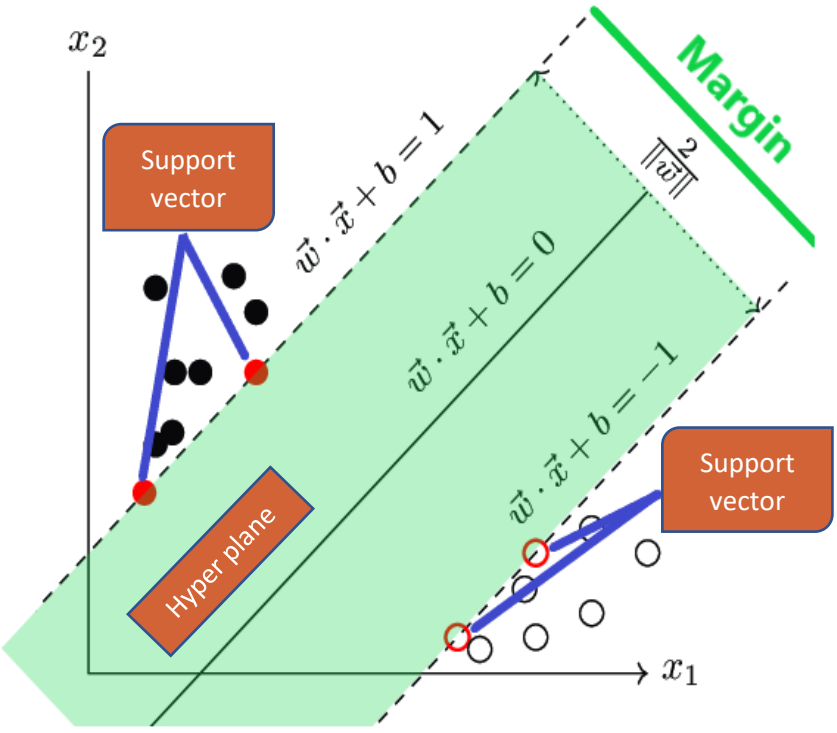

Figure 2.11. Support Vector Machine Model [Url-8]

There are different approaches to classifying groups with many different data, but the most common approach is that the classifier looks for which group the data belongs to in binary groups. As a result, SVM performs classification action on a strong theoretical basis in a powerful model and is used for data classification in many areas, especially in the field of health. [40]

2.4.1.2. K- Nearest Neighbor (KNN)

One of the algorithms used to classify with machine learning is the k-Nearest neighbor algorithm. Known as the simplest algorithm, KNN is also known as lazy learning. The reason why it is known as lazy learning is that there is no real learning in the KNN algorithm, a classification process is made by memorizing the data. The working principle of the KNN algorithm is based on the classification of the data in the feature area according to their distance from the entered training data. The distance calculation in question is made with many formula. İn this study the Euclidean formula given in Figure 2.12 was used to calculate distance.

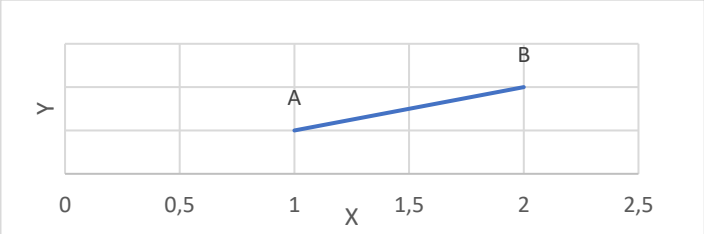

Figure 2.12. Euclidean Distance Graph

$$\boldsymbol{Distance} = \sqrt{(\boldsymbol{X2} - \boldsymbol{X1})^2 + (\boldsymbol{Y1} - \boldsymbol{Y2})^2} \quad (2.17)$$

The K value represents the number of neighbors to be scanned. For example, as shown in Figure 2.13, the number of neighbors to be scanned, ie. k selections, is shown to classify the unknown shapes. In summary, "k" neighbors are scanned in the data set in the algorithm. This K value is determined before running the algorithm. As can be seen in Figure 2.13, the selection of the K value is critical. If the K value is too small or too much, the model may lose its purpose or cause undesirable situations such as noise in the model [39]. In this study, the k number was taken as 5 because the KNN algorithm includes the classification of 5 sleep stages.

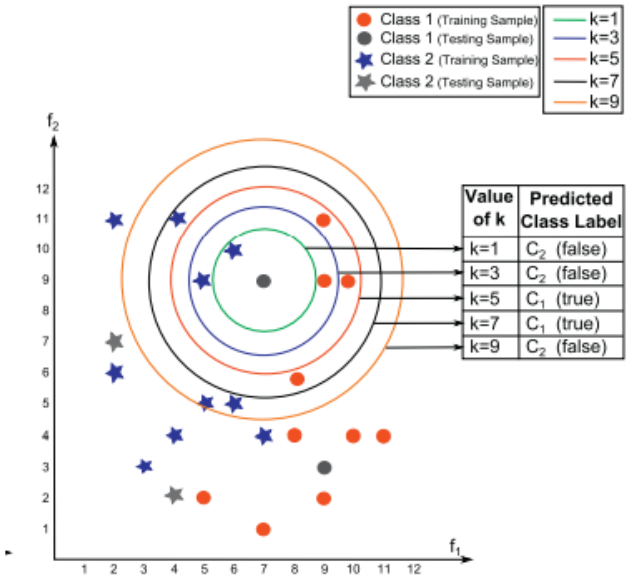

Figure 2.13. Different Number of Neighbor Values [41]

KNN advantages:

- ✓ Since the process is simple, any error can be detected immediately.
- ✓ There are special noise reduction techniques belonging to the KNN algorithm.

KNN disadvantages:

- ✓ If the training set is large, KNN may work with difficulty.
- ✓ It is sensitive to unrelated features, so careful feature selection should be made.
- ✓ It may be necessary to use a few additional algorithms to this algorithm in terms of classifier. [42] [43]

2.4.1.3. Decision Tree (DT)

One of the other classification algorithms is decision trees. Decision trees are a highly preferred classification algorithm due to their easy application and easy interpretation. Purpose of it is the same as for other classifiers. It is to assign the test data set according to the training data set. [44]

The decision tree diagram is schematized in Figure 2.14:

-Decision Node

-Change Node

-Terminal Node [44]

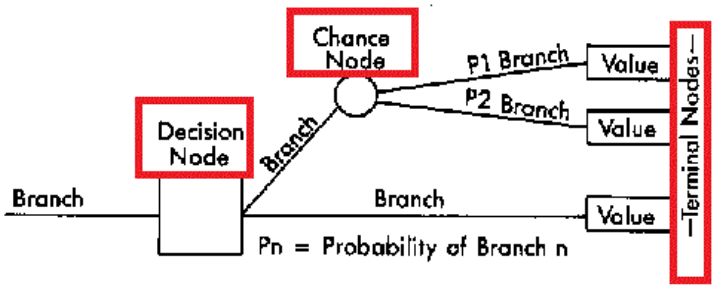

Figure 2.14. Decision Tree Algorithm Model [Url-9]

When a decision-making entry is given to this model, an estimate is made by determining the best quality for classification from the decision node, and a chance node is created for this prediction. This classification continues until a decision is made and the clusters created in the classification process are included in a single class, then decision node is created. In short, it has a model that extends from the root part to the resulting part in the form of branching. [46]

Decision tree classifications are divided into 2 group:

- Entropy-based classification trees (ID3, C4.5)
- Regression trees (CART) [39]

a) ID3 (Iterative Dichotomiser 3)

In the ID3 algorithm put forward by Quinlan Ros, a classification and branching process is applied according to the gain of top-down entropy.

