AUTOMATIC SLEEP STAGE SCORING BASED EEG EVOKED RESPONSE

WASSIM DIAI

Master's Thesis

Graduate School
İzmir University of Economics
İzmir
2021

AUTOMATIC SLEEP ATAGE SCORING BASED EEG EVOKED RESPONSE

WASSIM DIAI

A Thesis Submitted to
The Graduate School of Izmir University of Economics
Master of Science Program in Electrical and Electronics Engineering

İzmir
2021

ABSTRACT

AUTOMATIC SLEEP STAGE SCORING BASED EEG EVIKED RESPONSE

Diai, Wassim

M.Sc. in Electrical and Electronics Engineering

Advisor: Asst. Prof. Dr. Faezeh Yeganli

January, 2021

Sleep experts frequently perform sleep assessments by analyzing the neurophysiological signals obtained by the patient in sleep laboratories. It’s a very tough, boring and time costly job basically. Restrictions of manual sleep stage recording expanded the need for the production of Automatic Sleep Stage Classification systems (ASSC). The designation of the sleep phases applies to the recognition of the different periods of sleep and is a crucial step in allowing doctors to recognize and treat associated sleep abnormalities. In an effort to define the study gaps and potentially incorporate a realistic approach, this work seeks to analyze progress and difficulties with many Electroencephalograms (EEG) including the evoked response and other approaches used in each of the phases for sleep staging, englobing the data processing, feature extraction and classification. In this thesis, the sleep-edf dataset of healthy subjects measures several classifiers. For a testing accuracy of over 85%, the optimized model proves remarkable progress. The findings illustrate the disparity between the different classifiers. Finally, respectable classification

accuracies can be obtained using the 2 EEG channels combined by healthy people. In fact, it is possible to generalize algorithms further so they can be used by more individuals.

Keywords: Machine Learning, Electroencephalograms (EEG), Biomedical Signal Processing, Feature Extraction, Feature Engineering, Sleep.

ÖZET

AUTOMATIC SLEEP STAGE SCORING BASED EEG EVIKED RESPONSE

Diai, Wassim

Elektrik ve Elektronik Mühendesliği Yüksek Lisans Programı

Tez Danışmanı: Asst. Prof. Dr. Faezeh Yeganli

Ocak, 2021

Uyku uzmanları sıklıkla uyku laboratuvarlarında ki hastalardan elde ettiği nörofizyolojik sinyalleri analiz ederek uyku değerlendirmeleri yaparlar. Bu iş temelinde çok zor, sıkıcı ve zamana mal olan bir iştir. Manuel uyku aşaması kaydının kısıtlanması, Otomatik Uyku Aşaması Sınıflandırma sistemlerinin (ASSC) üretimine olan ihtiyacını artırdı. Uyku evrelerinin belirlenmesi, farklı uyku dönemlerinin tanınması için gereklidir ve doktorların ilişkili uyku anormalliklerini tanımasına ayrıca tedavi etmesine izin veren çok önemli bir adımdır. Çalışma boşluklarını tanımlamak ve potansiyel olarak gerçekçi bir yaklaşımı dahil etmek için, bu çalışma Elektroensefalogram (EEG) konusunda yer alan uyku evreleme aşamalarının her birinde kullanılan uyandırılmış yanıt ve diğer yaklaşımlar dahil, veri işleme, özellik çıkarma ve sınıflandırma konularında ki ilerlemeleri ve zorlukları analiz etmeyi amaçlamaktadır. Bu tezde, sağlıklı deneklerin uyku-edf veri seti birkaç sınıflandırıcıyı ölçmektedir. % 85'in üzerinde bir test doğruluğu için optimize edilmiş model, kayda değer bir ilerleme olduğunu kanıtlıyor. Bulgular, farklı sınıflandırıcılar arasındaki eşitsizliği göstermektedir. Son olarak, sağlıklı kişiler tarafından birleştirilen 2 EEG

kanalı kullanılarak ilgili sınıflandırma doğrulukları elde edilebilir. Aslında, algoritmaları daha fazla kişi tarafından kullanılabilmesi için daha fazla genellemek mümkündür.

Anahtar Kelimeler: Makine Öğrenme, Elektroensefalogramlar (EEG), Biyomedikal Sinyal İşleme, Özellik Çıkarma, Özellik Mühendisliği, Uyku.

To Mom, Dad, Ayman and Salma

Dedicated to everyone who wonders if I’m writing about them, I am.

Dedicated to People with neurological disabilities and diseases.

ACKNOWLEDGEMENTS

In the name of ALLAH (SWT) the most generous and the most merciful. All thanks is to Allah (SWT) with all the blessings that the author can accomplish this work. More- over, may the Prophet Muhammad (PBUH), who brought every man from darkness to lightness, be granted peace and salutations.

Firstly, I would love to thank Asst. Prof. Dr. Faezeh Yeganli, my Thesis supervisor from Izmir University of Economics, Computer Engineering department. When I stumbled into a tough situation or had a concern about my work and writing, the door of Prof. Yeganli's office was always open. This study was constantly allowed to be my own project, but when she felt it was required, she directed me in the right direction.

I would also like to appreciate the second reader of this dissertation, Dr. Prof. Aydın Akan Head of the Electrical and Electronics Engineering Department at the Izmir University of Economics for his valuable comments. I thank him for his useful remarks. The validation survey could not have been accomplished without his passionate participation and feedback.

Not least of all, for their continuing aid and motivation through the years of study, I am appreciative to my parents as well as during the research and writing process of this research. Without them, this success would not have been possible. Thank you. Thank you. Thank you.

Finally, for their constant encouragement and their relentless conviction that I achieve best, it's so important to my whole family. I can't thank you all unfortunately, on behalf of you because it lasts a lifetime, so I do want you to know that you are so counting. I would never have finished this study if this weren't about all your prayers and blessings, if it weren't for your true affection and support. Thank you all. Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Sleep is a vital necessity that allows our body and brain to relax and benefit from more physiological balance, keeping us refreshed when we get up. Having enough lively sleep allow the body to refresh and survive and continue recharged and avoid diseases. Without enough sleep, our brain wouldn't behave properly. It can damage our mental skills, our capacities to work, think and store memories (Mukherjee et al., 2015).

Seven to nine hours of nocturnal sleep would be enough for adults. Various factors can forbid us from enough sleep (stress, work schedules…). A good nutrition and concrete habit can guarantee an adequate amount of sleep nocturnally. However, persistent in-sufficiency of sleep may be the red alarm to ring for a sleep disorder.

When we mention "sleep disorder" we can't turn our eyes off without mentioning the most common abnormality which is the sleep apnea, in which the breathing operation stops and starts repeatedly. Another common type of sleep disorder is Insomnia, in which the patient has difficulty falling asleep or staying asleep, most of times treatment connect a combination of medical treatments and new lifestyle regime.

Indeed, sleep stages scoring can be decisive to regulate sleep and ensuring the patient comfort. It can give us an idea about the process which remain closely to understand better how the brain handles sleep. The EEG surveys the sleep classes, then analyzed and interpreted by specialists. It can be our guiding light to present a clear picture about sleep disorders.

In heart, brain and mental disorders, diagnosis and treatment in fields of medicine, electroencephalography (EEG) have long been involved (Siuly, Li, and Zhang, 2016). EEG signals analysis has proved its worth and recognized as a notorious method to the issue of withdrawing information of Human brain.

Nowadays, EEG are widely analyzed for several health care reading: brain computer interfaces, sleep stage scoring, apnea detection and epilepsy seizure detection (Siuly, Li, and Zhang, 2016). An EEG major advantage is to breakthrough brain problems detecting disorders. However, a plenty of EEG channels are visualized, analyzed by clinicians to have better idea about abnormalities which can propagate and be more dangerous in the nearest future. This project target to classify sleep into different stages with respect to the AASM and Rechtschaffen Kales (Moser et al., 2009), implementing various signal processing and segmentation techniques, essentially filtering, feature extraction and data normalization, we adopted the evoked response which is an electrical potential in a specific pattern recorded from a definite side of the nervous system, especially the brain, applying Fpz-Cz and Pz-Oz. We relate to a certain time frame while introducing EEG Epoching, which is derived from a permanent EEG signal called ’epochs’ and is usually time-closed in reference to an activity or describing an operation, in another word, the epochs define a set of EEG signals which represent the same sleep phase.

In this thesis, we test and evaluate some machine learning models including Support Vector Machine (Chatterjee, and Bandyopadhyay, 2016), Random Forest (Reddy Edla et al., 2018), Decision Tree (Guan, Zhao, and Yang, 2019), Multilayer Perceptron (Chatterjee, and Bandyopadhyay, 2016), and many others. The evaluation focuses on the task of sleep classification using the sleep physionet dataset (Imtiaz, and Rodriguez-Villegas, 2015); for this purpose, we preprocessed the data in a way that we clean it from noise and obtain the two ordered EEG channels from the polysomnography, 30s events from annotations then were extracted.

In this project, we will clarify how the statistical values based discrete wavelet transform coefficients were viable to use as input for our classification models. The classifiers take two vectors (X, Y), first represents the extracted features, while Y present the target Labels (Sleep Stages). The work demonstrates that with first divided and merged EEG channels later we have identified a short Methodology in which sleeping phases can be graded with the best methods of machine learning as well as to achieve 85 % of testing accuracy. As a bonus we will investigate the use of other channels like EOG, ECG and EMG submental.

CHAPTER 2: LITERATURE REVIEW

2.1 Sleep Overview

Sleep is defined as an indispensable need for our daily regime, that it makes up about a third of our lifetime's. As water and food, benefiting from enough healthy sleep at recommended times is as crucial to survival. sleep confusion, deprivation were behind different mental and emotional issues (Brown, 2012) (Killgore, 2010). Sleep can occur in all vertebrates, including birds, fish and reptiles, alongside human beings and animals (Cirelli, and Tononi, 2008). In order to guarantee an efficient communication between the body organs and the brain neurons, sleep represent the bridge between the mental and physical health in parallel with daily life challenges. It lies behind all essential functions within human beings, including the nervous system, brain, skin, leaf and physiology, cognitive system, behavior and sensitivity to disease. Scientists prove the sleep deficiencies or sleep usage in un- healthy situations or in conjunction with disabilities of some sort, can threaten human health and disrupt the wheel of the human body's mechanisms, as well as the risk of developing diseases such as diabetes, heart, arterial problems and obesity in some cases beside depression. Thus, research suggest sleeping healthy in order to remove fatigue and help the body work to balance blood pressure and get rid of toxins acquired during the day that constitute a negative addition the organs responsible for energy production and antibiotics beside strong natural immunity.

2.2 Sleep Stages

2.2.1 Two Types of Sleep

The normal phases of sleep have already been identified correctly by the specialists. Sleeping specialists choose to be separated into two main phases, rapid eye movement o (REM) and non-rapid eye-movement (NREM). The above may in turn be classified into phases 1, 2, 3 and 4, which are a continuum in relative depth. Everyone has special properties, involving changes in the rhythm of the brain wave, eye motions and tone of muscle. Electro-encephalographic (EEG) recording tracing electrical signals of brain functions has exposed sleep cycles and phases (Loomis, Harvey, and Hobart, 1937) (Dement, and Kleitman, 1957).

NREM and paradoxical sleep switch cyclically over a duration of sleep (Figure 1). The function of varying between these two sleep categories is not yet established, but sleep disturbances include abnormal cycling and/or missing sleep classes (Kryger, Roth, and Dement, 2005).

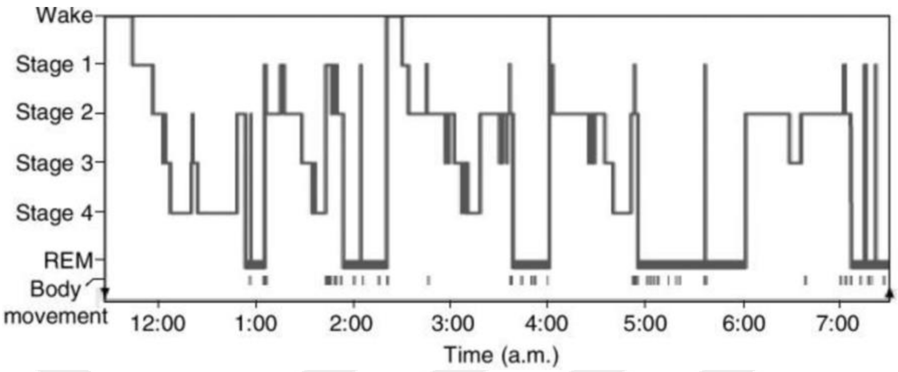

Figure 1. Sleep Stages During a Single Night. (Source: Kryger, Roth, and Dement, 2005)

2.2.2 NREM and REM Sleep Cycles

The sleep ration starts in parallel with NREM stage 1, gradually moving to the stage 2, stage 3, 4 until it reaches the REM stage. It must be noted that the individuals do not settle into one single stage of sleeping throughout the hall night, but rather there is an alternation between the REM and NREM stages (Figure 1). REM sleep accounts for between 20 and 25% of overall sleep spent, while NREM accounts for 70% to 80% of sleep spent. A Mean period of 70 to 100 minutes for the first stage of NREM-REM sleep. The second and subsequent periods last about 90 to 120 minutes (Kryger, Roth, and Dement, 2005). Sleep is paradoxical and is longer in the last third of the night period of ordinary adults as the evening progresses. Stage 2 continues to account for the bulk of orthodox sleep as the series of sleep increases, while stage 3 and stage 4 frequently totally disappear.

2.2.3 NREM Sleep Classes

During NREM sleep, 4 stages are each correlated with distinct physiology and brain action. The signals of EEG patterns typical of the four NREM phases are shown in the figure 2. Other instruments are being used to map the signature variations in muscle sound and eye motion. The image shows a K-complex in stage 2, with two spindles in sleep for underlining.

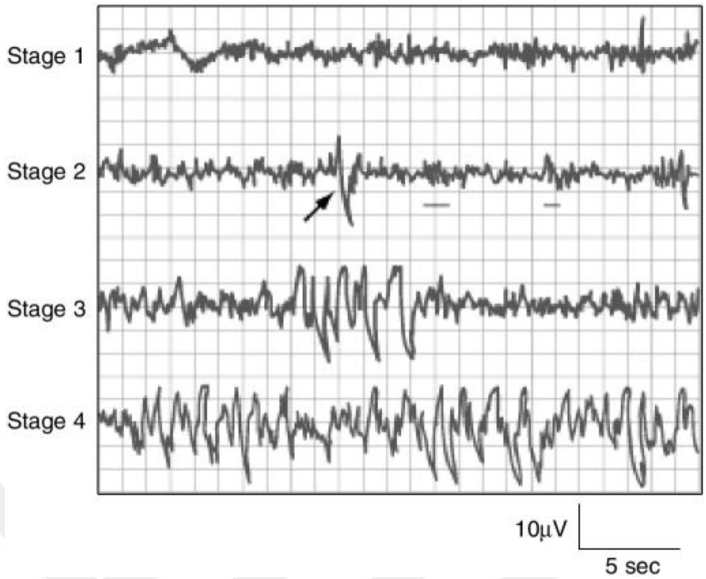

Figure 2. NREM Four Stages Characteristics (EEG). (Source: Kryger, Roth, and Dement, 2005).

