BRNO UNIVERSITY OF TECHNOLOGY
VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF BIOMEDICAL ENGINEERING
ÚSTAV BIOMEDICÍNSKÉHO INŽENÝRSTVÍ

SLEEP SCORING USING ARTIFICIAL NEURAL NETWORKS
VYUŽITÍ UMĚLÝCH NEURONOVÝCH SÍTÍ PRO HODNOCENÍ SPÁNKU

BACHELOR'S THESIS
BAKALÁŘSKÁ PRÁCE

AUTHOR — **Zuzana Vašíčková**
AUTOR PRÁCE

SUPERVISOR — **Ing. Martin Králík**
VEDOUCÍ PRÁCE

BRNO 2018

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Bachelor's Thesis

Bachelor's study field **Biomedical Technology and Bioinformatics**
Department of Biomedical Engineering

Student: Zuzana Vašíčková ***ID:*** 186699

Year of study: 3 ***Academic year:*** 2017/18

TITLE OF THESIS:

Sleep scoring using artificial neural networks

INSTRUCTION:

1) Get acquainted with issues of sleep physiology and polysomnography 2) Get acquainted with different classifiers focusing on artificial neural networks. 3) Summarize the issue in a form of literary research and design automatic sleep stage classifier using artificial neural networks. 4) Find suitable features and test their suitability. Use data availible at UBMI and in open databases. 5) Implement your solution for sleep scoring, comment your results and compare them with available articles.

RECOMMENDED LITERATURE:

[1] DOUGHERTY, Geoff. Pattern recognition and classification: an introduction. New York: Springer Science+Business Media, 2012. ISBN 9781461453222.

[2] Review of Sleep Medicine. 3rd Ed. Editor Alon Y. AVIDAN, editor Teri J. BARKOUKIS. Philadelphia: Elsevier Saunders, 2011. ISBN 9781455703197.

Date of project specification: 5.2.2018 ***Deadline for submission:*** 25.5.2018

Leader: Ing. Martin Králík
Consultant:

prof. Ing. Ivo Provazník, Ph.D.
Subject Council chairman

ABSTRACT

The main objective of the thesis is to create an artificial neural network for sleep-staging. Firstly, information about sleep and sleep stages is summarized. However, more profound overview of signal processing methods and methods of classification is found in next chapters. After summarizing the theoretical knowledge necessary to carry out practical part of thesis, signal features were extracted according to the theoretical proposal and using statistical analysis, convenient features were used as in input for artificial neural network, able to classify the sleep data into sleep stages after the learning stage.

KEYWORDS

sleep, sleep stages, classification, signal processing, artificial neural networks, polysomnography, features

ABSTRAKT

Hlavným cieľom semestrálnej práce je vytvorenie umelej neurónovej siete, ktorá bude schopná roztriediť spánok do spánkových epoch. Na začiatku je uvedené zhrnutie informácií o spánku a spánkových epochách. V ďalších kapitolách sa nachádza dôkladnejší prehľad metod na spracovávanie signálov a na klasifikáciu. Po zhrnutí teoretických znalostí potrebných na uskutočnenie praktickej časti práce boli na základe tohto rozboru vypočítané zo signálov potrebné znaky. Tieto znaky boli podrobené štatistickej analýze a na jej základe boli vybrané niektoré znaky, ktoré boli vhodné ako vstup do neurónovej siete, ktorá je po naučení schopná triediť spánkové epochy do príslušných fáz.

KLÍČOVÁ SLOVA

spánok, spánkové štádiá, klasifikácia, spracovanie signálu, neuronové siete, polysomnografia, príznaky

VAŠÍČKOVÁ, Zuzana. *Sleep scoring using artificial neural networks*. Brno, 2018, 51 p. Bachelor's Thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication, Department of Biomedical Engineering. Advised by Ing. Martin Králík

DECLARATION

I declare that I have written the Bachelor's Thesis titled "Sleep scoring using artificial neural networks" independently, under the guidance of the advisor and using exclusively the technical references and other sources of information cited in the thesis and listed in the comprehensive bibliography at the end of the thesis.

As the author I furthermore declare that, with respect to the creation of this Bachelor's Thesis, I have not infringed any copyright or violated anyone's personal and/or ownership rights. In this context, I am fully aware of the consequences of breaking Regulation § 11 of the Copyright Act No. 121/2000 Coll. of the Czech Republic, as amended, and of any breach of rights related to intellectual property or introduced within amendments to relevant Acts such as the Intellectual Property Act or the Criminal Code, Act No. 40/2009 Coll., Section 2, Head VI, Part 4.