As a first step, the entropy on the data set is calculated according to Equation (2.18). The entropy of property vectors is calculated as the second step. The difference between entropy on the data set and the entropies of property vectors is calculated The feature that is the biggest difference is the highest earning property, and branching is maintained through this property. In this way, classification of classes with different structures will be provided in a good way.

$$Entropy(S) = \sum_{i=1}^{c} -p(I)log_2 p(I) \qquad (2.18)$$

The ratio of S of class S is represented by p(I). [46] [47]

b) C4.5 Classification Algorithm

This algorithm technique is also a technique put forward by Quinlan Ros. In the C4.5 algorithm, considered an extension of ID3, threshold method is used unlike the ID3 algorithm. The values in the feature vector are considered as binary and their average is assigned as threshold. With whichever threshold value the gain is higher, the process continues with that threshold value. The features are divided into groups according to this threshold value and ID3 is made after this separation. [46]

c) Classification and Regression Trees (CART)

The classification and regression trees were introduced by Breiman in 1984. When the classification is created, the regreston trees can also be created. The model performed by this algorithm is based on the binary separation of features. The Gini index given in Equation (2.19) is used for dual separation. This index is used to determine the frequency of misperception. The reason why this algorithm is different from other algorithms is that it can also be used to perform regression analysis. [46]

Regression analysis is a system used to understand the relationship between the given variables. In short, it is a statistical method used to predict the effects of one group of variables on another group of variables. [48]

$$Gini = 1 - \sum_{i=1}^{c} (p_i)^2 \tag{2.19}$$

p_i represents the ratio of states reaching that node. [49]

2.4.1.4. Artificial Neural Network (ANN)

a) Relationship Between Brain and Neuron

Human's headquarters is the brain. It is the part that determines how the body will react by examining the information coming from the body. It shapes this examination according to the needs of the body. The central nervous system includes the brain and spine and consists of two structures. Neurons and Glia. While Glia cells, which are much more than neurons, play a role in the healthy and trouble-free functioning of the system, neurons are seen as a messenger of information from the body. It is responsible for providing communication between the parts of the body and the brain, and it conducts this communication by examining and using electrical and chemical stimuli. In short, many actions taken during the day are provided by neurons and the glial cells that support them. [Url-10] [Url-11]

b) Neuron

The number of neurons in the human body has been the subject of many studies. With the latest studies, this number has been said to be 86 billion. [50]

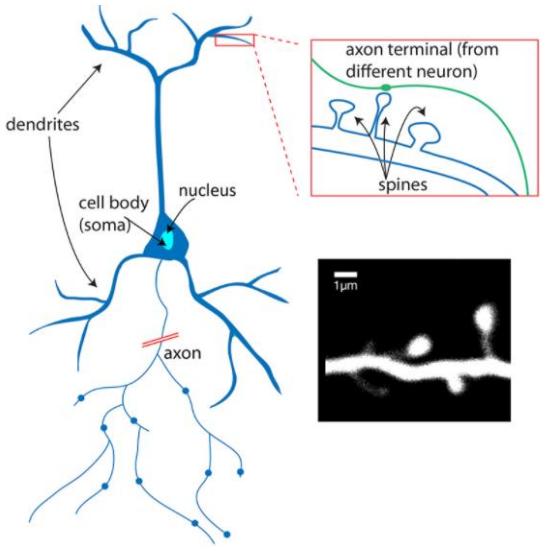

Figure 2.15. Parts of the Neuron [Url-14]

Neurons can be examined under 3 parts as seen in Figure 2.15.

a) Axon: It has a long and thin structure and acts as the message transmission part.
b) Dendrites: Shaped like a tree branch or root. This part is responsible for receiving the intracellular message.
c) Cell body: It is the part that contains a nucleus in its structure and is responsible for controlling the activities of the cell.

Neurons communicate with other neurons. This communication is achieved by the release of chemical structures called neurotransmitters from synapses between the axon of one neuron and the dendrites of the other neuron, as shown in Figure 2.16. [Url-12] [Url-13] [Url-14]

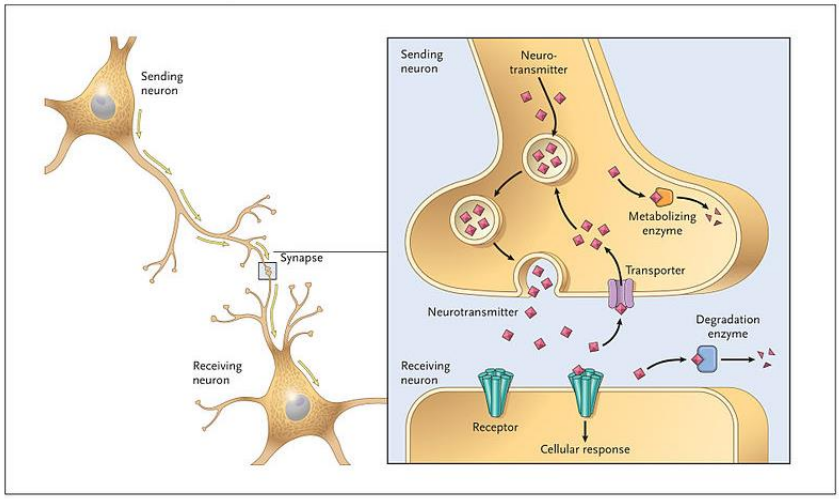

Figure 2.16. Neurotransmitter Transfer in the Synaptic Cleft [Url-15]

c) Artificial Neural Networks and Structure

The model created by using the neuron structure of the brain described earlier is called artificial neural networks. The model, which has been used in many areas recently, is also a preferred model for classification. Artificial neural networks can be schematized as inputs, weight addition function section, bias, activation function section and output as shown in Figure 2.17. [51]

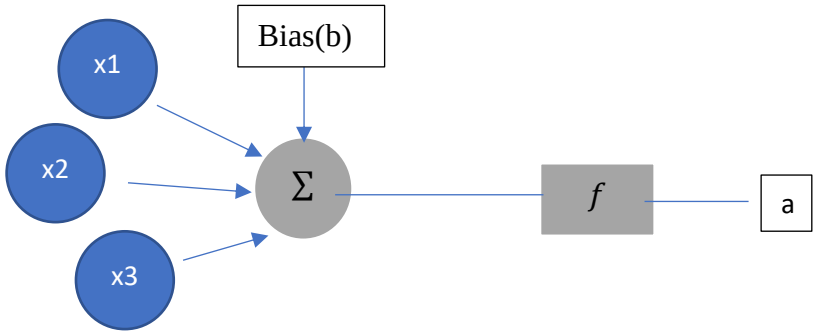

Figure 2.17. Basic Scheme of Artificial Neural Network

$$a = f\left(\sum_{i=1}^{N} w_i x_i + b\right) \tag{2.20}$$

Entries: This is the section where incoming entries are shown. Entries are expressed as x_i in Equation (2.20).

Adder: The part that adds up entries multiplied by their own weights. Weights are expressed in w_i.
Bias: It is a structure used to increase or decrease the net entry entering the activation function section.
Activation function section: The activation process is done in order to limit the amplitude of the output from a neuron, this activation process is done in this section.
Output section: It is the section that gives the output that occurs after the operations. [51][52]

d) Activation Function Types

- Threshold Function

$$f(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \tag{2.21}$$

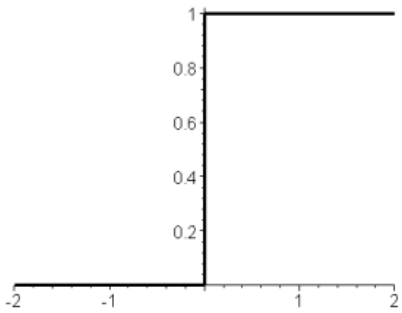

Figure 2.18. Threshold Function Formula and Graph

- Piecewise Linear Function

$$f(x) = \begin{cases} 1, & x \geq \frac{1}{2} \\ x + \frac{1}{2}, & \frac{1}{2} \geq x \geq -\frac{1}{2} \\ 0, & -\frac{1}{2} \geq x \end{cases} \tag{2.22}$$

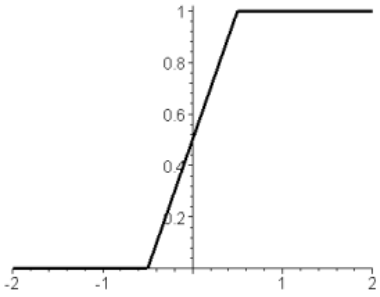

Figure 2.19. Piecewise Linear Function Formula and Graph

- Sigmoid Function (Logistic Function)

$$f(x) = \frac{1}{1+e^{x}} \tag{2.23}$$

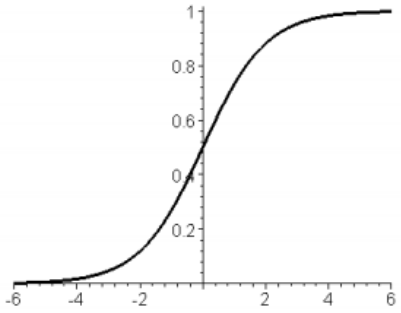

Figure 2.20. Sigmoid Function Formula and Graph [53] [54]

e) Multilayer Neural Network

In multilayer neural networks, neuron structures are arranged in layers. As seen in Figure 2.21, there are hidden layers between the structure consisting of input and output layers. Nodes in the entrance part of the network create the inputs that will be applied to the neurons in the second layer. These entries contain the elements that will enter the function section from the activation mentioned earlier. Like the conduction logic in the neuron, the outputs from the second layer are given as input to the third layer, and this branching continues in this way. Each layer of the formed networks has the output of the previous layer. In short, the inputs given from the input part in the artificial neural network pass through the activation model and the resulting responses are transferred to the other layer, thus, outputs are created from the neurons in the final output layer after branching. [39] [53] [54]