Sleep Stage 1:

In sleep-stage pedaling, NREM phase 1 sleep forms a temporary function. Besides newborns, people with narcolepsy and other specific neurological conditions, the series of sleep of the normal human starts in first phase of NREM. In the initial period, this period typically holds out from 1 to 7 minutes, accounting for 2 to 5% of overall sleep, and is quickly disturbed by rowdy sounds. In stage 1, brain function on the EEG shifts from wakefulness (marked by rhythmic alpha waves) to mixed-frequency low-voltage waves. A wakeful relaxation condition is synonymous with alpha waves and it has an eight to 13 period per second frequency (Kryger, Roth, and Dement, 2005).

Sleep Stage 2:

Step 2 sleep spends around 10 to 25 minutes in the first period and rises over every subsequent cycle, essentially reflecting 45 to 55% of sleep total. In second sleep phase a person needs to wake up more extreme stimuli than in stage 1. Various frequencies operation with the participation of sleep spindles and K-complexes, comparatively low voltage, can be seen in brain functions of EEG (Figure 2). Sleep spindles are hypothesized to be essential for memory consolidation. Individuals learning a new task have a much higher sleep spindle density than people in a control group (Gais et al., 2002).

Stage 3 and 4, Slow-Wave Sleep:

Third and fourth sleep phases are listed mutually as slow-wave sleep (SWS), much of which happens in the first night's third. Each has distinctive features. Step 3 duration is just a few minutes and accounts for around 3 to 8% of night sleep. The EEG demonstrates elevated behavior of high voltage, slow wave (Figure 2). Step four is the last NREM process in which 10 to 15% of the sleep takes approximately 20-40 minutes in the significant levels. The threshold for anticipation is the greatest in the fourth step of all NREM phases. This is highlighted by high-voltage activity levels on the EEG and slow waves (Kryger, Roth, and Dement, 2005).

2.2.4 REM Sleep

REM sleep is characterized by the occurrence of desynchronized activity of brain waves (low voltage, mixed frequency), muscle atonia and rapid eye movement bursts (Carskadon, and William C, 2005)." Saw- tooth," theta activity and sluggish alpha activity are both characteristic to REM sleep, and they have 3-7 counts per second. The REM time can last just 1-5 minutes during the initial cycle; however, it steadily extends with the progress of the sleep episode (Carskadon, and William C, 2005). NREM and REM sleep are subject to different physiological variations (Table 1).

2.3 Electroencephalogram and Polysomnography

2.3.1 EEG Definition

EEG (Electroencephalography) indicates the potential calculation that represents electrical behavior of a human brain. Which is an easily accessible operation which demonstrates how over time the brain behaves. EEG is commonly used in the study of brain processes and neurological diseases by doctors and biomedical experts. The study of electricity in the brain using the EEG documents is one of the main methods for neurology evaluation such as seizures, brain tumors, headache, sleep disturbances, dementia and anesthetic control throughout surgery (Hazarika et al., 1997) (Adeli, Zhou, and Dadmehr, 2003). Hans Berger brought the first EEG recording machine into the world in 1929. The German word" electrenkephalogram," used by Berger (Collura,

1993) a neuropsychiatrist at Jena University, Germany, to prove the graphical representations of electrical subjects in the brain.

Table 1. Differences Between NREM and REM

Physiological act	NREM	REM
Brain activity	Wakefulness declines	Engine and sensory rises
Heart rate	Bogs down from consciousness	Increases and varies
Blood pressure	Wakefulness declines	Increases (up to 30 percent)
Sympathetic nerve activity	Wakefulness declines	Wakefulness rises
Blood flow to brain	Wakefulness declines	Increases from NREM

It indicated that currents in the brain shifted due to the physiological state of the brain, including sleep, anesthesia and epileptics. It was a groundbreaking theory that helped develop a modern medical research branch called neurophysiology. A variety of small disks labeled electrons are positioned with short term glues on the scalp zone at various positions during the EEG test. The amplifier (One boost per electrode couple), an EEG grabbing system are attached to each electrode. Last but not least, on a computer screen, brain neural waves are translated to wavy lines to monitor results. Figure 3 gives an example of how electrodes on the scalp are mounted on a computer screen when the EEG signal is registered. The electrodes sense minor electrical charges generated by brain cell activity. The charges are amplified and can be written on paper as a graph on the screen of the device. The read is then read by a professional.

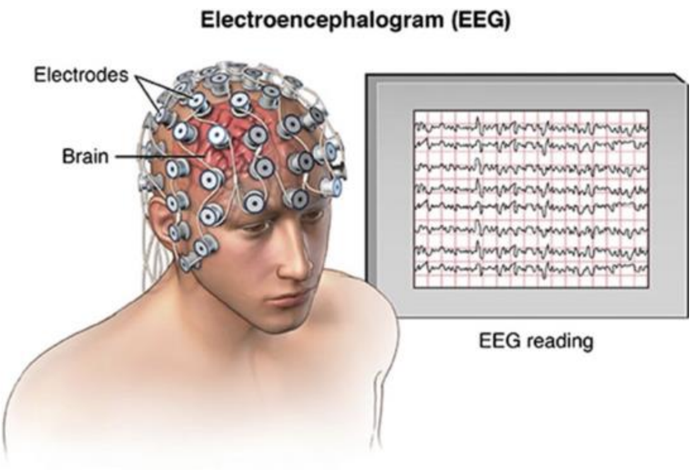

Figure 3. A Sample of EEG Recording (Source: Siuly, Li, and Zhang, 2016)

EEG records may have between 1 and 256 electrodes registered in parallel depending on their application. The word EEG records are multi-channel. A channel normally consists of one pair of electrodes. During EEG recording, each channel generates a signal. Depending on where the message in the head is taken, two forms of EEGs exist: scalp and intracranial. Small electrodes with strong electrical and mechanical interaction are mounted on the scalp of the scalp EEG. During procedure the intracranial EEG results in special electrodes inserted in the brain. The EEG is designated the electrocorticogram from the cortical surface. by using subdural electrodes (ECOG). In normal adults, the EEG magnitude usually varies from 1 to 100 V and when tested with subdural electrodes, such as needle electrodes, it is about 10-20 mV. Although the brain architecture is standardized and the cortex functional, EEG can differ based on where the recording electrodes are located. The topic of how the electrodes is mounted is critical because the various lobes of the brain cortex process different kinds of activity. The international 10-20 electrode scheme is the traditional way of locating scalp electrodes (Klem et al., 1961). The ”10” and the ”20” reflect true differences of 10 to 20 percent of the entire front or the left lengths of the skull between the neighboring electrodes. The following two points decide the positions: nasal location between the forehead and the nose, eye level and ossify prominence on the midline at the base of the skull behind the head.

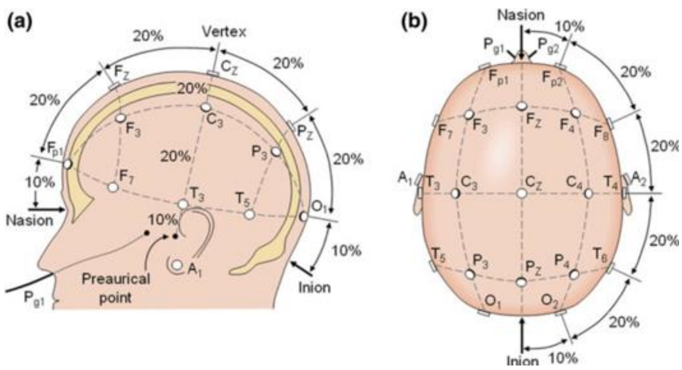

Figure 4. 10- 20 Electrode Localization Model (Source: Siuly, Li, and Zhang, 2016)

The figure 4 indicates the location of the electrode in the brain by the international 10-20 method. The lobe is marked by a letter at any position and the location is identified by a number. The letters F, T, C, P and O are Frontal, Temporary, Central, Parietal and

Occipital. An electrode positioned in the midline is referred to as a" z". The electrode locations on the right hemisphere are even numbers, while those on the left hemisphere are odd numbers. Since an EEG tension signal is a variation within the voltages of two electrodes, the EEG monitor can be mounted in many ways for the EEG reading system. The electrode positioning is known as an assembly.

2.3.2 EEG Signal's Structure

The human brain is made up of about 100 billion nerve cells called neurons, and these neurons sustain the brain's electrical burden. Neurons have the same features and have the same components as other cells; however, they relay electric signals and transfer orders to one another For long trips due to their electrochemical character. There are three essential bits of neurons: cell body (soma), axon and dendrites (Carlson, 2005) (Ghaedi et al., 2019) as shown in Figure 5 Neurotransmitter receptor sites which can also be sent with a paired axon.

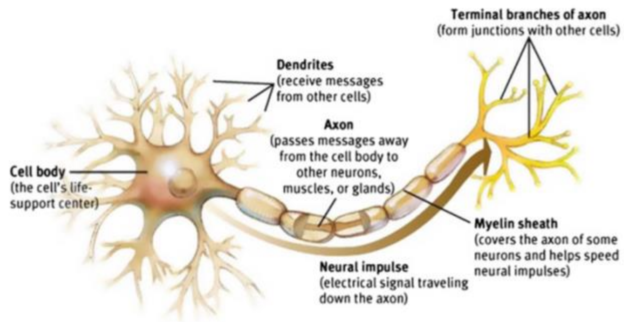

Figure 5. Structure of a Neuron (Source: Sanei, and Chambers, 2013)

On either of these ends of the cell, dendrites can be found. Neurons can interact with each other via the axon-dendrite connection. The action potential is an occurrence in which the ion pumps along the exterior side of an axon, quickly altering the ionic structure of the axon so that an electric pulse will pass rapidly from the axon to the next dendrite (Siuly, Li, and Zhang, 2016). As the ionic charge varies immediately, a voltage appears both within and outside the neuron's cell membrane (Carlson, 2005) (Schaul, 1998) (Ghaedi et al., 2019). These neurons emission a chemical called neurotransmitters (Carlson, 2005) (Ghaedi et al., 2019). Figure 5 displays the interneuron coordination system. The current flow that contributes to surface EEG is demonstrated in figure 6 during a net excitational input. Local current flows arise when neurons are stimulated by an electrochemical concenter. Neurons can be separated into two sub-sets of electric activity: action potential (AP) and post-symptomatic potential (PSP). When the PSP exceeds the postsynaptic neuron threshold level, neuron fires are started and an AP is initiated (Siuly, Li, and Zhang, 2016).

Low frequency summed inhibitory and excitatory PSPs within pyramidal neuronal cells establish the electric dipoles between soma and apical dendrites that record electrical potential on the surface (Figure 6). These PSPs summarize the cortex and spread to the scalp surface of the EEG recorded.

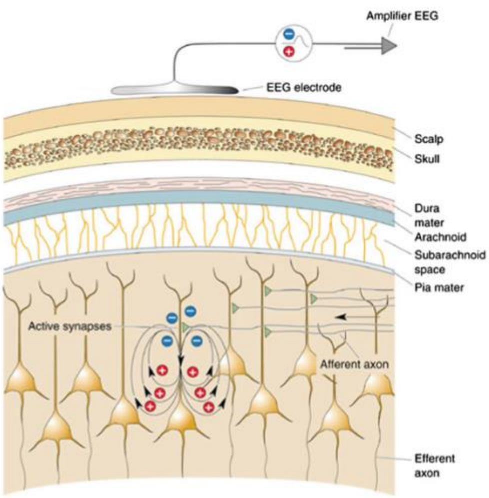

Figure 6. Exemplar of the Formation of Very Small Electrical Regions in Pyramid Cells by Neuronal Currents (Source: Freeman, and Zhai, 2009)

Nerve cell APs are much shorter in their theoretical field distribution than PSPs. Consequently, APs make no major contribution to scalp or therapeutic EEG registrations. On the scalp (Carlson, 2005) (Schaul, 1998) (Ghaedi et al., 2019) only large populations of activated neurons can produce record-able electrical activity. The voltage is usually too small to be calculated precisely with the current technology when created from a single cell.

2.3.3 EEG Signals Design and Features

Frequency is one of the key factors for the diagnosis of medical EEG abnormalities and for recognizing the cognitive research’s adaptive behaviors. Frequency means repeated rhythmic operation (in Hz). As frequency, the number of cycles in a second. The human EEG potential is manifested as an aperiodic random oscillation, with millions of vibrating neuron populations as its source. With sporadic oscillations. The amplitudes and rhythms of these signals vary in stable adults from state to state such as sleep and awakeness. 5 main brain waves of varying frequency ranges are present. Specific bands such as 0.5-4 Hz (delta, d), 4-8 Hz (theta, h), 8-13 Hz (alfa, a), 13-30 Hz (beta, b) and more than 30 Hz are listed in these frequency bands, from low to high frequency bands (gamma, c) (Niedermeyer, and Silva, 2005) (Fisch, and Spehlmann, 1999).