Brno ....

author's signature

ACKNOWLEDGEMENT

I would like to thank the supervisor of the semestral thesis Ing. Martin Králík for his professional mentoring, consultations, patience, and thoughtful suggestions concerning the bachelor's thesis.

Brno . . . . . .

. . . . . .

author's signature

CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

Sleep scoring is an inevitable part of sleep studies and diagnosis of sleep disorders. As sleep directly influences our bodily functions and therefore also the quality of life, it seems important to improve the diagnosis of such disorders and through sleep studies optimize the treatments for certain disorders. The cornerstone of these procedures is sleep scoring and hypnogram obtained from the scoring. Hypnogram shows the relative representation of sleep stages throughout the sleep and it is useful to save time when evaluating sleep. The bachelor's thesis focuses on automatizing the process of sleep scoring to save time and avoid subjective mistakes of manual sleep scoring.

The thesis is divided into two parts - theoretical and practical. The theoretical part is devoted to information concerning sleep, rules of sleep scoring, sleep disorders, polysomnography and most commonly recorded signals, automatic classifiers and basics of signal processing. The practical part is dedicated to creation of MATLAB based neural network for automatic sleep scoring. Firstly, the needed signals were processed and important feature signals were extracted, and according to the obtained information the signals were classified into sleep stages using the artificial neural network.

1 SLEEP

Sleep is one of the basic physiological needs and it is a must for appropriate function of immunity, thermoregulation, and other important processes in our bodies. As the faster lifestyle becomes more important the research of sleep became more important too as the quality of sleep significantly influences the quality of life. The diagnostics of sleep disorders and its consequences are being studied.

However, the analysis of sleep records, which are extensive, is extremely difficult and time consuming for manual evaluation. Due to this problem the need of automatized classification of the sleep stages and the following diagnostics arises [1].

1.1 Sleep Disorders

Insomnia

Insominia is one of the most commonly diagnosed sleep disorders nowadays. When suffering from this disorder the problems are associated with having problems falling asleep, waking up during night time, or with the length of sleep. Insomnia can occur in an acute or chronic form.

A sleep diary can lead to determination of medical, mental, circadian, or other sides of the problem. As a result, the sleep diary is necessary for the correct diagnosis of insomnia.

Insomnia can be divided into two classes, according to the reason of its occurrence. It can be either primary, when insomnia is present without any other disorder, or secondary, when insomnia is a co-morbidity to other disorders [2].

Hypersomnia and narcolepsy

All the cases when patient suffers from an excessive need to sleep during the day are summed up under the term of hypersomnia. Therefore, even people not sleeping more than usually, but still experiencing these feelings, suffer from hypersomnia. Narcolepsy can be also found under this term, although the patient with narcolepsy usually doesn't experience excessive need to sleep, they rather experience sudden falling asleep [1].

Narcolepsy is characterized by narcolepsy tetrade. This typical tetrade includes the excessive need to sleep, cataplexy, hallucinations, and sleep paralysis. The most distinguishing sign of narcolepsy is the cataplexy, when sudden loss of muscle tone occurs for several seconds or minutes.

The essential sign of hypersomnia is sleepiness. Often the affected confuses sleepiness with fatigue. Sleepiness has various symptoms which are not typical for

fatigue. First and the most fundamental sign of sleepiness is the ability of falling asleep under any circumstances. The second symptom is subjective perception of lowered ability to react. The third one is the failure of normal function during the day.

We differ between idiopathic hypersomnia and returning one. Idiopathic hypersomnia is further divided into idiopathic hypersomnia with long sleep times and idiopathic hypersomnia without long sleep times. When it comes to returning hypersomnia, we recognize two main disorders - Kleine-Levine syndrome and hypersomnia associated with menstrual cycle [2].

Sleep apnea

Sleep disorders induced by breathing problems are an important medical issue, endangering the patient's life. Two main types of apnea are distinguished - obstructive apnea or central apnea.

American Academy of Sleep Medicine (AASM) established criteria for scoring of sleep apnea and according to them apnea is characterized by at least 10 seconds long interruption of breathing. The difference between obstructive and central apnea lays in the cause of the interruption. During central apnea no effort to breathe is present, while when it comes to obstructive apnea we observe the effort. Mixed apnea is also common, when in the beginning no effort to breathe occurs and later, even despite of the effort to breathe, it is not possible anymore. Mixed apnea usually falls under obstructive apnea [2].