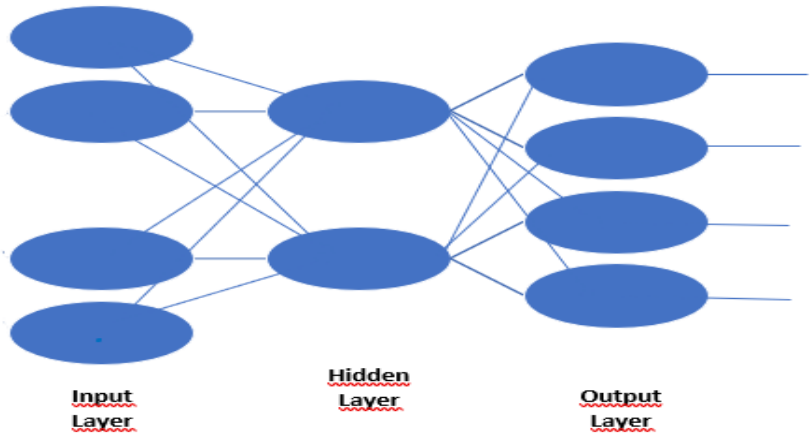

Figure 2.21. Multilayer Neural Network

f) Feed Forward Neural Network

In feed forward neural networks, the outputs are forward-oriented, return between 2 layers is not possible. Output occurs as a result of operations from inputs and these outputs are given directly to the other layer as input. [39]

g) Feedback Artificial Neural Network

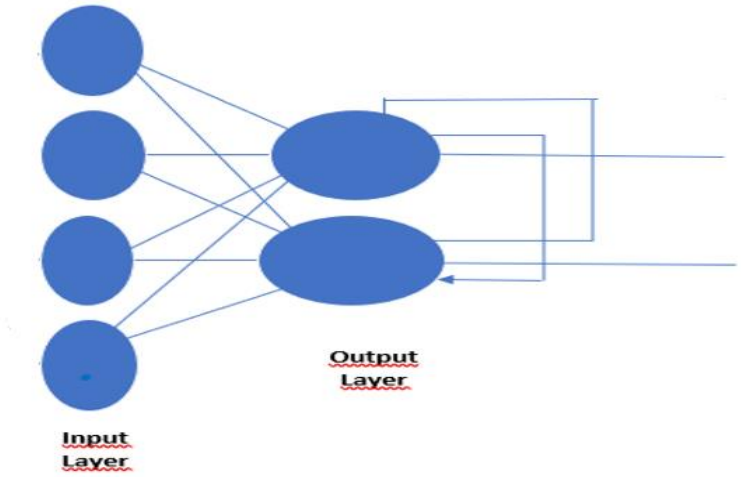

Figure 2.22. Communication Between the Layers

In feedback neural networks, as shown in Figure 2.22, while some outputs affect some inputs, there is a communication between layers. [55]

h) Advantages and disadvantages of Artificial Neural Networks

Advantages:

- ✓ Since it is parallel processed, it creates an ideal structure for linear program uses.
- ✓ Thanks to parallel processing, trouble in any element will not affect further steps.
- ✓ It has a preferable structure in order to simplify any complex situation.

Disadvantages:

- ✓ Since it has parallel processing, much time and processing will be required. [Url-16]

i) Artificial Neural Network Structure of This Study

In this study, Classification Learner App in MATLAB were used for ANN.

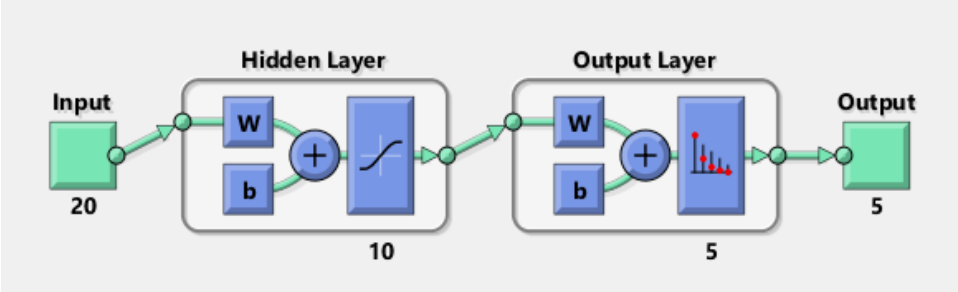

Figure 2.23. Neural Network Architecture For This Study

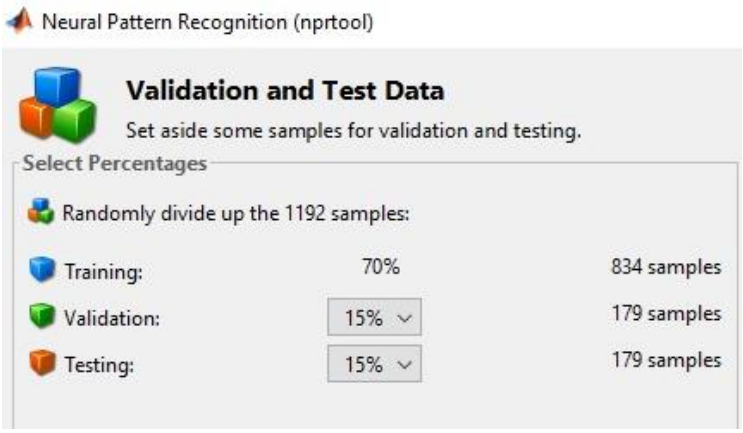

Figure 2.24. Percentages of Training Validation and Test Data Used in ANN

2.4.2. Test and Training Set Separation

The working scheme followed in this section is shown in Figure 2.25.

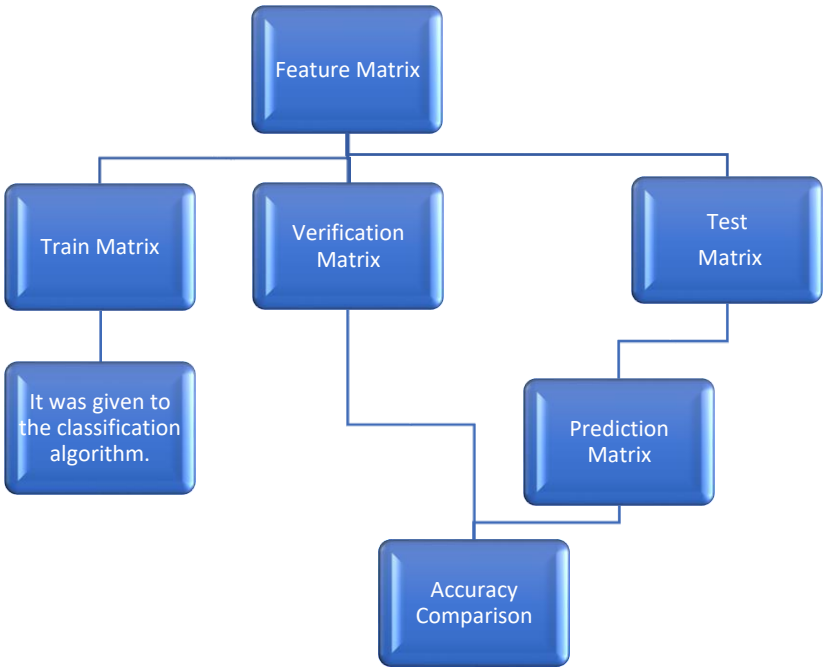

Figure 2.25. The Work Scheme Followed in Classification Section

After the EEG signals of 10 patients were passed through the aforementioned preprocessing and feature extraction sections, the resulting feature matrices were recorded. The 10 different feature matrices recorded were divided into test matrix and training matrix before being included in the classification process. 'cvpartition' command was used for this grouping operation. 'cvpartition' command is provided in Matlab 2018b environment. This command defines sections within the given dataset as random training and testing. [56] [57]

In this study, 70% of the values to be given was asked to be defined as training and 30% as test. Therefore, the 'cvpartition' was designed according to these values and the distribution was provided. Test train and total epochs for patients are shown in Table 2.2.

Table 2.2. Number of Epochs Taken From Patients

	Train	Test	Total
1	505	216	721
2	706	302	1008
3	688	295	983
4	746	320	1066
5	634	271	905
6	582	250	832
7	554	237	791
8	596	256	852
9	1149	493	1642
10	834	358	1192
TOTAL	6994	2998	9992

2.4.3. Training The Classifier

Using the training dataset obtained from the above mentioned 'cvpartition' command, the classifiers were trained before classification. Since all classification algorithms are for 5 classes, classifiers was modeled as one class against all classes.

2.4.4. Estimating The Classifier Using The Training Data Set

The classifier trained with the training data set was enabled to predict sleep stages using the test data set and estimates have been recorded.