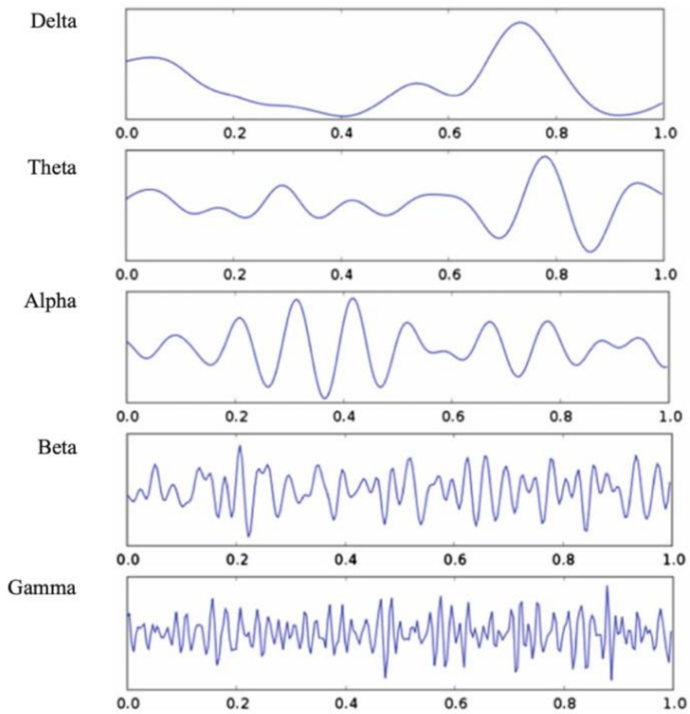

Figure 7. Normal EEG Rhythms (Source: Siuly, Li, and Zhang, 2016)

In pathological brain disorders such as epilepsy, higher frequencies are also more prominent. Descriptions of this pattern of EEG are given by the figure 7. The Delta wave varies from (0.5-4) Hz, and the form is the strongest and the largest in amplitude between waves. It is mostly linked to deep sleep, extreme brain and waking conditions. Amplitude is typically greater than 20 V for the Theta wave scales between 4 and 8 Hz. The theta comes from emotional tension, in particular from anger or deceit and unconscious content, artistic motivation and deep contemplation. The alpha contains the frequency spectrum of 8 to 13 Hz, with an amplitude of 30-50 m V, and is located predominantly in the back of the head (occipital lobe), when the eyes of this subject are closed or relaxed. It is typically related to extreme psychiatric activity, depression and stress. Mu activity is also considered alpha activity from sensorimotor zones. The beta is in 13-30 Hz. Beta is available. It is observed symmetrically on both sides in the frontal region at low amplitude and varying frequencies. The brain produces beta waves while it is excited and constantly involved in mental function. These waves are the qualities of a healthy power. Beta is the brain wave typically related to active things, active interest and an external emphasis or the resolution of concrete issues. The frequency of gamma waves is 30 Hz and beyond. Often this rhythm has a maximum frequency of about 80 to 100 Hz. It is related to multiple cognitive and motor functions. The figure 8 of regular EEG recording is shown. This illustrates the average or regular amount of beta activation in an awake EEG. Beta behavior is also easier to detect when awakening or drowsiness is comfortable. Electrical impulses from non-cerebral origins in EEG are known as artifacts.

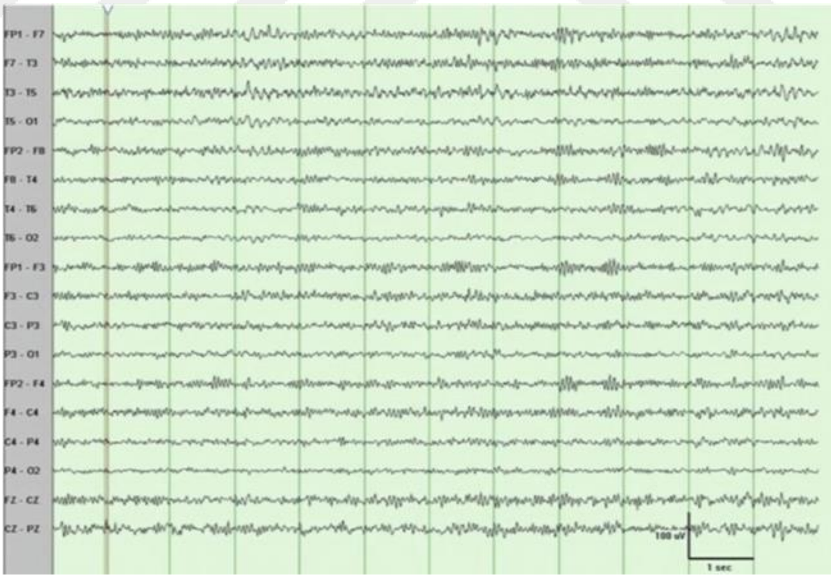

Figure 8. Regular EEG Recording (Source: Siuly, Li, and Zhang, 2016)

EEG data are virtually often infected by such devices. The artifacts' amplitude is widely related to the scale of the desired cortical signals. It belongs to substantial expertise causes of scientifically accurate interpretation of EEGs. Figure 9 indicates the most common four artifact forms of human EEG recording.

1- Electrooculographic system triggered by muscle arousal (related to blinking, for example). A major amplitude in frontal electrodes is prominent, sluggish, positive pulse.

2- The artefact of electrode caused by poor interaction between P3 electrode and skin (and hence greater impedance).

3- Swallowing artefact.

4- Common electrode reference anomaly due to terrible skin/depleting touch. Big wave on all channels identical.

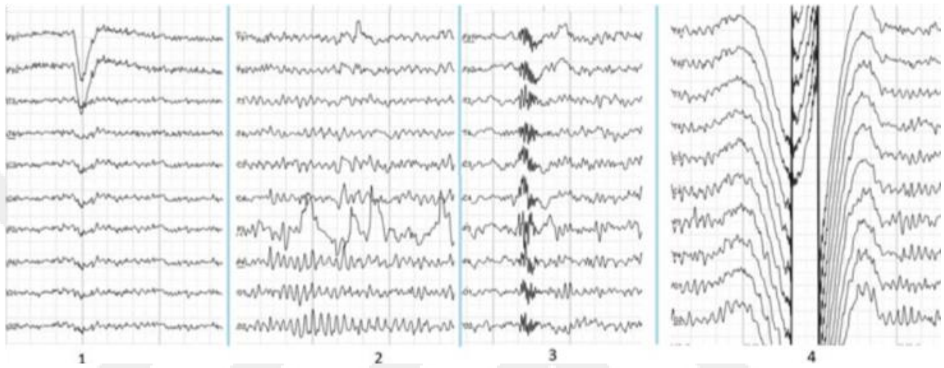

Figure 9. An example of the Greatest Kinds of EEG Artefacts (Source: Bin Heyat, and Siddiqui, 2015)

2.4 Data Preprocessing

Data are used to input and provide the learner with decision-making data when the algorithm is a grading or regression. Ideally, the extraction or collection of features as a separate procedure is not necessary in machine learning; the classifier must use some function to exclude those that are meaningless. The sophistication of a learner is determined by the data, consistency and quantity. This affects the nature of both space and time as well as the amount of training examples needed for a learner (Alpaydin, 2020). In the following lines we discuss about methods of extraction which are less new to the original features. Where the data has a large number of characteristics, its size must either be reduced, or a lower-dimensional representation can be sought that retains those properties.

2.4.1 Feature Extraction Using Wavelet Transform

Feature extraction works to enhance the success of data collection and retrieval by identifying the most distinguishable, descriptive and minimized elements or features. Significant vectors for classification problems remain the most common and suitable signal presentation (Subasi, 2019). The wavelet transform defines the time-frequency properties of a waveform in another way. But rather than time segments, the waveform may be divided into scale segments (Semmlow, 2004). Wavelets consist of two parameters, one for time scaling and one for time sliding. A wavelet is a time-concentrated energy oscillating function for improved transient signals. The band-pass filter functionality is only one of several mathematical aspects a wavelet function should include. Wavelet analysis attempts at both time and frequency to reach a suitable position. Two new degrees of freedom, sliding and scaling, allow study of precisely built structures and global waveforms in signals. The essential concept of evaluation of signals in various scales with an increasing resolution level is described by a multiresolution analysis (Sörnmo, and Laguna, 2005)

2.4.1.1 The Continuous Wavelet Transforms (CWT)

A wavelet family $\Psi(s, \tau)$ is differentiated by sliding and scaling $\Psi(t)$ mother wave with τ and s parameters which are continuously evaluated.

$$\Psi(s, \tau)(t) = \frac{1}{\sqrt{s}}\left(\frac{t-\tau}{s}\right) \tag{1}$$

Where $\frac{1}{\sqrt{s}}$ component assures the same energy is available to all the scaled functions. The wavelet expands for s >1 and contracts for 0 <s <1. The $\Psi(s, \tau)(t)$sampling function still has an oscillating pattern. The mother's wavelet Ψ(1,0) (t) ≡ $\Psi(t)$, takes its normal shape for s=1 and τ =0, along with other members of its family generated with dilation and contracture. The fact that a wavelet is contract for a smaller time scale means that it is more located and located more often due to the increased bandwidth of the bandpass response and the change to higher frequency ranges. A contrast of the x(t) signal with the probing function $\Psi(s, \tau)$, is defining a continuous wavelet transforming (CWT) $(s, \tau)(t)$of an x(t) continuous-term signal.

$$\omega(s,\tau) = \int_{-\infty}^{+\infty} x(t)\Psi(s,\tau)(t)dt \tag{2}$$

Build two-dimensional time domain mapping. The CWT can be treated as a linear filter since the above equation reflects the convolution between the signal $\mathrm{X}(t)$ and the impulsively response filter $\frac{\Psi(t/s)}{s}$. As the CWT splits the waveform into coefficients between the two factors s and τ, the initial waveform of wavelet coefficients needs to be restored with a double integration (Bodenstein, Schneider, and Malsburg, 1985):

$$x(t) = \frac{1}{C(\psi)} \int_{-\infty}^{\infty} \int_{0}^{+\infty} \omega(s,\tau)(t) \frac{d\tau ds}{s^2} \tag{3}$$

Where:

$$C(\psi) = \int_{0}^{\infty} \frac{|\psi(f)|^2}{|f|} df < \infty \tag{4}$$

and Ψ(f) represents the Fourier transform of Ψ(t). The easiest wavelet is the haar wavelet, which belongs to the Walsh's. The Mexican hat wavelet, defined by the equation bellow, is another common wavelet:

$$\Psi(t) = (1 - 2t^2)e^{-t^2} \tag{5}$$

The Morlet wavelet is given by the equation, named according to a pioneer of wavelet analysis:

$$\Psi(t) = e^{-t^2} cos\left(\pi \sqrt{\frac{2}{ln2}} t\right) \tag{6}$$

There have been proposals for a vast range of wavelets, each of which has something especially suited for those applications. Wavelets interchange the location of time and frequency (Subasi, 2019).

2.4.1.2 The Discrete Wavelet Transforms (DWT)

It is a redundancy which produces innumerable coefficients that are very important to correctly represent the original input signal. This is the main issue for CWT. This redundancy is only costly if you have to rebuild the first signal, since the estimation struggle is rendered quite pointless with all the coefficients. The discrete wavelet transform (DWT) generates the frugality of the coefficient by limiting scaling variance and sliding to power of 2; thus, it is often called the transformation of the dyadic wavelet transform with the same abbreviation (DWT). However, from the discrete coefficients of the dyadic Wavelet Transform, we can still correctly construct the original signal (Bodenstein, Schneider, and Malsburg, 1985). The DWT also constitutes a non-redundant bilateral change as it is a part of orthogonal family (Barni, and Bartolini, 2004). the discrete wavelet transform (DWT)is presented as:

$$\omega(j,k) = \int_{-\infty}^{+\infty} x(t)\psi(j,k)(t)dt \tag{7}$$

2.5 Machine Learning

2.5.1 Machine Learning Overview

Machine learning is characterized as methodologies using experience to boost efficiency or accurately predict. Experience reveals the prior knowledge for the learner, which is naturally gathered and made available for investigation by the electronic data records. These data may be in the form of digitized training sets for human beings or other knowledge gathered from ecosystem experiences. Machine learning is meant to con- struct predictive algorithms competently and reliably. As in other programming fields, their space and time complexity are important parameters for the consistency of these approaches. Since a learning method's output is dependent on the input and functions used, machine learning is usually related to analytical information and statistics. Algorithms of interest are usually input based approaches that incorporate essential computer science principles with concepts such as probability, statistics and optimization. In addition, such applications are dealing with large trees of learning concerns. The major categories of knowledge difficulties are grouping, regression, classification, clustering and Minimization of dimension (Mohri, Rostamizadeh, and Talw, 2018).

2.5.2 Machine Learning Framework:

A machine learning background must begin with the compilation of data and then turn it into helpful data using machine learning techniques in order to solve an empirical problem or to discover actual machine learning. A map of machine learning consists basically of input recovery, extraction components, training, simulation and implementation (Harrington, 2012) (Sarkar, Bali, and Sharma, 2018).

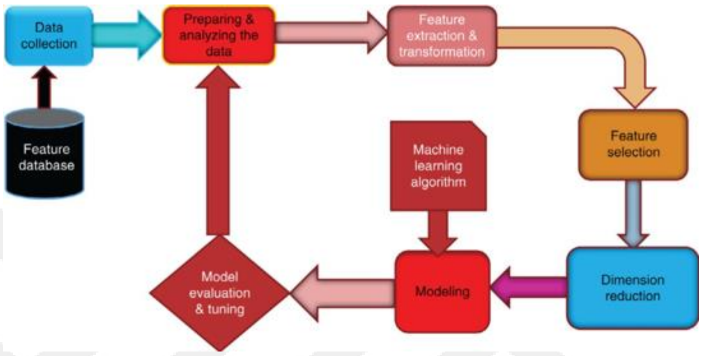

Figure 10. Machine Learning Framework.

2.5.3 Supervised Machine Learning

To generate models by supervised, labelled data and forecast effects for hided test data samples we employ the supervised algorithms. Some techniques, such as feature scaling, extraction and collection, have to begin in the way they are used for training or teaching the model and in the testing or evaluation process, the same functionality has to be retrieved from unseen test data samples. Figure 11 shows a classical map of supervised algorithms. The two stages of model preparation and estimation are emphasized, as can be seen in the figure 11. As previously stated also, for both model-training data and new data samples, the model forecasts performance, similar phases of data analysis, feature scaling, extraction, selection and dimensional reductions are used. This is a key point to consider when constructing every supervised model. The model also incorporates a controlled machine learning process, training data and associated labels, as seen in the figure. During the prediction process (testing), the generated model takes new data samples features and yield predicted labels (Sarkar, Bali, and Sharma, 2018).

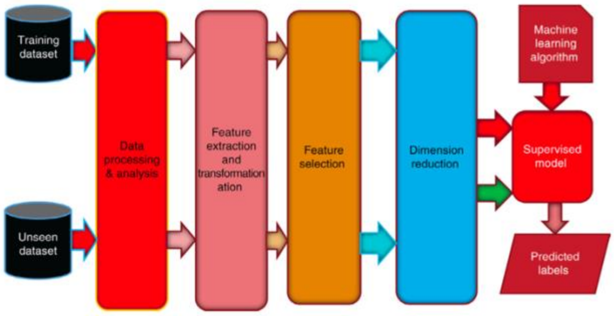

Figure 11. Supervised Machine Learning Framework.

2.5.4 Classification

Classification-based functions are a subsection of supervised machine learning in which the primary purpose of predicting responses or outputs in relation to what the model learned in the training process is a categorical input data. Any output reply therefore belongs to a particular discrete class or group. A broad variety of different problem areas can be overcome by machine learning, and the common classification algorithms are logistic regression, linear discrimination, the artificial neural network, the supported vector machines, the k-nearest neighbors, the naive bayes and the decision tree. A correct classifier requires to be trained from the training data during a classification task. Classification is just one set of missions that a model should learn through. Regression and clustering are the other types (Flach, 2012).

2.5.4.1 Logistic Regression

While it is called, it is not a regression, but a classification model. Logistic regression is an easy, more powerful way of dealing with binary and linear problems. It is a simple to carry out classification model that produces very good results with linear groups. It is a commonly used algorithm for business classification. The logistic regression model such as Adaline and Perceptron is a binary classification statistical approach that can be applied to a multi-class classification. Scikit-learn has an excellent version of the implementation of logistical regression which supports multi-level classification tasks (Raschka, 2015).