1.2 Standards of Sleep Scoring

The initial knowledge of mankind about sleep led only to observation of regular alternation between wakefulness and sleep. Only during the 1920s the sleep research started on both healthy and ill subjects which induced the urge to create unified system for scoring sleep. Initially, the scoring system focused entirely on electroencephalogram (EEG) records and as other records, such as electromyogram (EMG) or electrooculogram (EOG) were added, the scoring system became more complicated.

In 1968 Rechtschaffen and Kales published the *Manual of Standardized Terminology, Techniques, and Scoring System for Sleep Stages of Human Subjects* and unified the scoring system. According to this manual sleep is divided into 20 or 30 seconds long epochs. The main problem of this publication was its focus solely on healthy subjects. Therefore the diagnosis of a patient with a disorder was problematic. Regardless of the manual's inaccuracy it persisted as the standard for sleep scoring until recently.

Today the common standard for sleep scoring was created by AASM. The manual of AASM divides sleep into three main stages – W (Wakefulness), Rapid Eye Movements (REM), and Non-rapid Eye Movements (NREM), divided further into N1, N2, and N3. The epochs of 30 seconds are still applied and if we observe two stages in one epoch we mark the prevailing stage [1],[3].

1.3 Sleep Stages

Stage W – "Wakefulness"

Stage W can be defined as a stage where the subject is in a state of wakefulness or drowsiness. The distinguishing feature for this stage is at least 50 % prevalence of alpha waves (Fig. 1.1) in EEG recordings. These alpha waves are in the range of frequencies from 8 to 13 Hz, but mostly 9-11 Hz in adults.

Fig. 1.1: Alpha waves [2].

Alpha rhythm can be present only with closed eyes, when the eyes are opened it is replaced with low-amplitude mixed-frequency EEG pattern.

Additional signs of this stage, even without the alpha rhythm being present, are reading eye movements, REM or eye blinks with frequency of 0.5-2 Hz [4].

Stage N1

On EEG the low-amplitude mixed-frequency waves within the range of 4-7 Hz (Fig. 1.2) and sharp vertex waves, defined as sharp negative waves followed by positive wave shorter than 0.5 s, can be regarded. Slow eye movements shown on the EOG can be seen as well. EMG is slightly elevated, yet it is still lower than during stage W.

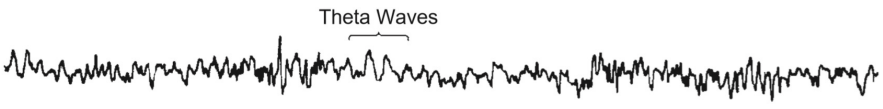

Fig. 1.2: Theta waves in stage N1 [2].

Stage is clearly marked as N1 if at least 50 % of alpha rhythm was replaced by low-amplitude mixed-frequency waves. When no alpha rhythm is shown on the

EEG, the stage is determined by presence of either low-amplitude mixed-frequency waves in the range 4-7 Hz, slow eye movements, or vertex waves. If no k-complexes associated with external arousal are exhibited, the epoch is still marked as N1 [4].

Stage N2

Appearance of k-complexes and sleep spindles within frequencies of 13-16 Hz is essential for stage N2. They intertwine with theta waves of frequency 12-14 Hz as seen on Fig. 1.3.

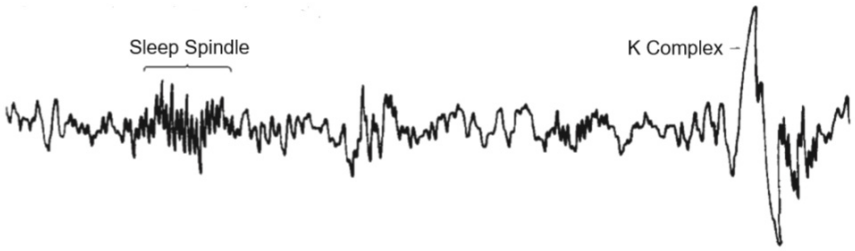

Fig. 1.3: Sleep spindles and k-complexes [2].