2.4.5. Checking Classifier Estimates

The estimates taken separately for 10 patients were compared with the validation scores to measure the percentage correct estimation of the classifier. The comparison result was recorded and the confusion matrices were created. The accuracy rate, sensitivity and specificity were calculated from the confusion matrices.

2.4.6. Making Calculations from the Confusion Matrix

Confusion matrices are a good structure to observe between the prediction of the classifier and the actual values. Confusion matrix and calculations for 2 classes are given below.

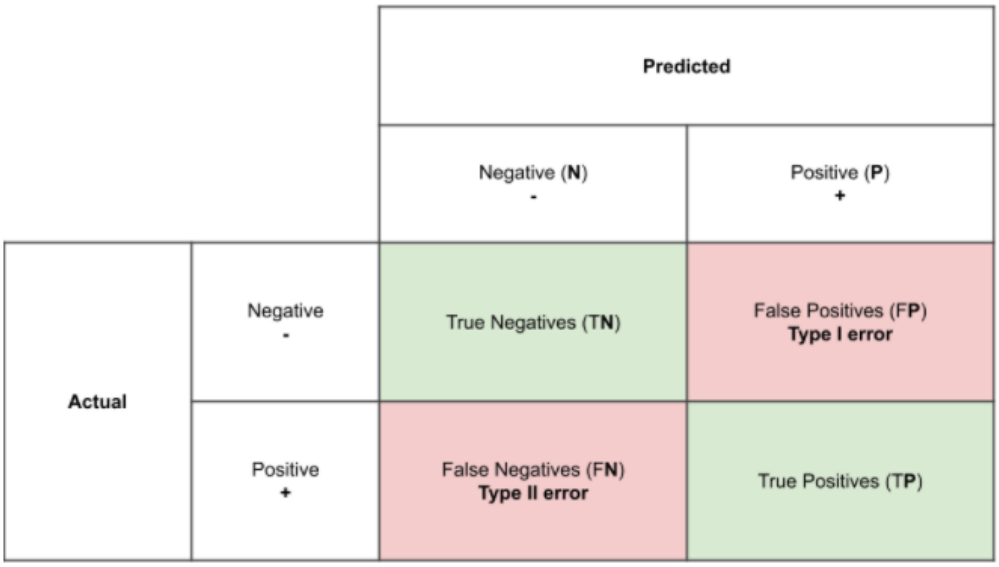

Figure 2.26. Confusion Matrix with Predicted and Actual Values [Url-17]

TN = The part where the classifier predicts the negative state as negative

TP = The part where the classifier predicts the positive state as positive

FP = the part where the classifier predicts the negative state as positive

FN = the part where the classifier predicts the positive state as negative

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.24)$$

$$Sensivity = \frac{TP}{TP+FN} \quad (2.25)$$

$$Specificity = \frac{TN}{TN+FP} \quad (2.26)$$

Since 5 classes were examined in this study, calculations were made using a 5x5 confusion matrix. [Url-17]

3. RESULT

In this study, which aims to classify the sleep stages using the database named Sleep-EDF Database (Expanded). After the signals obtained from 10 selected patients were preprocessed, the features were obtained. By using these features, classification has been made by Support Vector Machines, k-Nearest Neighbor, Decision Tree and Artificial Neural Networks algorithms. Confusion matrices were obtained from results of the classification. Sensitivity, specificity and accuracy ratios were also calculated from the confusion matrices.

3.1. KNN Results

Table 3.1. KNN Result Table

Patient		NonREM1	NonREM2	NonREM3	REM	Wake	Overall Accuracy
1	Acc.	93.06	86.57	91.20	90.28	94.44	77.78
	Sens.	33.33	76.40	88.71	70.27	72.00	
	Spec.	7.14	89.47	82.09	72.22	78.26	
2	Acc.	99.64	92.53	80.43	92.17	85.05	74.83
	Sens.	0.00	0.00	74.17	82.98	68.06	
	Spec.	0.00	0.00	78.76	92.86	72.06	
3	Acc.	92.52	76.87	94.90	83.67	96.94	72.45
	Sens.	73.68	73.33	92.00	53.57	58.33	
	Spec.	45.16	92.77	63.89	30.00	63.64	
4	Acc.	94.67	82.45	96.87	89.34	97.18	80.25
	Sens.	75.00	80.26	88.00	77.78	75.00	
	Spec.	28.57	94.92	75.86	59.32	46.15	
5	Acc.	94.46	86.35	93.73	90.41	99.63	82.29
	Sens.	87.50	86.22	75.00	68.00	100.00	
	Spec.	33.33	94.41	48.00	77.27	50.00	
6	Acc.	91.16	83.53	96.79	83.94	97.99	76.71
	Sens.	0.00	84.76	70.00	63.24	66.67	
	Spec.	0.00	89.68	58.33	74.14	33.33	
7	Acc.	93.67	82.70	94.94	85.23	98.73	77.64
	Sens.	37.50	84.17	93.33	66.23	50.00	
	Spec.	23.08	82.11	73.68	85.00	33.33	
8	Acc.	95.69	81.57	98.04	86.67	98.82	80.39
	Sens.	90.00	80.28	0.00	72.73	0.00	
	Spec.	66.67	97.16	0.00	36.36	0.00	
9	Acc.	96.54	98.85	97.36	92.07	99.34	86.59
	Sens.	0.00	87.50	84.85	70.87	94.58	
	Spec.	0.00	77.78	94.92	89.02	96.48	
10	Acc.	94.12	87.11	---	91.32	95.52	84.03
	Sens.	50.00	85.85	---	55.17	90.74	
	Spec.	19.05	91.92	---	47.06	94.23	
Average	Acc.	94.55	85.05	93.80	88.51	96.07	79.30
	Sens.	44.70	73.88	74.01	68.08	67.54	
	Spec.	22.30	81.02	63.95	66.33	56.75	

The highest overall accuracy for KNN among all patients is 86.59%. The highest accuracy for KNN among all stages belongs to wake stage with 96.07%. The lowest accuracy for KNN among all stages belongs to nonREM2 stage with 85.05%.

3.2. SVM Results

Table 3.2. SVM Result Table

Patient		NonREM1	NonREM2	NonREM3	REM	Wake	Overall Accuracy
1	Acc.	93.52	89.35	91.67	95.83	95.37	82.87
	Sens.	50.00	81.93	91.53	84.62	76.00	
	Spec.	35.71	89.47	80.60	91.67	82.61	
2	Acc.	99.65	91.87	87.63	94.35	89.40	81.13
	Sens.	0.00	22.22	83.62	88.64	74.36	
	Spec.	0.00	11.11	85.84	92.86	85.29	
3	Acc.	93.88	85.71	96.94	89.46	98.64	82.31
	Sens.	88.24	84.07	93.55	67.92	81.82	
	Spec.	48.39	92.17	80.56	72.00	81.82	
4	Acc.	92.16	90.28	97.18	91.85	97.81	84.64
	Sens.	40.91	90.69	85.71	80.00	80.00	
	Spec.	42.86	93.91	82.76	74.58	61.54	
5	Acc.	94.10	90.04	95.20	92.99	98.89	85.61
	Sens.	66.67	91.30	73.08	77.78	0.00	
	Spec.	47.62	93.85	76.00	79.55	0.00	
6	Acc.	90.76	89.96	97.99	88.35	98.39	82.73
	Sens.	33.33	90.12	81.82	75.44	75.00	
	Spec.	27.78	94.19	75.00	74.14	50.00	
7	Acc.	94.09	89.45	96.20	89.03	99.16	83.97
	Sens.	46.15	91.53	93.94	74.29	66.67	
	Spec.	46.15	87.80	81.58	86.67	66.67	
8	Acc.	96.86	88.24	97.25	93.33	99.22	87.45
	Sens.	88.00	89.67	25.00	82.93	100.00	
	Spec.	81.48	93.75	20.00	77.27	33.33	
9	Acc.	96.34	93.50	98.58	92.68	96.95	89.02
	Sens.	0.00	87.05	94.83	73.96	96.46	
	Spec.	0.00	89.63	93.22	86.59	95.98	
10	Acc.	92.16	90.76	---	94.68	95.52	86.55
	Sens.	18.18	89.86	---	69.23	94.00	
	Spec.	90.52	93.94	---	79.41	90.38	
Average	Acc.	94.35	89.92	95.41	92.26	96.93	84.63
	Sens.	43.12	81.84	80.34	77.48	74.43	
	Spec.	42.05	83.98	75.06	81.47	64.76	

The highest overall accuracy for SVM among all patients is 89.02%. The highest accuracy for SVM among all stages belongs to wake stage with 96.93%. The lowest accuracy for SVM among all stages belongs to nonREM2 stage with 89.92%.