2.5.4.2 Artificial Neural Networks

An artificial neural network (ANN) is a concept inspired by the brain and workings of the human mind. The nodes are like neurons in our brain and their interconnections. The ANN and the human brain are somewhat distinct. The brain operates parallel to several neurons, while the computer has only a small number of processors. In contrast to the digital processors, neurons are simpler and slower in rate. Another distinction between the brain and computer structures is the computing capacity on a broader scale. Neurons are made up of parallel operating networks called synapses. The operating system processor is active when the system's memory is passive. In the brain, though, the memory and the retrieval unit are separated over the neurons, and the memory is positioned in the synapses (Alpaydin, 2020). A regular ANN has a layer of input, an output layer, and a minimum layer of one hidden layer between input and output as seen in Figure 12. ANN has often a variety of node layers, specified link patterns and layer ties, relation weights and triggering node (neuron) functions that map outputs. The weights are adjusted during the preparation period. There is a similar arrangement in a deep neural Network (DNN), but there are two or three" hidden layers" of the inputs of neurons

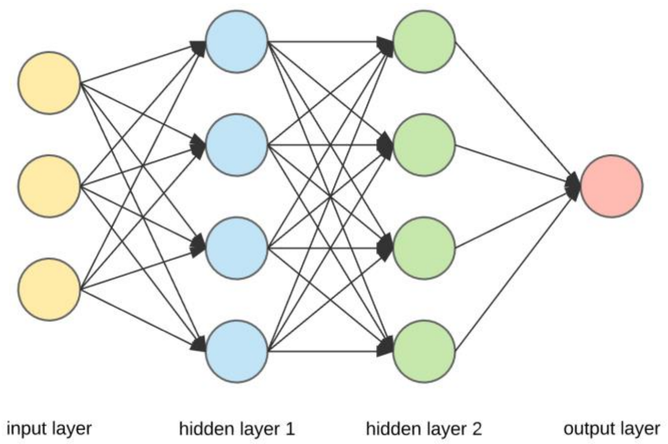

Figure 12. Artificial Neural Networks (ANN) model (Source: Subramanian, 2018)

2.5.4.3 Support Vector Machines

Support Vector Machines (SVMs), one of the most reliable and robust machine-learning algorithms. SVMs looks to find the most effective classification feature for class labels in the training data in the course of the two-class learning process. In the case of the linearly separable dataset, a comparison of the dividing hyperplanes passes through the middle and distinguishes all groups is the function of the linear feature of classification. Due to the presence of a range of linear hyper-aircraft, the SVM's function is expanded to ensure the most appropriate margin, is used by increasing the margin to the limit of the groups. The concept of margin is instinctively the class-to-class space. This margin is mathematically the shortest space possible between the hyperplane point and the closest data points. The motive behind the investigation of SVMs is to decide the most severe margin of hyperplanes which make best use of generalization. It makes the highest results in classification of training data and classifies potential data completely (Wu et al., 2008).

2.5.5 Performance Evaluation Parameters

A significant aspect of its architecture are standards for the assessment of a methodology's performance. Different types of performance assessment measurements occur in the pattern recognition field. The stability of a method's efficiency is measured based on conventional requirements for bio-medical signal analysis in this project. These include accuracy of classification, sensitivity (or true positive rate TPR) or recall, False alarm rate (FAR) and receiving operating characteristic (ROC) curve. We often use the k-fold cross validation approach to assess the efficiency of the existing methods to minimize overfitting. The definitions of these performance parameters (Siuly, Li, and Wen, 2011) (Guo et al., 2009) (Siuly, and Li, 2015) are provided below:

- **Classification accuracy:** Total of acceptable assessments divided by total number of cases
- **Sensitivity:** The estimate of real positive decisions divided by the actual set of successful cases
- **Specificity:** The total of real negative judgments divided by the actual number of negatives
- **FAR:** the proportion of false positive by negative class expected
- **ROC:** A useful tool to display, organize and pick a performance dependent classifier (Fawcett, 2006). The ROC curve shows the sensitivity on the X axis (true positive rate) and the 1-specificity on the Y axis, respectively. The region of the ROC curve is an essential factor for measuring the output of a binary classifier and

the value is always from 0 to1. If the ROC curve region is 1, it implies a perfect descriptive capacity for the classifier. If the area is equal to 0.5, the classifier has no distinguishing force and there can be no appropriate classifier less than 0.5 (Fawcett, 2006).

- **k-fold cross-validation procedure:** Cross-validation is an analysis outcome research validation method. It is commonly used in predictive environments, and we want to estimate how correctly a predictive model is used in practice. The cross-validation method includes the separation of multiple subsets, the analysis on one subset (called the training set) and the analysis on the other subset are validated (called the validation set or testing set). Multiple cross validation rounds are conducted with separate components and the validation to minimize uncertainty. Over the rounds, the results are summed. The data set is partitioned into k mutually excluding subsets roughly of the same size and repeated k times in the k-fold cross validation step (folds) (Sengür, 2009) (Ryali et al., 2010) (Siuly, and Li, 2012). Every time, one of the subassemblies is used as a test set and the other k 1 sub - assemblies are assembled to create a training set. The average accuracy is then estimated over all k samples. Figure 13 demonstrates a design of how the extracted vectors of this analysis are distributed by a k-cross-validation-system, mutually dependent subgroups. The feature vector set is separated into k sub-sets and the method repeats k times as seen in figure 13 (folds). As seen in the figure, one subset is used every time as a test set, the other nine subsets as a training set. Every k-time outcome on the test set is averaged over the iterations k-fold cross validation efficiency. A substantial explanation for cross validation rather than conventional validation is the fact that there are not enough data for partitioning into different training and testing sets without compromising considerable modeling or testing capabilities. (e.g., devising the data into two sets of 70% for training and 30% for testing).

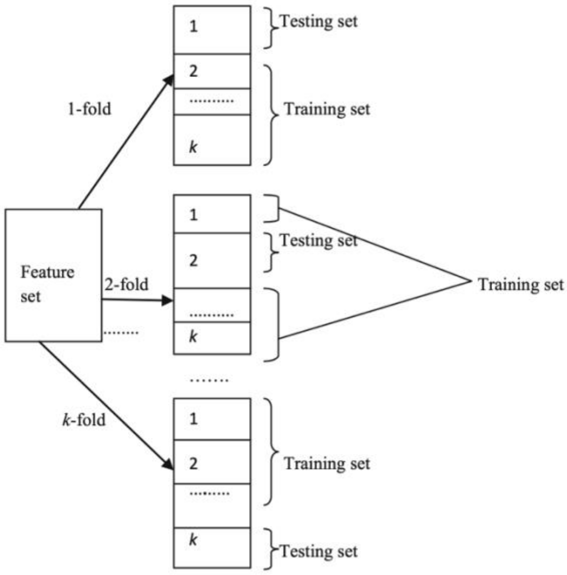

Figure 13. K-Fold Cross-Validation Design

2.6 Related Works

In (Huang et al., 2013), 2 EEG (FP1 and Fp2) channels were divided into EEG signals in near stationary elements, features extraction, reduction based Short Time Fast Fourier (STFT) and Fuzzy C-means (FCM)respectively. Multiclass SVM was used to build an ASSC system that matched an accuracy of 70.92 %. The study by (Radha et al., 2014) employed 6 major EEG signals and various signal processing characteristics including time domain, frequency domain and non-linear capabilities were employed. In addition, Random Forests and SVM were listed for five stages of sleep. The findings revealed the optimum efficiency using the frontal EEG signals, with spectral linear characteristics and an RF, which was better than the SVM. The precision data from 6, 5, 4 stages, 3 and 2 stages as 81, 57 percent, 86,53 percent, 87,49 percent and 95,05, respectively, were recorded by (Hassan, Bashar, and Hassan Bhuiyan, 2015), using the Bootstrap Aggregation Algorithm with a number of statistics and specimen features extracted from a single EEG channel. On the other hand, (Rodríguez-Sotelo et al., 2014) Generated entropy metrics, Q-algorithm for dimensionality reduction, and J-mean clustering for 2 EEG-channel automatic stage sleep score. The efficiency of automated data achieves maximum classification accuracy of up to 80%.According to (Lan et al., 2015), spectral functionality was derived on the basis of Fast Fourier Transform (FFT) of PSG data for the classification of sleep stages with a DT classifier based on multichannel and achieved a precision of 84%.Moreover, (Zhu, Li, and Wen, 2014) applied classification for the sleep stage based on extraction, using a single EEG channel and a multiclass SVM as classifier, nine graph field functions in the Visibility Map (VM) and Horizontal VG (HVG).

Table 2. Sleep Studies EEG Feature Extraction Techniques-Based Signal Processing

Technique	Features	sleep stage classification
Time Domain	• standard statistics • Integrated EEG • Tsallis entropy	• (Aboalayon, and Faezipour, 2014), (Aboalayon et al., 2016), (Khalighi et al., 2011) • (Correa, and Leber, 2010), (Correa, Orosco, and Laciar, 2014), (Kumari B. M, 2014) • (Khalighi et al., 2013), (Khalighi et al., 2011)
Frequency Domain	• Non-parametric Analysis • coherence analysis • median frequency	• (Khalighi et al., 2013), (Liang et al., 2012), (Chen, Wang, and Wang, 2013) • (Krakovská, and Mezeiová, 2011) • (Gudmundsson, Runarsson, and Sigurdsson, 2005)
Time-Frequency Domain	• WT • STFT • Choi Williams	• (Khalighi et al., 2013), (Zoubek et al., 2008) • (Sanders, McCurry, and Clements, 2014) • (Fraiwan et al., 2012)

The specificity of the six-stage classification was 87.5% using the SVM Classification. Hsu et al (Hsu et al., 2013) categorized 5 sleep phases with 87.2% precision based on six energy features from the same EEG channel implementing Elman recurrent neural classifier (ERNC). (Čić, Šoda, and Bonković, 2013) approached with a total accuracy of 90% by SVM classification and the use of a single EEG channel-based time-frequency features created by the EMD system and using the Generalized ZERO Crossing Method (GZC) on obtained intrinsic mode functions (IMF). Three features, the average power method, the preferential frequency band approach and an EEG uni-channel were fed into a sleep stage classification with (Sanders, McCurry, and Clements, 2014) as inputs to the LDA classification method. Using the combination of the average power and CFC characteristics, the proposed method graded correctly up to an average of 75%, which outperformed each solution individually. Table 2 listed different EEG-based signal processing techniques used in each ASSC stage. Koch El in (Koch et al., 2014) tackled the issue of classification of the sleep stage by the use of a multi-PSG signals based Latent Dirichlet model that included 2 EEG and 2 EOG channels. The average accuracy of the model was 68.3%.

CHAPTER 3: RESEARCH METHODS

Generally, health care's data is sensitive compared to others, hence as shown in figure 14 input collection and segmentation, processing of data, feature selection or dimension reduction and classification can be the focus of a four-phase classification process. The biomedical data are typically recorded and pre-processed from the human body.

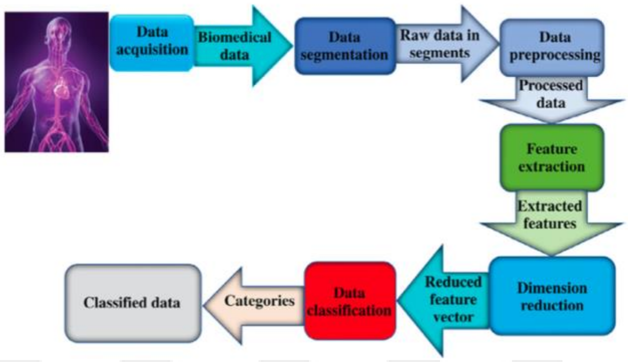

Figure 14. Biomedical Data Classification Substructure (Source: Subasi, 2019)

The input may contain noise beside missing parts or labels which should be eliminated or investigated, the noise was and is still a nightmare for the data pre-processing, the artefacts might take us far away from our main goal, it can lead to low accuracy and distorted results. Therefore, the pre-processing step is crucial to delete cropped epochs to amplify informative details of EEG raw signals. Some multi-channel-based techniques for automated sleep stages scoring are investigated here. The features are then examined and converted from biomedical data into a functional vector. A dimension reduction is used in the next stage to avoid the data redundancy from the functional group, building a minimized functional. During the latest phase, the reduced vector is identified by a classifier (Subasi, 2019). In the following section the database and algorithms used to attend the goal of this project will be investigated. The following gives a short overview of the methods for pre-processing data, extraction of the features and classification procedures. A new concept is created with EEG signal to identify sleep phases in various groups, in the following sections As shown in figure 15, our procedure commence with generating a new channel order where we use only 2 EEG channels instead of multi different channels, then we create EEG epochs based on different events found in annotation files, the evoked response from each group of epochs will be used as input for our classification model after extracting the valuable features using the wavelet transform and statistical methods, several machine learning classifiers will be examined in the next lines.

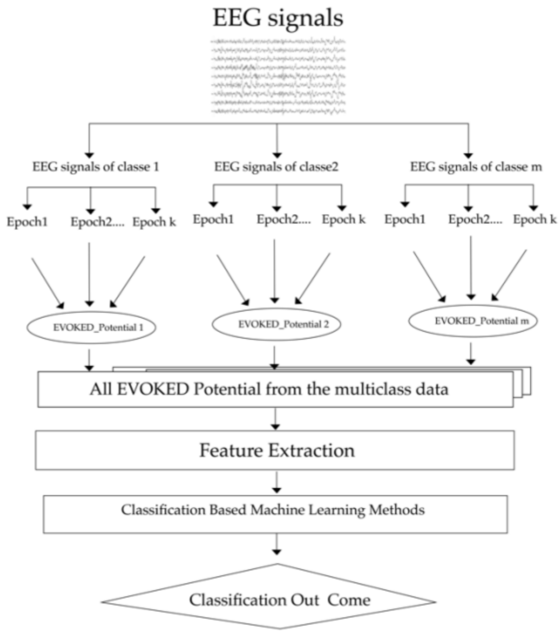

Figure 15. Proposed Automatic Sleep Stage Classification Procedure.

3.1 Dataset

For this dissertation, the dataset was collected from the online data archive of Physionet (Goldberger et al., 2000). Physionet provides is an open door to a range of physiological signals and signal processing's open-source applications. There are various PSG (Polysomnographic) collections, but Sleep-EDF Expanded Database were used for this study (Kemp et al., 2000). The Sleep-EDF database includes 197 polysomnographic full-night sleep records, including EEG, EOG, EMG and event indications. Such reports also include breathing and body temperature. According to the Rechtchaffen and Kales manual the associated sleep cycles have been marked manually by well qualified technicians and are also accessible.