K-complex consists of a negative short wave followed by a positive component standing out from a background EEG. The length of one k-complex is usually not more than 0.5 s. No association with arousal can be made with k-complex, meaning no external arousal can be within 1 second range around k-complex. Sleep spindles are a train of distinctive waves with frequency 11-16 Hz and its total duration is no longer than 0.5 s.

The epoch is marked as N2 if at least one k-complex or sleep spindle can be found in the first half of that epoch or the second half of the epoch before. None of these k-complexes or sleep spindles can be associated with external arousal.

The same activity as on EEG can be found on EOG. EMG is elevated, but still lower than in stage W.

Stage N2 ends when it changes into the stage N3, W, or REM. It changes into N1 or W either if an arousal is observed in EEG, or the conditions of stage N2 are not met. Major Body Movement (MBM) can be found during this stage. MBM are usually followed by slow eye movements and low-amplitude mixed-frequency waves with no k-complexes or sleep spindles caused by arousal. If no eye movements are observed after MBM, the stage is continued to be marked as N2. MBM can cover up more than half of the epoch. In that case we mark it as stage W if stage W precedes or follows this epoch, or if the alpha rhythm is present. If no alpha rhythm is present, we mark it as the following stage [4].

Stage N3

In stage N3 the typical sign is the presence of slow waves occurring at least in 20 % of the epoch, they reach frequencies of 0.5-3 Hz and the amplitude of slow waves has to be at least 75 µV from peak to peak. Sleep spindles are not considered important while marking stage N3 and k-complexes are considered to be slow waves, if they meet the definition of slow waves when it comes to amplitude and frequency. Even though the amplitude drops with growing age, it should always reach the minimum of 75 µV. The activity of EOG is similar to EEG and the amplitude of EMG is changing and it is relatively low [4].

Stage R

The name of R stage stems in the REM captured on EOG and they are the definite sign of this stage. REM are conjugate, irregular and sharp waves with initial deflection (Fig 1.4). The duration of these waves is generally shorter than 500 ms. On EEG sawtooth waves can appear as they can precede rem. These triangular toothed waves are in frequencies of 2-6 Hz. Other than sawtooth waves also theta waves within the range of 4-7 Hz and less frequent alpha waves can occur.

Fig. 1.4: Sawtooth waves occuring in REM sleep [2].

EMG signal sensed from the chin is during this stage in total minimum out of all the stages and the muscle tone is quite low in other places as well.

Transient muscle activity, which lasts less than 0.25 s, can be found on EMG and is quite significant on the lowered EMG. The usual place for observation of transient muscle activity is chin or anterior tibialis. To mark an epoch as a stage R must be present low-amplitude mixed-frequency waves on EEG, minimal values of chin EMG, and REM.

An epoch is still marked as stage R even if no REM occur, if low-amplitude mixed-frequency EEG signal persists, EMG is still at its minimum, and no k-complexes or sleep spindles are present.

An epoch is no longer marked as R if it changes into stage W or N3, the muscle tone in chin increases, an arousal or MBM followed by slow waves occur, or a k-complex or sleep spindle that is not associated with arousal appears [4].

SLEEP STRUCTURE

Human sleep consists out of repeated stages then forming sleep cycles. Two main

stages of sleep are REM and NREM sleep. The cycle of NREM-REM sleep occurs 4-5 times a night in young adults and lasts 90-110 minutes.

REM sleep altogether makes up 25 % of sleep in adults, while in children NREM and REM are more balanced, usually in ratio 1:1. The remaining 75 % is made up of NREM sleep, which is further divided into three stages - N1, N2, N3. Stage N3 is mostly slow waves and with increasing age makes a smaller part of sleep and it is slowly replaced by stage N2, making up almost 50 % of sleep in adult age. NREM sleep usually occurs in the first half of the sleep and it is unusual to have stage W followed by REM for someone older than 3 months. On the other hand, REM is prevailing in second half of the sleep, which is related to circadian rhythms.

Sleep begins with stage R representing state of wakefulness, characteristic with its presence of alpha waves. Onwards the NREM stage of sleep follows, particularly stage N1. Typical slow eye movements are present along with increased EMG, even though it is lower than in stage R as the muscles start to relax. In the following stage N2 the eye movements cease and the muscle tone lowers even further. The electrical brain activity decreases and EMG values are low in stage N3 points out the beginning of deep sleep. After NREM stage the REM sleep usually follows, but stages W, N1, and N2 can occur in-between. For stage R are typical REM and minimal muscle activity interrupted only by transient muscle activity. After that one full sleep cycle is enclosed and this cycle repeats itself 4-5 times a night [2],[5].