3.3. DT Results

Table 3.3. DT Result Table

Patient		NonREM1	NonREM2	NonREM3	REM	Wake	Overall Accuracy
1	Acc.	88.43	87.04	90.74	87.50	93.52	73.61
	Sens.	21.05	82.43	84.06	63.64	71.43	
	Spec.	28.57	80.26	86.57	58.33	65.22	
2	Acc.	99.64	92.53	86.12	91.46	89.32	79.14
	Sens.	0.00	28.57	87.76	81.25	71.59	
	Spec.	0.00	11.11	76.11	92.86	92.65	
3	Acc.	90.14	82.99	95.24	90.48	95.58	77.21
	Sens.	54.17	82.22	86.67	72.00	40.00	
	Spec.	41.94	89.16	72.22	72.00	36.36	
4	Acc.	89.97	84.95	95.92	88.09	97.18	78.06
	Sens.	32.26	84.98	80.77	73.33	100.00	
	Spec.	47.62	91.88	72.41	55.93	30.77	
5	Acc.	92.62	85.98	95.20	88.56	98.52	80.44
	Sens.	52.94	91.72	70.00	62.75	25.00	
	Spec.	42.86	86.59	84.00	72.73	50.00	
6	Acc.	89.56	85.14	95.98	87.95	97.19	77.91
	Sens.	0.00	85.54	75.00	72.58	44.44	
	Spec.	0.00	91.61	25.00	77.59	66.67	
7	Acc.	92.83	85.23	93.25	88.19	99.16	79.32
	Sens.	33.33	86.67	80.56	73.53	100.00	
	Spec.	30.77	84.55	76.32	83.33	33.33	
8	Acc.	95.29	85.10	98.04	88.63	99.22	83.14
	Sens.	82.61	85.20	0.00	71.43	100.00	
	Spec.	70.37	94.89	0.00	56.82	33.33	
9	Acc.	96.54	86.99	97.56	90.45	94.31	82.93
	Sens.	0.00	72.90	91.23	71.60	92.96	
	Spec.	0.00	83.70	88.14	70.73	92.96	
10	Acc.	89.08	86.27	---	94.68	93.00	81.51
	Sens.	17.86	89.01	---	70.27	89.11	
	Spec.	23.81	85.86	---	76.47	86.54	
Average	Acc.	92.41	86.22	94.23	89.60	95.70	79.33
	Sens.	29.42	78.92	72.90	71.24	73.45	
	Spec.	28.59	79.96	64.53	71.68	58.78	

The highest overall accuracy for DT among all patients is 83.14%. The highest accuracy for DT among all stages belongs to wake stage with 95.70%. The lowest accuracy for DT among all stages belongs to nonREM2 stage with 86.22%.

3.4. ANN Results

Table 3.4. ANN Result Table

Patient		NonREM1	NonREM2	NonREM3	REM	Wake	Overall Accuracy
1	Acc.	92.60	89.82	93.52	92.59	96.30	82.41
	Sens.	11.11	92.68	84.21	92.86	83.33	
	Spec.	100.00	82.61	96.97	65.00	62.50	
2	Acc.	96.03	86.76	94.70	92.05	96.03	82.78
	Sens.	0.00	80.36	93.33	90.91	72.73	
	Spec.	0.00	83.33	89.36	76.92	72.73	
3	Acc.	93.20	87.76	96.60	93.88	98.64	85.03
	Sens.	50.00	92.41	90.91	78.26	85.71	
	Spec.	80.00	85.88	86.96	81.82	85.71	
4	Acc.	96.88	88.75	96.88	90.00	98.75	85.63
	Sens.	50.00	90.99	80.00	73.91	87.50	
	Spec.	80.00	92.66	72.73	62.96	87.50	
5	Acc.	95.59	91.91	96.32	94.12	100.00	88.97
	Sens.	50.00	97.70	63.64	88.47	0.00	
	Spec.	100.00	90.43	87.50	82.14	0.00	
6	Acc.	93.60	86.40	96.80	88.00	98.40	81.60
	Sens.	0.00	87.67	62.50	91.18	50.00	
	Spec.	0.00	88.89	83.33	72.09	100.00	
7	Acc.	94.96	93.28	97.48	91.60	99.16	88.24
	Sens.	42.86	92.06	90.91	91.67	50.00	
	Spec.	60.00	95.08	83.33	82.50	100.00	
8	Acc.	98.44	92.97	96.88	94.53	100.00	91.41
	Sens.	80.00	96.88	0.00	86.96	0.00	
	Spec.	80.00	93.94	0.00	83.33	0.00	
9	Acc.	97.97	92.28	97.56	92.68	96.75	88.62
	Sens.	0.00	85.29	86.21	86.49	96.26	
	Spec.	0.00	86.57	92.59	71.11	96.26	
10	Acc.	94.41	90.50	-----	94.97	95.53	87.71
	Sens.	16.67	94.74	-----	85.00	92.31	
	Spec.	100.00	88.24	-----	73.91	92.31	
Average	Acc.	95.37	90.04	96.67	92.44	97.96	86.24
	Sens.	30.06	91.08	72.41	86.57	61.78	
	Spec.	60.00	88.76	76.97	75.18	69.70	

The highest overall accuracy for ANN among all patients is 91.41%. The highest accuracy for ANN among all stages belongs to wake stage with 97.96%. The lowest accuracy for ANN among all stages belongs to nonREM2 stage with 90.04%.

3.5. Final Results

In order to compare and evaluate the obtained results, a general result table has been created.

Table 3.5. Final Result Table

Overall	Accuracy Rate
KNN	79.30%
SVM	84.63%
DT	79.33%
ANN	86.24%

As can be seen from the final result table, the best accuracy rate was obtained from the Artificial Neural Networks classification algorithm with the accuracy rate of 86.24%. Later, the best accuracy rate was obtained from SVM algorithm as 84.63%, while the lowest accuracy rate belongs to the KNN algorithm with 79.30%.

Table 3.6. The Final Result of the NonREM1

NonREM1	Accuracy Rate	Sensitivity	Specificity
KNN	94.55	44.70	22.30
SVM	94.35	43.12	42.05
DT	92.41	29.42	28.59
ANN	95.37	30.06	60.00

The highest accuracy for NonREM 1 belongs to ANN with 95.37%, then KNN algorithm with 94.55%. DT has the lowest accuracy with 92.41%.

Table 3.7. The Final Result of the NonREM2

NonREM2	Accuracy Rate	Sensitivity	Specificity
KNN	85.05	73.88	81.02
SVM	89.92	81.84	83.98
DT	86.22	78.92	79.96
ANN	90.04	91.08	88.76

The highest accuracy for NonREM 2 belongs to ANN with 90.04%, then SVM algorithm with 89.92%. KNN has the lowest accuracy with 85.05%.

The highest sensitivity value for NonREM 2 belongs to ANN with 91.08%. At the same time, ANN has the highest specificity value for NonREM 2 with the rate of 88.76%.

Table 3.8. The Final Result of the NonREM3

NonREM3	Accuracy Rate	Sensitivity	Specificity
KNN	93.80	74.01	63.95
SVM	95.41	80.34	75.06
DT	94.23	72.90	64.53
ANN	96.67	72.41	76.97

The highest accuracy for NonREM 3 belongs to ANN with 96.67%, then SVM algorithm with 95.41%. KNN has the lowest accuracy with 93.80%.

The highest sensitivity value for NonREM 3 belongs to SVM with 80.34%.

Table 3.9. The Final Result of the REM

REM	Accuracy Rate	Sensitivity	Specificity
KNN	88.51	68.08	66.33
SVM	92.26	77.48	81.47
DT	89.60	71.24	71.68
ANN	92.44	86.57	75.18

The highest accuracy for REM belongs to ANN with 92.44%, then SVM algorithm with 92.26%. KNN has the lowest accuracy with 88.51%.

The highest sensitivity value for REM belongs to ANN with 86.57% and the highest specificity value for REM belongs to SVM with 81.47%.

Table 3.10. The Final Result of the Wake

WAKE	Accuracy Rate	Sensitivity	Specificity
KNN	96.07	67.54	56.75
SVM	96.93	74.43	64.76
DT	95.70	73.45	58.78
ANN	97.96	61.78	69.70

The highest accuracy for wake stage belongs to ANN with 97.96%, then SVM algorithm with 96.93%. DT has the lowest accuracy with 95.70%.

4. DISCUSSION

In this study, sleep stage classification was made using single channel EEG. 20 time and frequency features were used. SVM, KNN, DT and ANN were used as machine learning algorithms.

The highest overall accuracy rate is 86.24% and was obtained from ANN algorithm. At the same time, the highest accuracy rates for all stages have been achieved with ANN algorithm and the accuracy rates are 95.37%, 90.04%, 96.67%, 92.44% and 97.96% for non-REM1, non-REM2, non-REM3, REM and wakefulness, respectively.