3.1.1 Data and Annotation Files

PSG (polysomnography) files include full-night polysomnography of sleep EEG registrations (from Fpz- Cz and from Pz-Oz electrodes), EOG files (horizontal and submental chin EMG and an incident symbol). The dossier containing sleep-cassette PSG.edf files also often include oronasal respiration and rectal body temperature. Hypnograms include sleep stages labeled W, R, 1, 2, 3, 4, M (Movement time) and ? (not scored). The qualified technicians have manually labelled all hyperlinks (recognized by the eighth letter of the filename) in compliance with the 1968 Handbook of Rechtchaffen and Kales. But based on Fpz-Cz/Pz-Oz EEG instead of C4-A1/C3-A2 EEG.

3.1.2 Sleep Cassette Study and Data

A 1987-1991 research on the impact of age in sleep on stable Caucasians between 25 and 101 years of age, without sleep drugs collected 153 Sleep Cassette files (Mourtazaev et al., 1995). For two posterior days-night cycles at the subject homes two PSGs of about 20 hours were reported each. Though subjects continue their usual work, a changed Walkman-like tape recorder was worn as it is mentioned in Bob's 1987 thesis chapter VI.4 (Kemp, 1987). Sampling frequency for both EOG and EEG were fixed at 100 Hz. Electro rectified and high pass filtered, the submental EMG signal, which was later low pass filtered after sampling at 1Hz the resulting EMG cover expressed in μV rms (root-mean-square). The airflow, the temperature of the rectal body and the incident predictor were all sampled at same sampling frequency (1Hz).
we need to mention that there are two-night recordings per subject except for subjects 13, 36 and 52 which have one file missing each owing to missing recording hardware.

3.2 Data Pre-processing

After the 67 data subjects have been downloaded, the aim is to acquire epochs and its associated ground real, in order to do this, we use the MNE, open-source Python software, to explore, simulate and evaluate the neurophysiological data of humans such as MEG, EEG, sEEG, ECoG and more. We use the fetcher to load the data and extract for each a pair of files for all individuals; PSG.edf: polysomnography included. The raw data that is compatible with EEG cask data. Hypnogram.edf with the expert's annotations. Now that we have our input recording, we combine the two files (PSG & Hypnogram) in an object of Raw then Events can be derived from annotations including descriptions to meet the epochs (Figure 16). An annotation is defined by an onset, a duration, and a string description. It can contain information about the sleep, but it can contain details on signals marked by a human: bad data segments, sleep scores, sleep events (spindles, K-complex) etc. An Annotations object is a container of multiple annotations.

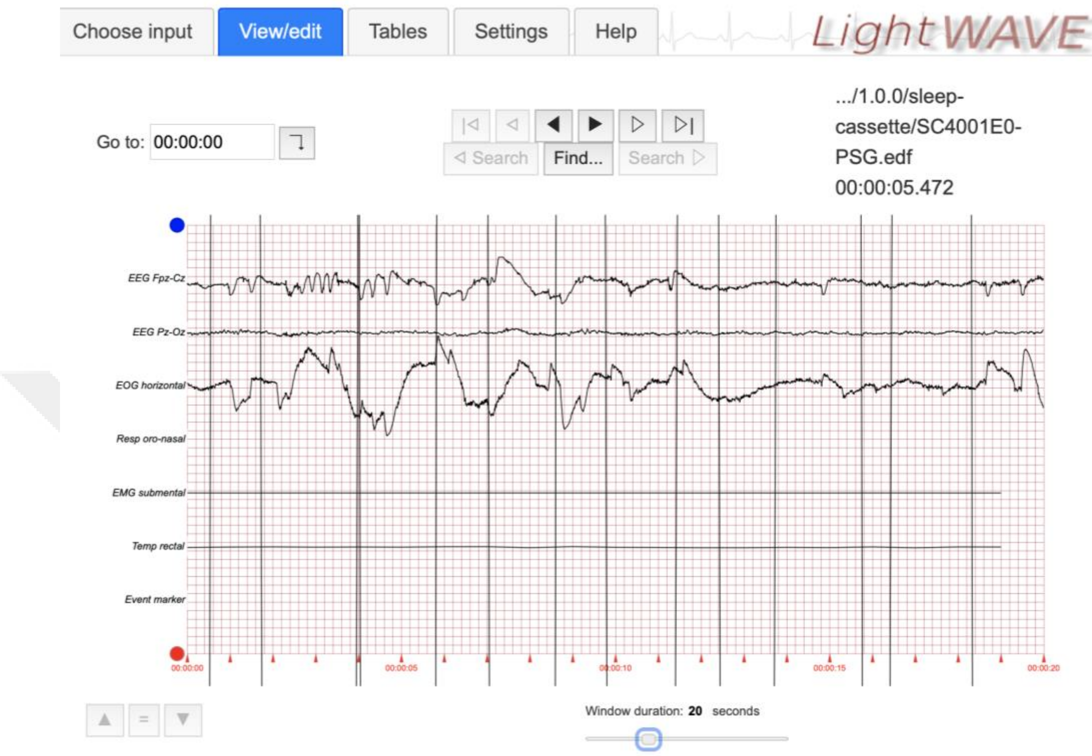

Figure 16. Subject 1 First Night's PSG Recording

We can observe that our raw object contains 7 channels, EEG Fpz-Cz, EEG Pz-Oz, EOG horizontal, Resp oro-nasal, Temp rectal and Event marker.

3.2.1 Staging Criteria

The American Academy of Sleep Medicine's sleep rating norm was used to obtain improved outcomes. This manual is even before RK rules and this standard is adopted by most of the studies conducted in recent years. Previous findings typically transform sleep from RK to AASM simply by inserting slow wave sleep stages S3+S4, which produce stage N3. This transition cannot be treated as completely correct, as the updated regulations have modified the length of each night sleep level. In addition, the new regulations recommend a 500Hz sampling frequency, whereas most data sets (including the data used in this study) contain 100Hz sampling frequency of signals.

In addition, the EEG channels suggested are F4-M1, C4-M1, O2-M1, and F3-M2, C3-M2, and O1- M2 backup channels. In this study, EDF-x EEG signals from Fpz-Cz and Pz-Oz were analyzed. We will establish a new channel order consisting of just two channels that we will use for the remainder of the study.

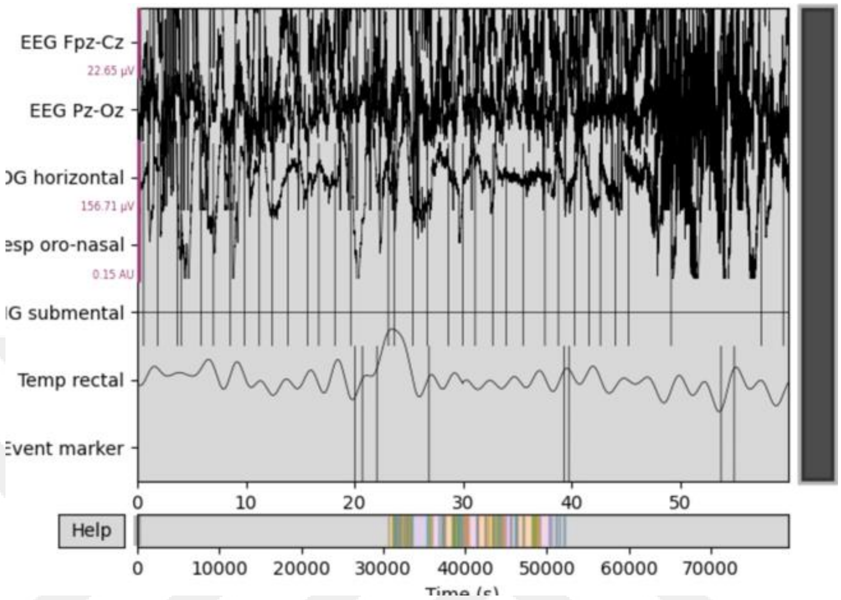

Figure 17. Raw Object Obtained After Combining Hypnogram and PSG for Subject 1

The Sleep Physionet dataset is labeled with eight labels: wake (W), Stage 1, Stage 2, Stage 3, Stage 4, from light to deep sleep, REM sleep (R) where REM is the briefing for Rapid Movement Sleep, Motion (M) and Stage (?) in all segments not scoring. We only operate with 5 classes: Wake (W), Stage 1, Stage 2, Stage 3/4 and Sleep for REM (R). To do so, we pick which events we want and apply an event identifier to each event. In addition, there are long wake regions before and after the night for the recordings. We only slice per recording 30 mins of wake time before the first occurrence and 30 mins after the last occurrence of sleep stages so as to limit the effect of class imbalance. As a result, for this analysis, an AASM manual is used, meaning that samples are classed at 30 seconds or 3000 data points (f=100Hz) using 5 classification stages (W, N1, N2, N3, R).

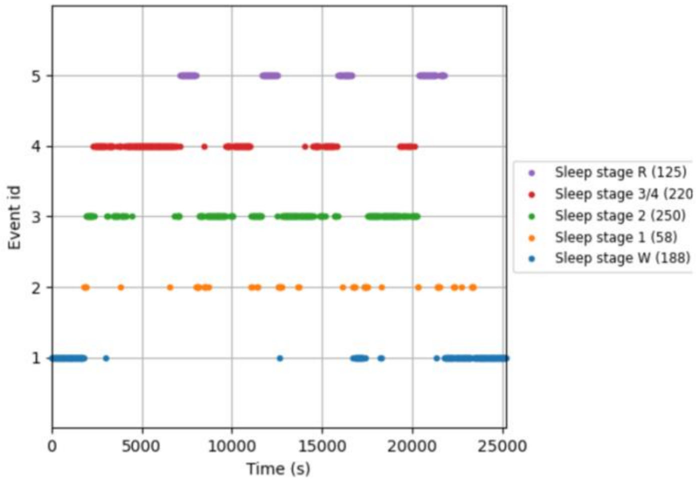

Figure 18. Sleep Stages-Based Event-Id from the First Night (1st Subject)

3.2.2 Epoch Determination

The EEG signals are considered to be non-stationary and aperiodic, and the amplitude of their signals is continually changing. We split the EEG signals of a class into certain subsets, that we name epochs in this study, for representative values of a particular time frame.

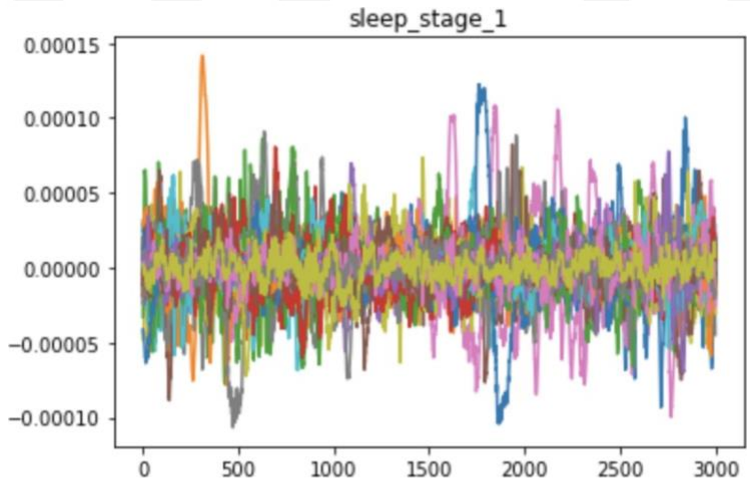

Figure 19. Sleep Stage 1 Epochs Extracted from the Raw EEG from Subject 1

Thus, in a time window, each epoch contains EEG data. It should be remembered that there are many EEG channel data in any epoch. In general, EEG raw input is the columns of an epoch as shown in figure 19, Our EEG data is 3D (number of epochs, number of channels, number of times), after the epoching procedure, where time is a duration in each period, and epochs are the number of components we have derived of the continuous EEG. In order to make the process of visualization simpler, we reshape our signal to become (number of times, (number of channels * number of epochs))).

3.2.3 Evoked Response Potential

A pattern reported by electrodes from a certain part of the nervous system is an evocative response potential. The amount of work done to transfer a charging unit within a field without acceleration is an electric force. These electrical signals are used in disease control and diagnosis. However, ’Event-related potential’ (ERP) may be viewed as stereotyped neurophysiological reactions to a voluntary action or an external factor. ERPs are acquired by averaging the random electroencephalogram (EEG), for example repetitive displays of a similar stimulus, for multiple instances of the same phenomenon. Assuming the event’s start is known, the ERP is a time series that includes the brain reflex, known as a period or epoch, useful for the distinction of the subject’s brain states in response to the multiple stimuli in the sense of the test or study. This makes ERP analysis the optimal, long-term method in the interaction between brain and machine (BCI). Therefore, we use intervals of sleep to estimate evocative answers from various classes of sleep.

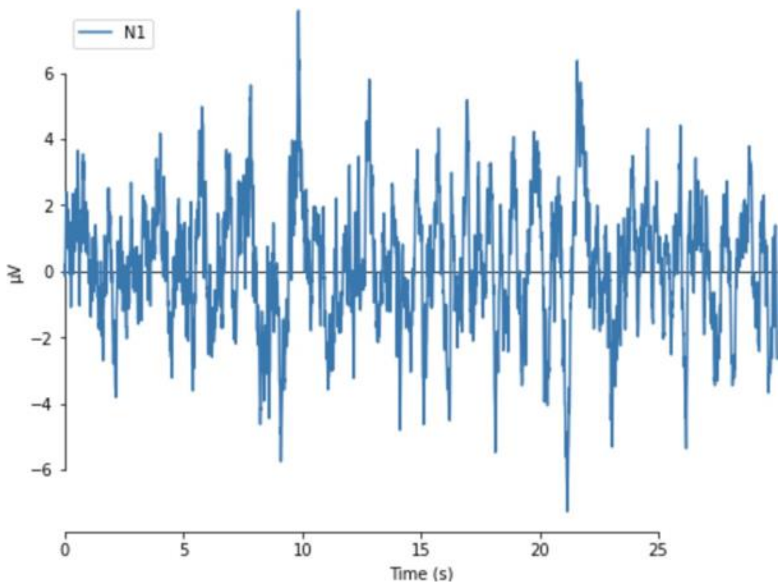

Figure 20. Sleep Stage 1 Evoked Response’s from Subject 1

3.2.4 Feature Engineering

After we dropped the bad epochs, we extracted the evoked response from the epochs obtained from our two channels, The Evoked response input must be translated into significant features which reflect the signal information. Power spectral density is one possible way in which the variations in the electromagnetic processing of the brain are defined in sleep stage scoring. This segment discusses the study of the electroencephalogram (EEG) based power spectrum density (PSD). If we analyze the spectral power track of the epochs clustered by sleeping periods, we can see that there are various sleep signature characters. These signatures continue to be similar among different data subjects (Figure 21).