2 POLYSOMNOGRAPHY

To obtain sleep records from which it can be continued to analyze them, we use polysomnography (PSG). The term PSG was first used in 1974 by Holland, Dement, and Raynal. Using PSG, records of various signals are obtained, for other analysis in order to diagnose sleep disorders or for deeper understanding of sleep itself.

The most crucial element of PSG are three electrodes meant for recording EEG, one electrode for EMG and two electrodes for recording EOG, otherwise no useful data can be collected. electrocardiogram (ECG) is a compulsory part of PSG as well. Moreover, the respiratory functions are monitored, being essential in the diagnosis of respiratory disorders during sleep, such as sleep apnea. Breathing can be monitored by different forms, most often nasal cannula, oral thermometer, throat microphone, thoracic and abdominal belt, and pulse oximeter are used. Other sensors or EEG electrodes can be added, as every PSG is issued due to different reasons [1],[2].

2.1 Electroencephalography

EEG captures activity of the brain in time. Multiple scalp electrodes are used to compare electrical signals (electrical potential) from different areas of the brain. The potential of extracellular space arises as a neuron produces neuronal activity, probably due to the extracellular currents generated from post-synaptic potentials and inhibitory postsynaptic potentials. Amplitude of the EEG is higher when more neurons fire at the same time. The electrodes are usually attached to the scalp, therefore the contribution of neurons is quite low as the tissue attenuates the signal significantly [1],[5].

Electrode Placement

The standard system used for measuring EEG is system 10-20, pictured on Fig. 2.1, and the minimum of electrodes is 21. Usually odd numbered electrodes are placed on the left side of the head and even numbered on the right side. The 10-20 system implies the distances from the nasion to inion, as the first electrodes are 10 % of the distance and the others are 20 % of the distance. Every section of the brain has assigned letters of EEG electrodes:

- Fp - frontpolar,
- F - frontal,
- T - temporal,
- C - central,
- P - parietal,
- O - occipital,

- A - auricular.

Subsequently, the signal from one part of the brain is measured in reference to another electrode. The signal gets gradually lower with the distance from the source. Thus the signal is clear usually only from one of the parallel combination of electrodes [1],[5].

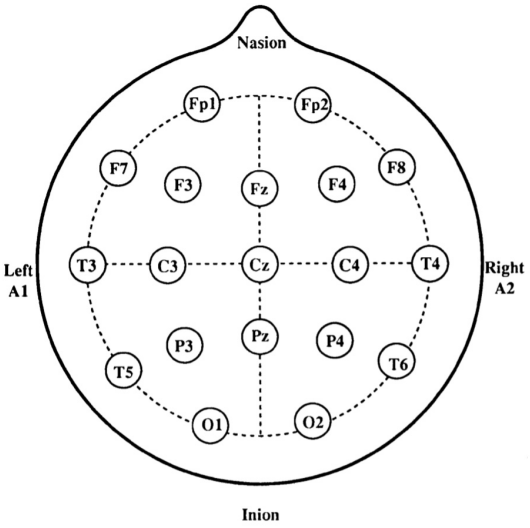

Fig. 2.1: Recommended placement of EEG electrodes [1].

EEG electrodes can be either affixed to the scalp or they can be intracranial. However, during PSG the later one is not used at all. Nowadays we recognize different ways of fixating the scalp electrodes sufficiently enough. Typically, they are made of silver chloride or gold and they are cup-shaped, designated to hold conducting paste, rarely glue to provide better results during recordings that exceed 24 hours [1],[5].

EEG Rhythms

We can observe various kinds of rhythms in EEG. These rhythms are extremely useful when scoring sleep from PSG data.

In normal EEG we recognize these 4 key rhythms:

- Beta waves fill the range of 13-30 Hz and they correspond with active state.
- Alpha waves correspond to range of 8-13 Hz and they are present during stage W.
- Theta waves are in the range of 4-8 Hz while showing drowsiness.
- Delta waves belong to the range up to 4 Hz and they are present during sleep [1],[5].

2.2 Electrocardiography

The heartbeat is controlled by pacemaker cells which produce electrical signals when firing. ECG is the expression of these most commonly recorded biomedical signals. The electrical signal starts at sinoatrial node which is regulated by autonomous system. Afterwards the signal propagates further through atrioventricular node where the signal is delayed, the His bundle, the bundle branches, and Purkinje fibres propagating the signal to the ventricle.