The sensitivity and specifity values of the non-REM3 stage when the patient cannot fall into deep sleep or the wake stage when the patient does not wake between sleep are obtained as 0.00. This situation decreases the average sensitivity and specifity values. At the same time, this situation can be caused by the unbalanced distribution of epoch numbers. For example, if there are 3 epochs from a stage out of 1000 epochs in total, these 3 epochs may be used for training and there may not be epochs for testing.

Literature research was made separately as single channel EEG and multichannel (EEG, EOG, EMG, ECG) signals.

4.1. Literature Comparison of Single Channel EEG Studies

In Table 4.1, previous studies are listed for classifying sleep stages as five class (wake, REM, N1, N2, N3).

Table 4.1. Previous Studies Classifying Sleep Stages Using Single Channel EEG Signal

Researchers	Feature	Classifier	Number of Epoch	Accuracy
Berthomier et al [58]	Frequency Domain	Fuzzy Analyze	8500	71.20%
Liang et al [59]	12 Different Features	Decision Tree	16105	77.98%
Proposed Method	**Time and Frequency Domain Features**	**KNN**	**9992**	**79.14%**
Proposed Method	**Time and Frequency Domain Features**	**Decision Tree**	**9992**	**79.33%**
Shen et al [60]	Improved Model Based Essence Features (IMBEFs)	Bagged Trees	104643	79.90%
Fraiwan et al [61]	Time-Frequency Domain	Random Tree	20269	83%
Hassan et al [62]	Statistical Features	Adaboost, Bagging	15188	83.49%
Liang et al [63]	21 Features; Multiscale Entropy (MSE) and Autoregressive (AR) Coefficients	Lineer Discriminate Analysis	8480	83.60%
Proposed Method	**Time and Frequency Domain Features**	**SVM**	**9992**	**83.94%**
Koley et al [64]	39 Different Features	SVM	15541	85.72%
Jinwoo Kim et al [65]	Statistical Feature	SVM	2305(5sec)	86%
Proposed Method	**Time and Frequency Domain Features**	**ANN**	**9992**	**86.23%**
Hassan et al [66]	Static, Spectral features	Bootstrap Aggregating (Bagging)	15188	86.53%
Zhu et al [67]	DVG and HVG Features	SVM	14963	88.90%
Hassan et al [68]	Spectral Features	SVM	15188	91.50%

As can be seen in Table 4.1, the accuracy rates obtained in this study are comparable to many studies. For example, Koley et al used SVM algorithm with 39 features and they obtained an accuracy rate of 85.72%. In this study, 83.94% accuracy rate was obtained with 20 features and SVM algorithm.

Almost the same accuracy rate was obtained with fewer features. In the same way Liang et al obtained a good result with an accuracy rate of 83.60% with 21 features. However in this study a better result was obtained with an accuracy rate of 86.24% with 20 features. In short, when compared to studies using single-channel EEG, in this study, higher accuracy rates were obtained than many studies.

4.2. Literature Comparison of Multichannel (EEG, EOG, EMG, ECG) Signal Studies

Table 4.2 shows previous studies that classified sleep stages into five classes using multichannel signals.

Table 4.2. Previous Studies Classifying Sleep Stages Using Multichannel Signals

Authors	Channel	Features	Number of Epoch	Clasifier	Accuracy percentages
Mora et al [69]	EEG, EMG, EOG	Symbolic Representation of EEG	10800	SVM	70%
Krakovská et al [70]	EEG, EMG, EOG, ECG	14Different Features	18058	ANN	81%
Proposed Method	**EEG**	**20Different Features**	**9992**	**SVM**	**83.94%**
Proposed Method	**EEG**	**20Different Features**	**9992**	**ANN**	**86.24%**
Özşen et al [71]	EEG, EMG, EOG	20Different features	4878	ANN	90.93 %

Mora et al obtained 70% accuracy rate by using 3 different signals in their study. Likewise, Krakovská et al. Obtained an accuracy rate of 81% with 4 different signals and 14 different features in their study. In addition to these, Özşen et al extracted 20 different features using 3 different signals and they achieved an accuracy rate of 90.93. In this study, an accuracy rate of 86% was achieved using a single channel EEG signal and is higher than many studies. As a result of the literature comparison, it can be said that the obtained accuracy rate in this study is acceptable and a higher accuracy rate can be obtained by increasing the number of signal channels used.

4.3. Literature Comparison of Studies with 5 Stages Classification

Table 4.3. Previous Studies Classifying 5 Sleep Stages

Researchers	Feature-Classifier-Epoch	WAKE	NREM1	NREM2	NREM3	REM
1. Diykh et al [72]	9 Features-SVM, kmean-5160	86.00%	88.00%	87.00%	79.00%	70.00%
2. Hsu et al [73]	Energy Based Features-Elaman Neural Classifier-2880	70.80%	36.70%	97.30%	89.70%	89.50%
3. Bajaj et al [74]	TFI based features-MC-LS-SVM-4700	88.00 %	75.48%	96.53%	92.42%	92.53%
4.Özşen et al [71]	57-ANN-4878	87.00%	88.47%	90.88%	86.42%	95.13%
5.Liang et al [59]	12-DT-16105	88.43%	35.12%	87.01%	90.8%	90.51%
6.Schaltenbrand et al [75]	17-ANN-61,949	91.73 %	4.67%	90.61%	86.86%	79.96%
7. Park et al [76]	4-Hybrid Rule-and Case-Based Reasoning-950	90.79 %	3.04%	87.38%	69.12%	53.76%
8.Proposed Method	**20 Time and Frequency Domain Features-KNN-9992**	**96.07%**	**94.55%**	**85.05%**	**93.80%**	**88.51%**
9.Proposed Method	**20 Time and Frequency Domain Features-SVM-9992**	**96.93%**	**94.35%**	**89.92%**	**95.41%**	**92.26%**
10.Proposed Method	**20 Time and Frequency Domain Features-DT-9992**	**95.70%**	**92.41%**	**86.22%**	**94.23%**	**89.60%**
11.Proposed Method	**20 Time and Frequency Domain Features-ANN-9992**	**97.96%**	**95.37%**	**90.04%**	**96.67%**	**92.44%**

The accuracy of the stages was taken separately and a literature comparison was made. Previous studies are listed in the Table 4.3.

It was observed that the 4 algorithms used in this study for NonREM1, NonREM3 and wake stages have higher accuracy than the studies in the literature.

NonREM3 values of the second, third and fifth study are higher than our study. However, NonREM1 values of these studies are 36.70%, 75.48% and 4.67% respectively.

The REM value of the fourth study was higher than our study. However, NonREM 1, NonREM 3 and wake values of the fourth study are 88.47%, 86.42% and 87%, respectively, which are lower than our study.

As a result, considering the results of the four algorithms used in this study, it can be seen that the values of all stages are above 85.00%. At the same time, the lowest overall accuracy is 79.30%. Therefore, this can be regarded as a reliable automatic sleep stage classification system.

4.4. Parameters Affecting the Accuracy Rate

4.4.1. Training Size

In this study, the feature matrix was divided into 50% training and 50% test matrix in order to train and test the algorithm. Since the obtained accuracy rate was 52%, it was found appropriate to divide the attribute matrix as 70% train and 30% test as a result of the researches. After this change, the accuracy rate was recorded as 65%. Therefore, the percentage of test and training matrices given to train the algorithm affects the accuracy rate.

4.4.2. Feature

In this study, firstly the first 8 features mentioned in the method section were used, but the features was tried to be increased since the accuracy rate was 56%. In new features, normalization was not applied because it affects the principle of the feature. The normalization process was used only for features deemed necessary. As a result of this observation, it has been proved that if the number of features raises, the accuracy rate will be increased.

4.4.3. Machine Learning Algorithm

The features values obtained in the study were first classified with the KNN algorithm at the classification stage. Since the obtained accuracy rate is 79.30% with this algorithm, it has been tried to be increased with other algorithms.

As the second algorithm, the classification was made with decision trees and an accuracy rate of 79.33% was obtained. In addition, the ratio was tried to be increased with the SVM algorithm and it was obtained as 84.63%.

Since it is thought that the accuracy rate will be higher with artificial neural networks, ANN was used and the accuracy rate was found to be 86.24%. In this way, it has been verified that the accuracy rate will be different with different classification algorithms. Accuracy rate can be increased with advanced algorithms.

In future studies, more efficient classification can be achieved by reducing the number of features by performing feature selection. That may also reduce the computational load. Moreover, higher accuracy rates can be achieved by changing the parameters that affect the accuracy rate.

REFERENCES

[1] An, P., Yuan, Z., Zhao, J., Jiang, X., & Du, B. (2021). An effective multi-model fusion method for EEG-based sleep stage classification. *Knowledge-Based Systems*, *219*, 106890.