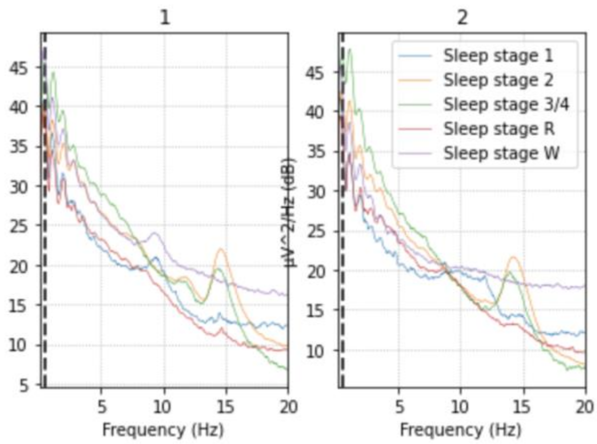

Figure 21. PSD of the Epochs Grouped by Sleep Stages from Subject 1 and 2

3.2.4.1 Feature Extraction

Feature extraction for any form of EEG-related research is considered the most effective protocol. An effective feature extraction technique is essential for an efficient classification to remove insightful and distinctive features from the initial data frame. Generally, if the derived functionalities do not precisely reflect the input investigated and are not useful, it can be difficult to describe the identification of the classes or features using the classification algorithms (Siuly, Li, and Zhang, 2016). In order to construct features that better represent raw data, the essence of the signal should be known. EEG signals are non-stationary, meaning the statistics of the signal are time shifting. This means that the time domain signal analysis is not enough. The use of time domain characteristics, frequency domain features, time-frequency domain features, entropy characteristics and non-linear attributes reveals multiple elements of the EEG signals.

3.2.4.2 Wavelet Transform

Wavelet Transform is a popular method of time-frequency distribution and in the last two decades has been widely used for signal visualization and interpretation in many sectors. It uses both time and frequency features. Thanks to its compact way to describe the time-frequency domain of a signal, wavelet transformation is sufficient for the study of a non-stationary signal. It is thus an effective method to evaluate and derive the EEG signal from its functionality. By moving and scaling wavelets over various frequency ranges, the input signal is breaking down. The vector of coefficients can be accessed and used as an input of the classifier by the use of the multi-resolution. A wavelet family is weighted and shifted such that the EEG signal from the Evoked Response is optimally approximated. Both fine and coarser signal resolution characteristics are recorded by the wavelets. The characteristics of the Evoked Response signal are the wavelet coefficients. The extraction of features using wavelets has been commonly used to examine EEG, opening up the possibility of a geographical, spectral and time highly detailed explanation.

3.2.4.3 Wavelet Families

The Wavelet Transform consists of several distinct wavelet families. The wavelet families vary because the compact and smooth look of the wavelet has been made for each family. Which implies that each form of wavelet has a different shape, compactness and usability, which suits the functions that our signal seeks best, and it is useful for a different purpose. For this analysis, we examine various discrete wavelet families using 2 distinct EEG channels (Daubechies, Coiflets, Biorthogonal, Symlets).

3.2.4.4 Discrete Wavelet Transform

The Discrete Wavelet Transform still acts as a filter bank in operation. It is thus used as a case of high-pass and low-pass filters. An explanation for this, the filter banks are a highly useful way to break a signal in many frequency sub-bands. The study can often produce the same characteristics correlated with the specific sleep phases during each Evoked response phase of the loop. For each level of sleep, sleep EEG signals have been analyzed using discrete wavelet transformation in this project (DWT). DWT is often used to dissolve the Evoked Response EEG waves into a comprehensive coefficient. We selected 6 stages of decomposition for this purpose beside a mathematical approach of statistical values related to the coefficients of each decomposition level. The features extracted from this study almost correspond to the desired features (Figure 22).

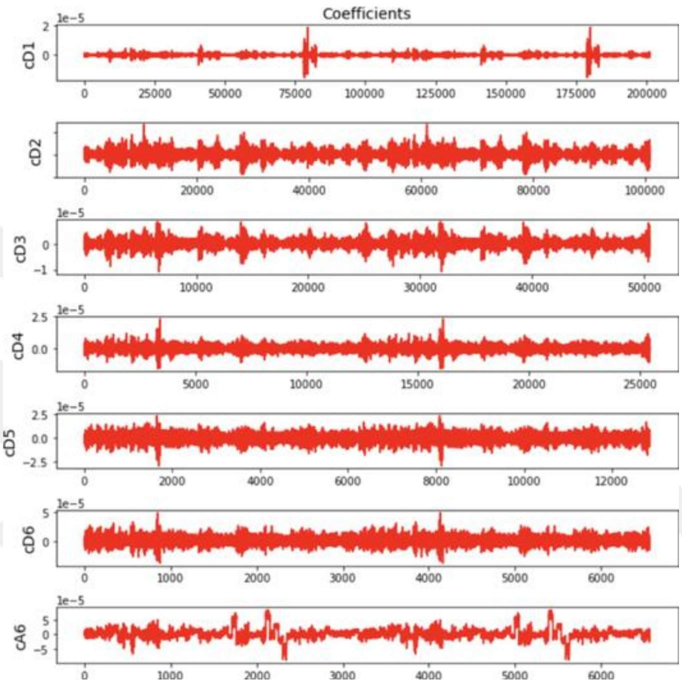

Figure 22. Coefficients Extracted from the DWT for Evoked Response 1 from First Subject.

3.2.4.5 Statistical Values

Now that we extracted valuable features from the EEG evoked response signals by employing discrete wavelet transform (DWT), it's time to compute statistical parameters of Discrete wavelet sub-bands. Several statistical values (features), which are the significant indication values for the distribution of the evoked response data, can be derived from each sampled data point. The statistical components derived from every coefficient of evoked potential signals are presented below. A number of data points of each evoked response were determined from various stages of sleep for each feature:

Mean: This statistical element has been used for EEG signals (2 channels). It's determined the average of an evoked potential. The data points of an epoch (electrical potential) were summed and consequently divided by the total number of data points (N), while Xi present every single data point.

$$\mu = \frac{\sum_{i=1}^{i=N} Xi}{N} \tag{8}$$

Skewness: The asymmetry frequency of a grouping measurement. The skewing of the normal grouping is zero, when the negative and positive skewing show the data is biased, respectively, left and right, where $\bar{Y}$ presents the average, v the standard deviation while the summed data points present N.

$$skewness = \frac{\sum_{i=1}^{i=N} (Yi - \bar{Y})^3}{N * s^3} \tag{9}$$

Kurtosis: That is the maximum of a frequency distribution relative to the normal distribution in the graph with regard to the concentration of variables adjacent to the mean. Where $\bar{Y}$ presents the average, v the standard deviation while the summed data points are N. Kurtosis is considered also a higher-order-statistics measure (fourth moment).

$$kurtosis = \frac{\sum_{i=1}^{i=N} (Yi - \bar{Y})^4}{N * s^4} - 3 \tag{10}$$

Beside calculating the median, Ratio, RMS and standard deviation from frequency sub-bands. Heretofore, we examinate three stages from our model, we defined the epochs from 67 subject recordings using 2 channels (EEG Fpz-Cz, EEG Pz-Oz), evoked response potential from the epochs were obtained describing the 5 sleep stages, discrete wavelet transform with different wavelets families takes lead to extract important features from our EEG evoked responses, 48 statistical values per evoked potential were extracted.

Table 3. Number of Statistical Values Extracted from DWT Sub Bands

Statistical values	Mean	std deviation	skewness	kurtosis	median	ratio	RMS
Number of values	7	7	7	7	7	7	6

The shape of total extracted features using the first channel or second separated will be:

[(Number of classes$*$ Number of signals), number of features] = [335, 48]

Using 2 EEG channels Combined:

[2 $*$ (Number of classes $*$ Number of signals), number of features] = [670, 48]

In order to organize our data via the corresponding labels, we associate with every feature extracted the appropriate sleep stage using the 5 labels (N1, N2, N3, REM, WAKE).

3.2.4.6 Features Normalization

Normalization using min-max scaling is a commonly used approach which is acceptable when the values are expected in a restricted length. For some machine learning techniques standardization could be more realistic as diverse linearly models allocate weight to 0 or small random parameters to around 0. The function columns are focused at mean 0 by standardization with standard deviation 1, in such a manner that the columns establish a regular distribution. Standardization also preserves useful knowledge on the outliers and makes it less vulnerable than the min-max scaling which scales the data into a small range of values. Standardization is also introduced in scikit-learn as a standard scaler category, like min-max scaler. The regular scaler must also be illuminated as it fits on the training data and these parameters can be used for converting the test set or other fresh data points. The features derived from evoked responses are of different ranges, and the outcomes of these measures may be influenced by this variation. This bias has been avoided by using the standardization function approach. We used normalization in this thesis. The classification of features has been measured and affected by this theory. The features have been revamped to have a zero mean and unit deviation in standardization. For several machine learning techniques, this rescaling is essential. It is necessary to refer that we used Label encoding, it involves converting each value in the columns to a number. Consider our label columns in the extracted features data frame, each label is assigned a unique integer based on alphabetical ordering.

3.3 Classification

Classification is called the method of marking the data in relevant groups. The initial phase in the classification plan is to define the traits or attributes that highlight or make for the greater discrimination between the various categories. A classification model is enhanced to afford a basis for applying the behavior of the classification processes. The optimal model of the classification system should be used, even though it can be updated as the construction of the classifier progresses. It is introduced and" trained" to detect the selected data components or to define the appropriate mapping input-to-output. After training and feeding, the model is able to identify particular inputs. The program can then be checked and analyzed using measurements such as machine speed and classification accuracy (Hastie, Tibshirani, and Friedman, 2009). For sleep stage classification variety of the classifiers have been used. In this thesis we will compare the recently emerged methods. Five different classifiers were chosen as follows: Linear discriminant Analysis (LDA), Multi-layer perceptron (MLP), Support vector machine (SVM), Random Forest (RF), Gradient Boost (GB) and Bagging.

3.3.1 Linear Discriminant Analysis

LDA is a classification algorithm employed to locate a linear feature composition that distinguishes two or more data groups. As a linear grouping, the following mixture can be used. The groups can usually be normally allocated in LDA. It can be used for either reduction or classification of dimensions like PCA. In our case, which we deal with five categories, the priory probabilities for category 1 (N1), category 2 (N2), category 3 (N3), category 4 (Wake), and category 5 (REM) areρ1, ρ2, ρ3, ρ4, ρ5, the class means, and overall mean are 1, 2, 3, 4, 5 and the class variances are cov1, cov2, cov3, cov4, cov5 respectively.

$$\mu = \rho 1 * \mu 1 + \rho 2 * \mu 2 + \rho 3 * \mu 3 + \rho 4 * \mu 4 + \rho 5 * \mu 5 \tag{11}$$

Then scatters inside and within classes are used to reflect the appropriate class dissimilarity criteria. The scatter measurements are determined as follows:

$$S_{\omega} = \sum_{i=1}^{N} \rho i * x * cov(i) \tag{11}$$

where N goes to the sum of sleep stages. In our practical case, class covariances and average are not defined, but from the training set they can be estimated.

3.3.2 Multi-layer Perceptron

A regular ANN seems to have an input layer, an output layer and at least one hidden layer between input and output. ANN often has a number of node levels, specified connection patterns and layer ties, linking weights and node activation functions that map output to weighted inputs. The weights are adjusted during the training. The back-propagation technique is an ANN-training procedure and has two principal phases; propagation and weight update.

3.3.2.1 Propagation

1. The input data sample vectors distributed through the neural network to produce from the output layer the output values.
2. Evaluate the generated output group for that input data vector to the real (wanted) output vector.
3. The output unit error is measured.
4. Back-propagate error for any link (neuron).

3.3.2.2 Update

1. Measure the gradient by multiplying the input activation and output errors.
2. Using the rate of learning in the first weight to calculate by subtraction the proportion of the gradient and then change the weight of the node. By using different epochs, these two steps are replicated several times before accurate findings are obtained. Back propagation is typically used along with algorithms for optimization such as stochastic gradient descent. Multi-layer sensor (MLP) is an

artificial neural totally connected feed forward network of three layers or more (input, output, and minimum one hidden layer). Back-propagation should be used to train the model of MLPs (Sarkar, Bali, and Sharma, 2018).

3.3.3 Support Vector Machine (SVM)

SVM does not allow multi-class classification by nature in its simplest form. It facilitates differential grouping and division into two groups of data points. The same theory is used to map input points into high dimensional surface for reciprocal linear differentiation between each two groups after splitting the multi classification problem into several binary classification problems. The one-to-one strategy splits up the dilemma of multiple classifications into multiple concerns. Another method we may use is One- to-Rest. A binary classifier is per class couple. The classification is set as a binary classification of each class. A single SVM is binary and can discriminate between two classes. In order to distinguish data points from 5 groups of our dataset according to the two overview approaches; the classifier can use 5 SVMs in the One-to-Rest approach. In one of the five phases of sleep each SVM will estimate membership. In the One-to-One approach, it will use 10 SVMS.

3.3.4 Random Forest

The Random Forest exercises a decision tree as the key classifier, which is to define the knowledge and apply this ensemble learning strategy. The ensemble methodology blends many qualified classifiers with a view to classifying new instances. A random forest is a classification category made up of tree-organized classifier. The individual random variables are indistinguishable in such classifiers. In comparison, for the best-known class each tree makes the unit vote. A random vector is independent of the previous random vectors and is used for the development of a tree. For random forests, the upper limit for the acquisition of generalization errors will be calculated to the degree of two fundamental parameters: the precision of the different classifications and their interrelationship. In the case of random forest, there are two parts for error generalization. These sections are defined as the power of the individual classifiers in the forest and their association according to the raw margin. The correlation must be minimized to improve the consistency of the random forest to ensure the power stays unchanged. (Goel, and Abhilasha, 2017).

3.3.5 Gradient Boost (GB)

With their “bagging” procedure, Breiman (Breiman, 1996) developed the concept of randomness in order to increase its reliability in the function estimation process. Original AdaBoost application often employed random sampling, although this was assumed a deterministic weighting approximation as long as the application of the simple learner does not endorse observational weights but rather as an integral feature. Breiman subsequently proposed a hybrid bagging enhancing technique (“adaptive bagging”) for additive expansion suitable to the least squares. The simple learner is replaced by the required bagged base learner in normal boosting operations, while at each boost, the “out of bag” residues are supplied with the ordinary residue. Driven by Breiman, the gradient boosting change has been limited, so randomness as an integral part of the method is combined. A sub-sample of the training data is taken from the whole training dataset at each run randomly (without replacement). This selected sub-sample is then used to match the simple learner and determine the update model for the real iteration instead of the entire sample (Friedman, 2002).