The propagation of the electrical signal throughout the heart causes characteristic heart movements necessary for blood circulation shown on the Fig. 2.2. These movements are captured as distinctive waves on ECG:

- P wave - represents slow depolarization caused by slow propagation of electrical activity through atrial muscles.
- QRS complex - the signal moves rapidly from apex to the top of the heart, depolarization arises as a result.
- T wave - represents repolarization of ventricles [6].

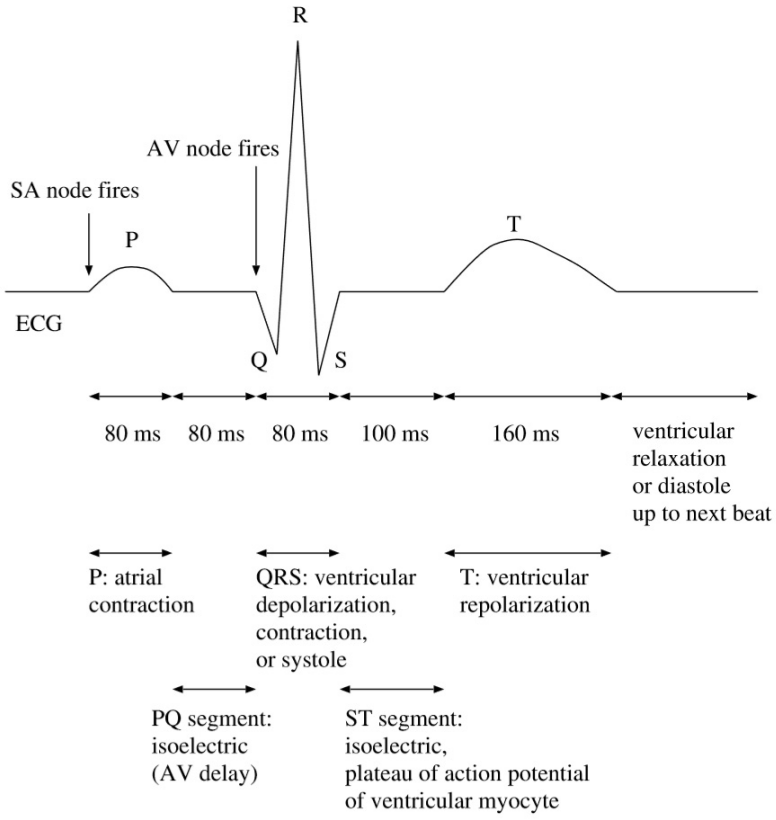

Fig. 2.2: Waves of circulation cycle [6].

2.3 Electrooculography

When considering retina against cornea, retina will act as a negative and cornea as a positive. Therefore EOG is a recording of the cornea moving in respect to retina - eye movement. To measure the signal one electrode is applied for each eye on the canthus, placement of canthuses is shown on Fig. 2.3. On the right eye, the electrode is set off by 1 cm above the horizontal and on the left side the electrode is set off by 1 cm below the horizontal. Two additional electrodes, used as reference electrodes for each side are placed infraorbitally and supraorbitally for each eye [1].

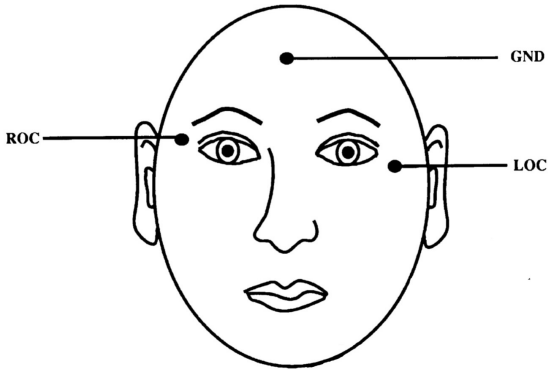

Fig. 2.3: Placement of EOG electrodes. ROC, LOC = right/left outer canthus; GND = ground (earth) [1].

2.4 Electromyography

To record the muscle activity, we measure the summation of activity taking place on the motor plates forming EMG. Being able to determine the muscle tone on chin, on anterior tibialis and other are needed to confirm several sleep stages. EMG electrodes are usually at least three gold or silver-silver chloride cups, to avoid artifacts in the recording [1].