[2] Rodríguez-Sotelo, J., Osorio-Forero, A., Jiménez-Rodríguez, A., Cuesta-Frau, D., Cirugeda-Roldán, E., & Peluffo, D. (2014). *Automatic Sleep Stages Classification Using EEG Entropy Features and Unsupervised Pattern Analysis Techniques. Entropy, 16(12), 6573–6589.*

[3] Khosla, A., Khandnor, P., & Chand, T. (2020). A comparative analysis of signal processing and classification methods for different applications based on EEG signals. *Biocybernetics and Biomedical Engineering*, *40*(2), 649-690.

[4] Rupasov, V. I., Lebedev, M. A., Erlichman, J. S., Lee, S. L., Leiter, J. C., & Linderman, M. (2012). *Time-Dependent Statistical and Correlation Properties of Neural Signals during Handwriting. PLoS ONE, 7(9), e43945.*

[5] Khalili, E., & Asl, B. M. (2021). Automatic Sleep Stage Classification Using Temporal Convolutional Neural Network and New Data Augmentation Technique from Raw Single-Channel EEG. *Computer Methods and Programs in Biomedicine*, 106063.

[6] Aboalayon, K., Faezipour, M., Almuhammadi, W., & Moslehpour, S. (2016). *Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation. Entropy, 18(9)*

[7] Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H., & Dickhaus, H. (2012). *Automated sleep stage identification system based on time–frequency analysis of a single EEG channel and random forest classifier. Computer Methods and Programs in Biomedicine, 108(1), 10–19*

[8] P. A. Abhang, B. W. Gawali, S. C. Mehrotra, Introduction to EEG- and Speech-Based Emotion Recognition, Academic Press, San Diego, CA, USA 2016, pp. 1–17.

[9] Şen, B., Peker, M., Çavuşoğlu, A., & Çelebi, F. V. (2014). *A Comparative Study on Classification of Sleep Stage Based on EEG Signals Using Feature Selection and Classification Algorithms. Journal of Medical Systems, 38(3).*

[10] Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberye, J. J. L. (2000). *Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. IEEE Transactions on Biomedical Engineering, 47(9), 1185–1194.*

[11] Yousif, E. S., Abdulbaqi, A. S., Hameed, A. Z., & Al-din, S. (2020, November). Electroencephalogram Signals Classification Based on Feature Normalization. In *IOP Conference Series: Materials Science and Engineering* (Vol. 928, No. 3, p. 032028). IOP Publishing.

[12] Akbulut, O. (2020). Feature Normalization Effect in Emotion Classification based on EEG Signals. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, *24*(1), 60-66.

[13] Ben-Gal. Outlier detection. In Data Mining and Knowledge Discovery Handbook, pages 131–146. Springer, 2005.

[14] Misra, S., Li, H., He, J., 2019. Machine Learning for Subsurface Characterization. Elsevier Science.

[15] A.Stine.R. (2001). *Statistic 910.* (Class notes). The wharton School of the Univeristy of Pennsylvania, Department of Statistic

[16] Christiano, L. J., & Fitzgerald, T. J. (2003). *The Band Pass Filter*. International Economic Review, 44(2), 435–465.*

[17] Saltzberg, B., Burch, N. R., McLennan, M. A., & Correll, E. G. (1957). *A New Approach, to Signal Analysis in Electroencephalography. IRE Transactions on Medical Electronics, PGME-8(0), 24–30.*

[18] Alakuş, T. B., & Türkoğlu, İ. (2017). Yapay sinir ağları kullanılarak epilepsi nöbeti öncesinin tahmin edilmesi. In *8th International Advanced Technologies Symposium–IATS* (pp. 510-516).

[19] Esteller, R., Vachtsevanos, G., Echauz, J., & Litt, B. (2001). *A comparison of waveform fractal dimension algorithms. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, 48(2), 177–183.*

[20] Bao, F. S., Lie, D. Y. C., & Zhang, Y. (2008, November). A new approach to automated epileptic diagnosis using EEG and probabilistic neural network. In *2008 20th IEEE International Conference on Tools with Artificial Intelligence* (Vol. 2, pp. 482-486). IEEE.

[21] Petrosian, A. (1995, June). Kolmogorov complexity of finite sequences and recognition of different preictal EEG patterns. In *Proceedings eighth IEEE symposium on computer-based medical systems* (pp. 212-217). IEEE.

[22] Vourkas, M. (2000). *Use of ANN and Hjorth parameters in mental-task discrimination. First International Conference on Advances in Medical Signal and Information Processing.*

[23] Davis PD, Parbrook GD, Kenny GNC: Basic Physics and Measurement in Anaesthesia. Oxford: Butterworth Heinemann; 1995:14–28

[24] Tonner, P. H., & Bein, B. (2006). *Classic electroencephalographic parameters: Median frequency, spectral edge frequency etc. Best Practice & Research Clinical Anaesthesiology, 20(1), 147–159*

[25] Moretti, D. (2004). *Individual analysis of EEG frequency and band power in mild Alzheimer's disease. Clinical Neurophysiology, 115(2), 299–308*

[26] Saby, J. N., & Marshall, P. J. (2012). *The Utility of EEG Band Power Analysis in the Study of Infancy and Early Childhood. Developmental Neuropsychology, 37(3), 253–273.*

[27] Salem, O., Naseem, A., & Mehaoua, A. (2014). *Epileptic seizure detection from EEG signal using Discrete Wavelet Transform and Ant Colony classifier. 2014 IEEE International Conference on Communications (ICC)*

[28] Krisnandhika, B., Faqih, A., Dewi Pumamasari, P., & Kusumoputro, B. (2017). *Emotion recognition system based on EEG signals using relative wavelet energy features and a modified radial basis function neural networks. 2017 International Conference on Consumer Electronics and Devices (ICCED).*

[29] Zhang, D. (2019). Wavelet transform. In *Fundamentals of Image Data Mining* (pp. 35-44). Springer, Cham.

[30] Adeli, H., Zhou, Z., & Dadmehr, N. (2003). Analysis of EEG records in an epileptic patient using wavelet transform. Journal of Neuroscience Methods, 123(1), 69–87. doi:10.1016/s0165-0270(02)00340-0

[31] J. Bruhn, H. Ropcke, and A. Hoeft, “Approximate entropy as an electroencephalographic measure of anesthetic drug effect during desflurane anesthesia,” Anesthesiology, vol. 92, pp. 715–726, 2000.

[32] Pincus, S. M., Gladstone, I. M., & Ehrenkranz, R. A. (1991). *A regularity statistic for medical data analysis. Journal of Clinical Monitoring, 7(4), 335–345.*

[33] Yan, R., & Gao, R. X. (2007). *Approximate Entropy as a diagnostic tool for machine health monitoring. Mechanical Systems and Signal Processing, 21(2), 824–839.*

[34] G. Wang, S. J. Shepherd, C. B. Beggs, N. Rao and Y. Zhang, "The use of kurtosis de-noising for EEG analysis of patients suffering from Alzheimer’s disease," in 4th International Conference on Biomedical Engineering and Biotechnology, Shanghai, 2015.

[35] Mendes de Paiva, L. R. et al. Analysis of the relationship between EEG signal and aging through Linear Discriminant Analysis (LDA). Rev. Bras. Eng. Bioméd. 28 (2)

[36] Hariadi M, Purnomo MH et al (2014) EEG signal identification based on root mean square and average power spectrum by using backpropagation. J Theor Appl Inf Technol 66(3)

[37] Tessy, E., Shanir, P. P. M., & Manafuddin, S. (2016). *Time domain analysis of epileptic EEG for seizure detection. 2016 International Conference on Next Generation Intelligent Systems (ICNGIS).*

[38] Bedeeuzzaman M, Farooq O, Khan YU. Automatic seizure detection using inter quartile range. Int J Comput Appl 2012 Apr;44(11):1–5.

[39] Karakuş C. (2001). *Makine Öğrenmesi.* (Class notes). Istanbul Technical University, Department of Engineer

[40] Noble, W. S. (2006). *What is a support vector machine? Nature Biotechnology, 24(12), 1565–1567.*

[41] Tharwat, A., Mahdi, H., Elhoseny, M., & Hassanien, A. E. (2018). *Recognizing human activity in mobile crowdsensing environment using optimized k -NN algorithm. Expert Systems with Applications, 107, 32–44.*

[42] Cunningham, P., & Delany, S. J. (2020). k-Nearest Neighbour Classifiers--. *arXiv preprint arXiv:2004.04523*.