3.3.6 Bagging

Bagging identified as a set learning classifier trained with randomly chosen inputs from a training data frame that improves design variation. In the event of a grouping or predictive forecast, the moderate solution is to adopt a weighted choice or a weighted mean. This enables one to distinguish test events in order to generate accurate predictions for a certain machine learning model. The bagging often covers the initial training results, instead of taking individual samples from the domains. Bagging is special because the initial training data is re-sampled instead of using independent domain datasets. Parallel to the final model with equal weights, separate models are composed with many samples. The classifier still succeeds from the initial training data with much better predictions than the real classifier and never has substantially bad results. This is intended to neutralize the gap by modifying the initial preparation details by canceling and duplicating those cases. The classifier promotes accuracy and eliminates variation and prejudices by preventing overfitting. However, the partial model, robust in the variations in samples training results, does not greatly improve (Witten, Frank, and Hall, 2011)

CHAPTER 4: RESULTS

In its expanded Sleep-EDF database, the proposed Automatic Sleep Stage scheme (ASSC) is standing on the two independent, EEG double-raw signal. The EEG signals have been filtered and segmented into 30-seconds epochs, indicating that the evoked responses mentioned in the processing step are obtained as averages. New feature sets were drawn from the EEG evoked responses of 67 EEG recordings using the discrete wavelet transform, to assess the efficiency of our proposed ASSC process. Our analysis is based on the criteria available to define the sleep cycle, which include 5 stages: wake, Stage 1, Stage 2, Stage 3 and REM. For classification tasks, various types of wavelets were used to train and analyze inputs at various proportion ranges alongside multiple machine learning algorithms, namely: 50% and 50%; 70% and 30%; and 80% and 20%. For each percentage stage, classifiers have been randomized and tested, and the best accuracy was recorded. We also evaluated the efficiency of various classifiers, as a certain classifier will produce better results than other classifiers in ASSC systems. According to (Hassan, Bashar, and Hassan Bhuiyan, 2015), A perfect Machine learning classifier does not exist, since a set of machine learning models which work very well in one field can work terribly in another. Therefore, we have done experiments using various types of classifiers to ensure that the logic of algorithm and it's practical in outside of this paper and have decided that it is best for the most diverse results. Moreover, on the basis of our survey, a range of ASSC programs have taken into account tests and enhanced results.

		PREDICTED CLASS	
ACTUAL CLASS		Class=Yes	Class=No
	Class=Yes	(TP)	(FN)
	Class=No	(FP)	(TN)

Figure 23. Confusion Matrix (Source: Subasi, 2020)

Our research involves SVM, LDA, MLP, Bagging, Random Tree, and Gradient Boost, amongst the most widely used classifications in the ASSC on the basis of our study. Calculation of the accuracy, sensitivity, and characteristics with TP, FP, FN and TN values will determine the output of these classifiers, where TP corresponds to true negatives, FP is false positive, and FN is false negatives. The following are the equations of precision, adaptation and specificity:

Firstly, the recall indicates the ratio of classes correctly identified by classes not correctly sorted:

$$Recall = \frac{TP}{TP + FN} \tag{12}$$

The ratio of proper classification of all relevant instances can be expressed by the formula to determine the precision of the model:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

The model's precision is the percentage of instances accurately identified positively over other positive instances:

$$Precision = \frac{TP}{TP + FP} \tag{14}$$

In the following tables we describe the benefits of the suggested approach and importance of the experimental results for each algorithm or method applied. These tables highlight each sleep stage's comparison, sensitivity and specificity with the general accuracy of all six classifications analyzed: SVM, LDA, MLP, Bagging, RF and Gradient Boost.

4.1 EEG Fpz Cz

Table 4. LDA Performance from First Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	52	50	85	72	36	73.5
	Precision	55	62	78	62	44	
30	F1	56	63	68	72	45	74.35
	Precision	53	62	81	68	45	
50	F1	43	56	67	49	39	83.23
	Precision	50	47	79	45	42	

Table 5. MLP Performance from First Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	40	36	86	63	55	100
	Precision	50	40	75	55	67	
30	F1	59	47	88	71	53	99.57
	Precision	59	50	90	75	48	
50	F1	47	48	82	54	31	97
	Precision	56	40	80	47	43	

Table 6. Bagging Performance from First Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	50	40	90	67	40	98.1
	Precision	44	50	88	67	42	
30	F1	59	63	84	65	33	97.43
	Precision	55	52	100	72	38	
50	F1	45	54	81	63	36	98.2
	Precision	45	50	78	62	41	

Table 7. RF Performance from First Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	45	50	91	65	43	100
	Precision	50	62	83	62	40	
30	F1	50	68	85	58	40	100
	Precision	53	59	94	54	47	
50	F1	43	63	88	64	51	100
	Precision	45	65	86	60	53	

Table 8. GB Performance from First Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	38	45	88	58	38	100
	Precision	44	50	79	56	38	
30	F1	51	46	83	68	46	100
	Precision	45	47	89	68	47	
50	F1	52	54	80	66	54	100
	Precision	57	51	89	57	56	

Table 9. SVM Performance from First Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	65	73	93	67	61	99.25
	Precision	63	73	93	73	58	
30	F1	60	56	75	39	37	81.19
	Precision	75	45	66	69	39	
50	F1	46	53	80	70	27	91.61
	Precision	36	50	76	74	47	

4.2 EEG Pz-Oz

Table 10. LDA Performance from Second Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	25	48	72	77	62	73.50
	Precision	33	61	72	76	32	
30	F1	47	48	84	65	32	78.63
	Precision	59	36	100	63	32	
50	F1	31	43	69	66	54	85.62
	Precision	30	35	80	70	59	

Table 11. MLP Performance from Second Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	43	62	74	76	48	99.25
	Precision	42	57	71	88	45	
30	F1	42	41	89	59	59	97.43
	Precision	53	35	95	52	58	
50	F1	30	50	82	59	54	97.60
	Precision	29	42	82	63	67	

Table 12. Bagging Performance from Second Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	32	62	77	70	30	99.62
	Precision	29	57	77	81	30	
30	F1	43	53	81	65	54	98.71
	Precision	56	45	78	63	53	
50	F1	43	66	78	66	51	98.2
	Precision	35	60	76	79	70	

Table 13. RF Performance from Second Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	50	65	91	71	86	100
	Precision	50	62	91	83	75	
30	F1	35	69	87	76	40	100
	Precision	33	64	90	88	33	
50	F1	45	63	81	70	51	100
	Precision	41	59	84	76	54	

Table 14. GB Performance from Second Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	42	62	74	76	60	98.13
	Precision	38	57	71	88	60	
30	F1	45	52	84	70	43	100
	Precision	56	47	84	64	42	
50	F1	42	55	82	71	43	100
	Precision	36	55	82	79	47	

Table 15. SVM Performance from Second Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	75	80	100	80	67	84.85
	Precision	75	75	100	86	67	
30	F1	60	51	69	76	48	91.45
	Precision	57	53	67	79	50	
50	F1	43	64	86	67	50	100
	Precision	46	63	83	66	50	

4.3 EEG Fpz-Cz & EEG Pz-Oz

Table 16. LDA Performance from Combined Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	25	48	72	77	62	74.3
	Precision	30	40	75	67	80	
30	F1	38	59	74	55	35	75.64
	Precision	35	52	100	50	40	
50	F1	35	48	63	60	36	82.63
	Precision	31	45	69	65	38	

Table 17. MLP Performance from Combined Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
48	F1	54	59	77	83	50	94.4
	Precision	58	48	77	77	75	
30	F1	33	58	86	60	44	97.86
	Precision	33	56	100	50	54	
50	F1	38	48	77	58	40	100
	Precision	33	48	74	58	48	

Table 18. Bagging Performance from Combined Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	48	62	88	79	26	99.62
	Precision	47	56	92	65	43	
30	F1	40	60	77	51	41	99.57
	Precision	32	56	100	53	47	
50	F1	39	44	85	66	21	100
	Precision	31	38	87	75	35	

Table 19. RF Performance from Combined Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	57	67	86	80	75	100
	Precision	67	62	90	86	60	
30	F1	46	68	86	71	41	100
	Precision	55	63	95	61	47	
50	F1	41	56	85	58	23	100
	Precision	35	45	87	66	37	

Table 20. GB Performance from Combined Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	47	56	74	75	19	100
	Precision	44	47	71	69	40	
30	F1	46	58	76	70	14	100
	Precision	38	52	89	70	20	
50	F1	44	53	83	57	31	98.80
	Precision	41	51	86	50	42	

Table 21. SVM Performance from Combined Channel. Accuracy: Training Accuracy

Test Percentage		Sleep Stages					Accuracy
		N1	N2	N3+N4	Wake	REM	
20	F1	40	80	96	94	67	84.85
	Precision	100	75	92	89	67	
30	F1	50	71	93	73	67	85.89
	Precision	45	65	100	83	67	
50	F1	42	61	81	67	53	95.20
	Precision	45	53	80	69	62	

Because we’re facing a multi classification problems. Especially a balanced model with five sleep stages, precision is calculated as the sum of true predicted values for sleep stages divided by the total of wrong positives and true positives for different sleep stages. We can define precision and recall for each of the classes. For example, if we take a look on the random forest in the 2 channels experiment with a test percentage of 20%, the precision for the sleep stage 1 is the number of correctly predicted sleep stages which is 16 out of all predicted sleep stages (15+1+0+4+7), which refers to 16/27=56%. Same process applied on all the sleep stages from different channels, for the training accuracy it appoint the sleep stages correctly defined or recognized from the train set which varies depending on the train test split percentage, for example, implementing the second channel for a 30-70 test train percentage respectively, by dint of MLP classifier we achieved 97.43 % which means that from 235 features (70% of the dataset), almost 229 classes were correctly classified and it’s the same process for all the training columns showed in tables above. On the other hand, recall and precision together form the F1-score. Above we saw a summary of the precision and F1 for the five classes, if we take a look on the table 19, for the Wake sleep stage using (20-80) test-train split:

$$F_1 - score(Wake) = 2 * \frac{80\% * 86\%}{80\% + 86\%} = 82\% \tag{15}$$

Where 80% and 86% present the precision and recall respectively of the wake stage. In a similar way, we can perform the F1 score for the other 4 sleep classes via different channel distribution. As seen from the tables above, the remaining sleep classes had different classification accuracies. From our observation, good accuracy results took place in the sleep stage 3 that reached more than once 100 % using different machine learning methods from different type of channel’s organization. Overall, the precision, and training accuracy results for random forest were somehow better for all sleep stages except the 2 first evoked responses belonging to sleep stage 1 and 2 respectively. About the test classification accuracy, the SVM proved that it can be a significant model for multiple sleep stages classification, despite the fact of being a binary classifier, it per- formed well face to 5 sleep stages. With using the symlet 12 as a type of wavelet, 20% and 80% as testing training percentage respectively, from first channel (EEG Fpz-Cz) we achieved a test accuracy of 71.64%. The next experiment investigates the use of SVM with second channel (EEG Pz-Oz), adopting the same data splitting percentage, the algorithm attained almost 86% accuracy, which is the best test result if we consider that we would like to avoid an overfitted model. Using the 2 channels above combined allowed us to reach various accuracies depending on the test-train percentage, 85.17% of precision using 80%-10%-10% of train-test-valid data respectively, 60% using the 50-50 test train, which was not an efficient split idea due to the limited size of dataset used in this experience. The sleep stage 3 was easy to predict even for a bad test-train split percentage, the wake stage was hard to predict for other classifiers, the random forest using EEG Pz-Oz achieved 88%. About the REM the top score of identification was 80% performed by linear discriminant analysis using the EEG channels combined.

CHAPTER 5: CONCLUSION

This Thesis provided a global, overall survey of automatic sleep stages classification applied evoked response and machine learning. Including preprocessing, feature extraction, normalization and classification. Our algorithm map has four main sections. The study affords important information for the reader to explore various biomedical signal handling methodologies that have been adopted for same purpose, also negotiate their efficiency. Moreover, a new method was adopted that could be easily applied using EEG signals and various evoked responses from different signals captured from human brain. The principal method of the EEG multi-canal input uses mainly denoising techniques to clear the raw of noise. The technique developed in this work, used evoked response instead of using the total of the grouped epochs which can be hard to do with limited hardware techniques. Statistical features based discrete wavelet transform provided a productive approach for extracting and well describing important features from different sleep stages. The use of normalization is due to change numerical columns in the data frame to a common scale, with respect to the ranges of values. It is needed to avoid different extracted features ranges. The experiments managed in this thesis used various machine learning classifiers. The proposed method achieves a test accuracy of 85% using SVM. Therefore, our feasible methodology makes our approach faster and easy to implement using evoked responses instead of heavy files of EEG channels, it is worth mentioning that the more our dataset is larger the more our model is flexible to achieve better accuracy, thing we realized after using the 2 EEG channels combined even with deficient machine learning methods.

5.1 Future Work

Since our model achieved a good accuracy compared to other’s seen in the Literature review, we hope to develop an enhanced model for Epileptic seizure prediction, which is a global neurological condition affecting more than 50 million people. In recent years, there were several initiatives in medical science to better meet modern diagnosis and therapies. Even though we can use a cloud-based application monitoring for epileptic seizure prediction. Wearable sensors, smartphones and other device can help to conduct the input EEG signal by using multiple channel (more than 100 channel), these devices are connected to the cloud using a communication protocol. In order to classify the patient case (Emergency, Home Care, doctor.), our methodology would be based on feature extraction from evoked responses, feature selection due to the large number of features and then classification using deep learning methods (LSTM, CNN, RNN…).

REFERENCES

Čić, M., Šoda, J. and Bonković, M. (2013) *Automatic classification of infant sleep based on instantaneous frequencies in a single-channel EEG signal.* Computers in *Biology and Medicine,* 43(12), pp. 2110-2117.

Aboalayon, K. A. and Faezipour, M. (2014) *Multi-class SVM based on sleep stage identification using EEG signal.* Seattle, WA, USA, IEEE, pp. 181- 184.

Aboalayon, K. A., Faezipour, M., Almuhammadi, W. S. and Moslehpour , S. (2016) *Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation.* Entropy, 18(9), p. 272-313.

Adeli, H., Zhou, Z. and Dadmehr, N. (2003) *Analysis of EEG records in an epileptic patient using wavelet transform*. Journal of Neuroscience Methods, February, 123(1), pp. 69-87.

Alpaydin, E. (2020) *Introduction to Machine Learning.* s.l.:MIT Press.

Barni, M. and Bartolini, F. (2004) *Watermarking systems engineering: enabling digital assets security and other applications.* 1st ed. s.l.:CRC Press.

Bin Heyat, M. and Siddiqui, M. M. (2015) *Recording of EEG , ECG , EMG Signal.* International Journal of Advanced Research in Computer Science and Software Engineering, October, 5(10), pp. 813-815.

Bodenstein, G., Schneider, W. and Malsburg, C. (1985) *Computerized EEG pattern classification by adaptive segmentation and probability-density-function classification. Description of the method.* Computers in Biology and Medicine, 15(5), pp. 297-313.