[43] Cheng, D., Zhang, S., Deng, Z., Zhu, Y., & Zong, M. (2014). *kNN Algorithm with Data-Driven k Value. Lecture Notes in Computer Science, 499–512*

[44] Kesici, M., Saner, C. B., Mahdi, M., Yaslan, Y., & Genc, V. M. I. (2019). *An Optimal PMU Placement Scheme for Early Prediction of Transient Instabilities in Power Systems. 2019 7th International Istanbul Smart Grids and Cities Congress and Fair (ICSG).*

[45] Kamiński, B., Jakubczyk, M., & Szufel, P. (2017). *A framework for sensitivity analysis of decision trees. Central European Journal of Operations Research, 26(1), 135–159.*

[46] Sharma, H., and Kumar, S., A survey on decision tree algorithms of classification in data mining. International Journal of Science and Research (IJSR) 5(4):2094–2097, 2016.

[47] Chandra, B., & Varghese, P. P. (2007). *On Improving Efficiency of SLIQ Decision Tree Algorithm. 2007 International Joint Conference on Neural Networks*

[48] Sykes, A.O. An Introduction to Regression Analysis. Am. Stat. 1993, 61, 101.

[49] Sivagama Sundhari, S. (2011). *A knowledge discovery using decision tree by Gini coefficient. 2011 International Conference on Business, Engineering and Industrial Applications.*

[50] Herculano-Houzel, S. (2009). *The human brain in numbers: a linearly scaled-up primate brain. Frontiers in Human Neuroscience, 3.*

[51] Tek, F. B. (2021). An adaptive locally connected neuron model: Focusing neuron. Neurocomputing, 419, 306-321.

[52] Abraham, A. (2005). Artificial Neural Networks. Handbook of Measuring System Design.

[53] Péter, Á. Informatikai Tudományok Doktori Iskola.

[54] Livingstone, D. J. (Ed.). (2009). *Artificial Neural Networks. Methods in Molecular Biology™*

[55] Baddari, K., Djarfour, N., Aïfa, T., & Ferahtia, J. (2010). *Acoustic impedance inversion by feedback artificial neural network. Journal of Petroleum Science and Engineering, 71(3-4), 106–111.*

[56] Steyerberg, E. W., Harrell, F. E., Borsboom, G. J. J. ., Eijkemans, M. J. ., Vergouwe, Y., & Habbema, J. D. F. (2001). *Internal validation of predictive models. Journal of Clinical Epidemiology, 54(8), 774–781*

[57] Austin, P. C., & Steyerberg, E. W. (2014). *Events per variable (EPV) and the relative performance of different strategies for estimating the out-of-sample validity of logistic regression models. Statistical Methods in Medical Research, 26(2), 796–808*

[58] C. Berthomier, X. Drouot, M. Herman-Stoca, P. Berthomier, J. Prado, D. BokarThire, O. Benoit, J. Mattout, M.-P. d'Ortho, Automatic analysis of single-channel sleep eeg: validation in healthy individuals, Sleep 30 (11) (2007) 15871595

[59] Liang, S.-F., Kuo, C.-E., Hu, Y.-H., & Cheng, Y.-S. (2012). A rule-based automatic sleep staging method. Journal of Neuroscience Methods, 205(1), 169–176

[60] Shen, H., Ran, F., Xu, M., Guez, A., Li, A., & Guo, A. (2020). An Automatic Sleep Stage Classification Algorithm Using Improved Model Based Essence Features. Sensors, 20(17), 4677

[61] Fraiwan, L., Lweesy K., Khasawneh N., Wenz H., Dickhaus H., 2012. Automated sleep stage identification system based on time-frequency analysis of a single EEG channel and random forest classifier. Computer Methods and Programs in Biomedicine, 108(1), 10-9.

[62] A.R. Hassan, M.I.H. Bhuiyan, Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting, Comput. Meth. Progr. Biomed. 140 (2017) 201–210.

[63] S.F. Liang, C.E. Kuo, Y.H. Hu, Y.H. Pan, Y.H. Wang, Automatic stage scoring of single-channel sleep eeg by using multiscale entropy and autoregressive models, IEEE Trans. Instrum. Meas. 61 (6) (2012) 1649–1657

[64] Koley, B., & Dey, D. (2012). An ensemble system for automatic sleep stage classification using single channel EEG signal. Computers in Biology and Medicine, 42(12), 1186–1195.

[65] Kim, J. (2014). A comparative study on classification methods of sleep stages by using EEG. *Journal of Korea Multimedia Society*, *17*(2), 113-123.

[66] A.R. Hassan, S.K. Bashar, M.I.H. Bhuiyan, On the classification of sleep states by means of statistical and spectral features from single channel electroencephalogram, 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015, pp. 2238–2243

[67] G. Zhu, Y. Li, P. Wen, Analysis and classification of sleep stages based on difference visibility graphs from a single-channel eeg signal, IEEE J. Biomed. Health Inf. 18 (6) (2014) 1813–1821

[68] A. R. Hassan, M. I. H. Bhuiyan, A decision support system for automatic sleep staging from eeg signals using tunable q-factor wavelet transform and spectral features, Journal of Neuroscience Methods 271 (3) (2016) 107–118.

[69] Mora, A. M., Fernandes, C. M., Herrera, L. J., Castillo, P. A., Merelo, J. J., Rojas, F., & Rosa, A. C. (2010). Sleeping with ants, SVMs, multilayer perceptrons and SOMs. 2010 10th International Conference on Intelligent Systems Design and Applications.

[70] Krakovská, A., & Mezeiová, K. (2011). Automatic sleep scoring: A search for an optimal combination of measures. Artificial Intelligence in Medicine, 53(1), 25–33.

[71] Özşen, S. (2012). Classification of sleep stages using class-dependent sequential feature selection and artificial neural network. Neural Computing and Applications, 23(5), 1239–1250.

[72] Diykh, M., & Li, Y. (2016). *Complex networks approach for EEG signal sleep stages classification. Expert Systems with Applications, 63, 241–248.* doi:10.1016/j.eswa.2016.07.004

[73] Y.L. Hsu, Y.T. Yang, J.S. Wang, C.Y. Hsu, Automatic sleep stage recurrent neural classifier using energy features of EEG signals, Neurocomputing 104 (2013) 105–114.

[74] Bajaj, V., & Pachori, R. B. (2013). *Automatic classification of sleep stages based on the time-frequency image of EEG signals. Computer Methods and Programs in Biomedicine, 112(3), 320–328.* doi:10.1016/j.cmpb.2013.07.006

[75] Schaltenbrand N, Lengelle R, Toussaint M, Luthringer R, Carelli G, Jacqmin A, et al. Sleep stage scoring using the neural network model: comparison between visual and automatic analysis in normal subjects and patients. Sleep 1996;19:26–35

[76] Park, H.-J., Oh, J.-S., Jeong, D.-U., & Park, K.-S. (2000). Automated Sleep Stage Scoring Using Hybrid Rule- and Case-Based Reasoning. Computers and Biomedical Research, 33(5), 330–349. doi:10.1006/cbmr.2000.1549

[Url-1] https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Understanding-sleep#1 Access Date: 15.04.21

[Url-2] https://www.helpguide.org/harvard/biology-of-sleep-circadian-rhythms-sleep-stages.htm Access Date: 15.04.21

[Url-3] http://web.bvu.edu/faculty/ferguson/Course_Material/2011_Courses/Sleep_2011/Brain.html Access Date: 14.04.21

[Url-4] https://courses.lumenlearning.com/wsu-sandbox/chapter/stages-of-sleep/ Access Date: 15.04.21

[Url-5]https://www.sleepfoundation.org/how-sleep-works/stages-of-sleep Access Date: 15.04.21

[Url-6] https://physionet.org/content/sleep-edfx/1.0.0/ Access Date: 15.04.21

[Url-7] https://medium.datadriveninvestor.com/support-vector-machines-important-questions-a47224692495 Access Date: 11.04.21

[Url-8] https://www.wikiwand.com/en/Support-vector_machine Access Date: 11.04.21

[Url-9] https://medium.com/@ekrem.hatipoglu/machine-learning-prediction-algorithms-decision-tree-random-forest-part-5-2970905c021e Access Date: 15.04.21

[Url-10]https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Life-and-death-Neuron Access Date: 15.04.21

[Url-11] cognifit.com/tr/brain Access Date: 15.04.21

[Url-12] https://qbi.uq.edu.au/brain/brain-anatomy/what-neuron Access Date: 18.04.21

[Url-13]https://www.khanacademy.org/science/biology/human-biology/neuron-nervous-system/a/overview-of-neuron-structure-and-function Access Date: 18.04.21,

[Url-14] https://www.healthline.com/health/neurons#types Access Date: 18.04.21

[Url-15]https://teachmephysiology.com/nervous-system/synapses/synaptic-transmission/ Access Date: 18.04.21

[Url-16]https://www.electronicshub.org/artificial-neural-networks-ann/ Access Date: 18.04.21

[Url-17]https://www.nbshare.io/notebook/626706996/Learn-And-Code-Confusion-Matrix-With-Python/ Access Date: 18.04.21