Breiman , L. (1996) *Bagging predictors.* Machine Learning , August, Volume 24, pp. 123-140.

Brown, L. K. (2012) *Can sleep deprivation studies explain why human adults sleep?.* Current Opinion in Pulmonary Medicine, November, 18(6), pp. 541-545.

Carlson, N. R. (2005) *Foundations of physiological psychology.* New Zealand: Pearson Education.

Carskadon, M. A. and William C., D. (2005) *Normal Human Sleep : An Overview.* In: 4th, ed. Principles and practice of sleep medicine. s.l.:Elsevier, pp. 13-23.

Chatterjee, R. and Bandyopadhyay, T. (2016) *EEG Based Motor Imagery Classification Using SVM and MLP.* Bhubaneswar, India, IEEE, pp. 84-89.

Chen, X., Wang, B. and Wang, X. (2013) *Automatic Sleep Stage Classification for Daytime Nap Based on Hopfield Neural Network.* Guiyang, China, IEEE, pp. 2671-2674.

Cirelli, C. and Tononi, G. (2008) *Is Sleep Essential?.* PLoS Biology, August, 6(8), p. 1605-1611.

Collura, T. F. (1993) *History and evolution of electroencephalographic instruments and techniques.* Journal of clinical neurophysiology, 10(4), pp. 476-504.

Correa, A. G. and Leber, E. L. (2010) *An automatic detector of drowsiness based on spectral analysis and wavelet decomposition of EEG records.* Buenos Aires, Argentina, IEEE, pp. 1405-1408.

Correa, A. G., Orosco, L. and Laciar, E. (2014) *Automatic detection of drowsiness in EEG records based on multimodal analysis.* Medical Engineering & Physics, 36(2), pp. 244-249.

Dement, W. and Kleitman, N. (1957) *Cyclic variations in EEG during sleep and their relation to eye movements, body motility, and dreaming.* Electroencephalography and Clinical Neurophysiology, 9(4), pp. 673-690.

Fawcett, T. (2006) *An introduction to ROC analysis.* Pattern Recognition Letters, Volume 27, pp. 861-874.

Fisch, B. and Spehlmann, R. (1999) *Fisch and Spehlmann's EEG Primer: Basic Principles of Digital and Analog EEG.* s.l.:Elsevier Health Sciences.

Flach, P. (2012) *MACHINE LEARNING The Art and Science of Algorithms that Make Sense of Data.* s.l.:Cambridge University Press.

Fraiwan, L., Lweesy, K., Khasawneh, N., Wenz, H. and 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), pp. 10-19.

Freeman, W. J. and Zhai, J. (2009) *Simulated power spectral density (PSD) of background electrocorticogram (ECoG).* Cognitive Neurodynamics, 3(1), pp. 97-103.

Friedman, J. H. (2002) *Stochastic gradient boosting.* Computational Statistics & Data Analysis, February, 38(4), pp. 367-378.

Gais, S., Mölle, M., Helms, K. and Born, J. (2002) *Learning-Dependent Increases in Sleep Spindle Density.* The Journal of Neuroscience, August, 22(15), pp. 6830-6834.

Ghaedi, A., Smaeili , M., Ghasemian Moghadam, M. R. and Izadi, B. (2019) *Neuro - Psychological Investigation of Famous Athletes' Endorsement on Decision Purchasing Sport and Non-Sport Products.* Communication Management in Sport Media, 7(2), pp. 41-51.

Goel, E. and Abhilasha, E. (2017) *Random Forest: A Review.* International Journal of Advanced Research in Computer Science and Software Engineering, 7(1), pp. 251-257.

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. Ch., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C. and Stanley, H. E. (2000) *PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals.* Circulation, 101(23), pp. E215-E220.

Guan, S., Zhao, K. and Yang, S. (2019) *Motor Imagery EEG Classification Based on Decision Tree Framework and Riemannian Geometry*. Computational Intelligence and Neuroscience, Volume 2019, pp. 1-13.

Gudmundsson, S., Runarsson, T. P. and Sigurdsson, S. (2005) *Automatic Sleep Staging using Support Vector Machines with Posterior Probability Estimates.* Vienna, Austria, IEEE, pp. 366-372.

Guo, L., Rivero, D., Seoane, J. A. and Pazos, A. P. (2009) *Classification of EEG Signals Using Relative Wavelet Energy and Artificial Neural Networks.* Shanghai, China, Association for Computing Machinery (ACM), pp. 177-183.

Harrington, P. (2012) *Machine Learning in Action.* s.l.:Manning Publications Co.

Hassan, A. R., Bashar, S. K. and Hassan Bhuiyan, M. I. (2015) *On the classification of sleep states by means of statistical and spectral features from single channel Electroencephalogram.* Kochi, India, IEEE, pp. 2238-2243.

Hastie, T., Tibshirani, R. and Friedman, J. (2009) *The Elements of Statistical Learning Data Mining, Inference, and Prediction.* 2nd ed. Verlag New York: Springer.

Hazarika, N., Chen, J. Z., Tsoi, A. C. and Sergejew, A. (1997) *Classification of EEG signals using the wavelet transform.* Signal Processing, Volume 59, pp. 61-72.

Hsu, Y. L., Yang, Y. T. C., Wang, J. S. and Hsu, C. Y. (2013) *Automatic sleep stage recurrent neural classifier using energy features of EEG signals.* Neurocomputing, Volume 104, pp. 105-114.

Huang, C. S., Ling Lin, C., Yang, W., Ko, L., Yi Liu, S. and Teng Lin, C. (2013) *Applying the fuzzy c-means based dimension reduction to improve the sleep classification system.* Hyderabad, India, IEEE, pp. 1-5.

Imtiaz, S. A. and Rodriguez-Villegas, E. (2015) *An open-source toolbox for standardized use of PhysioNet Sleep EDF Expanded Database.* Milan, Italy, IEEE, pp. 6014-6017.

Jasper, H. H. (1958) *The ten twenty electrode system of the international federation.* Electroencephalogr Clin Neurophysiol, Volume 10, pp. 371-375.

Kemp, B. (1987) *Model-based monitoring of human sleep stages.* s.l.:s.n.

Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C. and Oberyé, 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), pp. 1185-1194.

Khalighi, S., Sousa, T., Oliveira, D., Pires, G. and Nunes, U. (2011) *Efficient feature selection for sleep staging based on maximal overlap discrete wavelet transform and SVM.* Boston, MA, USA, IEEE, pp. 3306-3309.

Khalighi, S., Sousa, T., Pires, G. and Nunes, U. (2013) *Automatic Sleep Staging: A Computer Assisted Approach for Optimal Combination of Features and Polysomnographic Channels.* Expert Systems with Applications, 40(17), pp. 7046-7059.

Killgore, W. D. S. (2010) *Effects of sleep deprivation on cognition.* In: *Progress in Brain Research.* s.l.:Elsevier, pp. 105-129.

Klem, G. H., Lüders, H. O., Jasper, H. H. and Elger, C. (1961) *The Ten Twenty Electrode System: International Federation of Societies for Electroencephalography and Clinical Neurophysiology.* American Journal of EEG Technology, 1(1), pp. 13-19.

Koch, H., Christensen, J. A.E., Frandsen, R., Zoetmulder, M., Arvastson, L., Christensen, S. R., Jennum, P. and Sorensen, H. B.D. (2014) *Automatic sleep classification using a data-driven topic model reveals latent sleep states.* Journal of Neuroscience Methods, Volume 235, pp. 130-137.

Krakovská, A. and Mezeiová, K. (2011) *Automatic sleep scoring: A search for an optimal combination of measures.* Artificial Intelligence in Medicine, 53(1), pp. 25-33.

Kryger, M. H., Roth, T. and Dement, W. C. (2005) *Mammalian Sleep.* In: *Principles and Practice of Sleep Medicine.* s.l.:Elsevier, pp. 91-100.

Kryger, M. H., Roth, T. and Dement, W. C. (2005) *Normal Human Sleep: An Overview.* In: *Principles and Practice of Sleep Medicine.* s.l.:Elsevier, pp. 13-23.

Kumari B. M, K. (2014) *Review on Drowsy Driving: Becoming Dangerous Problem.* International Journal of Science and Research (IJSR), 3(1), pp. 49-51.

Lan, K. C., Chang, D. W., Kuo, C. E., Wei, M. Z., Li, Y. H., Shaw, F. Z. and Liang, S. F. (2015) *Using off-the-shelf lossy compression for wireless home sleep staging.* Journal of Neuroscience Methods, Volume 246, pp. 142-152.

Liang, S. F., Kuo, C. E., Hu, Y. H. and Cheng, Y. S. (2012) *A rule-based automatic sleep staging method.* Journal of Neuroscience Methods, 205(1), pp. 169-176.

Loomis, A. L., Harvey, N. and Hobart, G. A. (1937) *Cerebral States during Sleep as Studied by Human Brain Potentials.* Journal of Experimental Psychology, August, 21(2), pp. 127-144.

Mohri, M., Rostamizadeh, A. and Talw, A. (2018) *Foundations of Machine Learning.* 2nd ed. s.l.:MIT Press.

Moser, D., Anderer, P., Georg, G., Loretz, E., Boeck, M., Kloesch, G., Heller, E., Schmidt, A., Danker-Hopfe, H., Saletu, B., Zeitlhofer, J. and Dorffner, G. (2009) *Sleep classification according to AASM and Rechtschaffen and Kales: effects on sleep scoring parameters.* Sleep, Feb, 32(2), pp. 49-139.

Mourtazaev, M., Kemp, B., Zwinderman, A. and Kamphuisen, H. A. (1995) *Age and gender affect different characteristics of slow waves in the sleep EEG.*. Sleep, 18(7), pp. 557-564.

Mukherjee, S. R., Patel, S. N., Kal, S., T-Ayas, N., P-Strohl, K., Gozal, D. and Malhotra, A. (2015) *An Official American Thoracic Society Statement: The Importance of Healthy Sleep. Recommendations and Future Priorities.* American Journal of Respiratory and Critical Care Medicine, 15 Jun, 191(12), pp. 1450-1458.

Niedermeyer, E. and Silva, F. H. L. d. (2005) *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.* s.l.:Lippincott Williams & Wilkins.

Radha, M., Garcia-Molina, G., Poel, M. and Tononi, G. (2014) *Comparison of feature and classifier algorithms for online automatic sleep staging based on a single EEG signal.* Chicago, IL, USA, IEEE, pp. 1876-1880.

Raschka, S. (2015) *Python Machine Learning.* s.l.:Packt Publishing Ltd.

Reddy Edla, D., Mangalorekar, K., Dhavalikar, G. and Dodi, S. (2018) *Classification of EEG data for human mental state analysis using Classification of EEG data for human mental state analysis using Random Forest Classifier.* Procedia Computer Science, Volume 132, pp. 1523-1532.

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

Ryali, S., Supekar, K., Abrams, D. A. and Menon, V. (2010) *Sparse logistic regression for whole brain classification of fMRI data.* NeuroImage, June, 51(2), pp. 752-764.

Sörnmo, L. and Laguna, P. (2005) *Bioelectrical Signal Processing in Cardiac and Neurological Applications.* s.l.:Academic Press.

Sanders, T. H., McCurry, M. and Clements, M. A. (2014) *Sleep stage classification with cross frequency coupling.* Chicago, IL, USA, IEEE, pp. 4579-4582.

Sanei, S. and Chambers, J. (2013) *EEG Signal Processing.* s.l.:Wiley-Interscience.

Sarkar, D., Bali, R. and Sharma, T. (2018) *Practical Machine Learning with Python.* Berkeley: Apress.

Schaul, N. (1998) *The fundamental neural mechanisms of electroencephalography.* Electroencephalography and clinical Neurophysiology, 106(2), pp. 101-107.

Semmlow, J. L. (2004) *Biosignal and Biomedical Image Processing.* s.l.:Marcel Dekker, Inc.

Sengür, A. (2009) *Multiclass least-squares support vector machines for analog modulation classification.* Expert Systems With Applications, 36(3), pp. 6681-6685.

Siuly, S. and Li, Y. (2012) *Improving the Separability of Motor Imagery EEG Signals Using a Cross Correlation-Based Least Square Support Vector Machine for Brain-Computer Interface.* IEEE Transactions on Neural Systems and Rehabilitation Engineering, 20(4), pp. 526-538.

Siuly, S. and Li, y. (2015) *Designing a robust feature extraction method based on optimum allocation and principal component analysis for epileptic EEG signal classification.* Computer Methods and Programs in Biomedicine, 119(1), pp. 29-42.

Siuly, S., Li, Y. and Wen, P. (2011) *EEG signal classification based on simple random sampling technique with least square support vector machine.* International Journal of Biomedical Engineering and Technology (IJBET), 7(4), pp. 390-409.

Siuly, S., Li, Y. and Zhang, Y. (2016) *EEG Signal Analysis and Classification Techniques and Applications.* 1 ed. Cham, Switzerland: Springer International Publishing.

Siuly, S., Li, Y. and Zhang, Y. (2016) *Electroencephalogram (EEG) and Its Background.* In: *EEG Signal Analysis and Classification.* s.l.:Springer International Publishing, pp. 3-21.

Siuly, S., Li, Y. and Zhang, Y. (2016) *Electroencephalogram (EEG) and Its Background.* In: *EEG Signal Analysis and Classification: Techniques and Applications.* Cham, Switzerland: Springer International Publishing, pp. 3-21.

Subasi, A. (2019) *Electromyogram-controlled assistive devices.* In: *Bioelectronics and Medical Devices.* s.l.:Elsevier, pp. 285-311.

Subasi, A. (2020) *Practical Machine Learning for Data Analysis Using Python.* 1st ed. s.l.:Academic Press .

Subramanian, V. (2018) *Deep learning with PyTorch : a practical approach to building neural network models using PyTorch / Vishnu Subramanian..* Birmingham: Packt Publishing Ltd.

Witten, I., Frank, E. and Hall, M. (2011) *Data Mining: Practical Machine Learning Tools and Techniques.* 6th ed. s.l.:Morgan Kaufmann.

Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z. H., Steinbach, M., Hand, D. J. and Steinberg, D. (2008) *Top 10 algorithms in data mining.* Knowledge and Information Systems, 14(1), pp. 1-37.

Zepelin, H., Siegel, J. M. and Tobler, I. (2005) *Mammalian Sleep.* In: *Principles and Practice of Sleep Medicine.* 4th, ed. s.l.:Elsevier, pp. 91-100.

Zhu, G., Li, Y. and Wen, P. (2014) *Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal.* IEEE Journal of Biomedical and Health Informatics, November, 18(6), pp. 1813-1821.

Zoubek, L., Charbonnier, S., Lesecq, S., Buguet, A. and Chapotot, F. (2008) *Feature selection for sleep/wake stages classification using data driven methods.* Biomedical Signal Processing and Control, 2(3), pp. 171-179.