3 METHODS OF AUTOMATIC SLEEP STAGING

As it is necessary to score sleep records automatically, no optimal method to automatically stage sleep can be chosen. While wide variety of automatized classifiers is available, it is still not clear which classifier is the best applicable in this case. The reason behind this is the fact that the method of sleep scoring manually is heavily subjective and varies between doctors as well. The most commonly used classifiers occurring in other studies are mentioned further [7].

3.1 Artificial Neural Networks

Artificial neural networks (ANN) are models based on the human nervous system. Neural networks, both human and artificial, consist of neurons, compared on Fig. 3.1. Human neurons are nerve cells, specialized for transmitting, storing and processing the signal, to ensure the correct response to stimuli. Regarding artificial neurons, they are heavily based on the biological model. Through many dendrites, the spike of electrical activity follows through to nucleus and further the signal transmits along the axon, entering many dendrites again. The mathematical model is called perceptron and perceptron imitates biological neuron, parallel processing information. The artificial neuron has n inputs x_1,x_2,...x_n - modelling the signals coming from dendrites. Afterwards, each input is weighted by real weights w_1,w_2,...w_n and by these weights the input is connected to the output. The neuron fires when the output overcomes a threshold value.

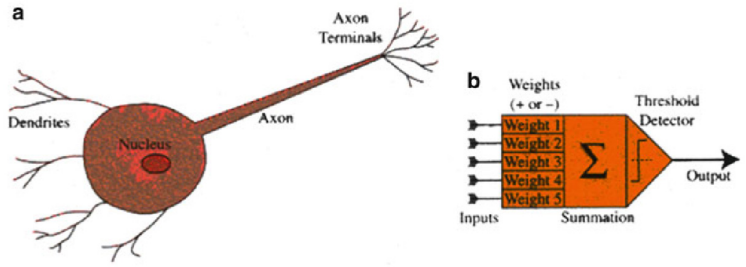

Fig. 3.1: Structure of (a) a neuron and (b) a perceptron [8].

ANN is a computing technique used for various purposes. It is a parallel system, solving problems that are not usually possible to solve through linear computing. Each ANN consist of minimally two layers - input layer and output layer, but usually third, hidden, layer is added. The hidden layer is a layer of cells summing the output as a product of preceding layer after this product is weighing it by weight factor.

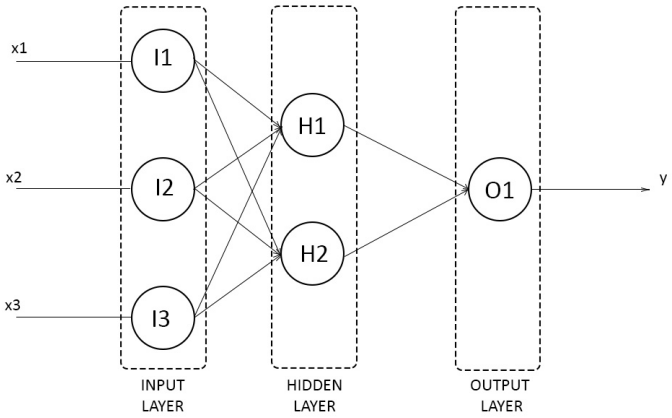

Fig. 3.2: Structure of feedforward three-layered ANN.

Therefore, each neuron produces as an output depending on the given input and weight, according to the activation function.

When designing an ANN only known layers are the input and the output layer and the determination of number of the hidden layers is necessary. On Fig. 3.2 a feedforward ANN with one hidden layer is pictured. During the training phase, the input layer is fed and through trial and error the weights, number of hidden layer and activation function are gradually calculated by the ANN [8],[9].

Learning Rule

As mentioned before, during the training phase the parameters are redefined in order to achieve the ideal values of the output layer. The learning rule is applied during the training phase and supervised and unsupervised learning is recognized.

During the unsupervised learning, the criteria of validity is unknown. The neural network itself must find the correct representation of the input data.

Supervised learning is based on the provision of the input and output data as well - a guide for the network. The weights are computed until adequate values in the output layer are achieved or the maximum of epoch of learning was exceeded. The most common learning rule is the backpropagation algorithm, where input vector and expected output vector is presented to the network and weights and thresholds are set to random low values. After one epoch of learning the values of these parameters are redefined according to the comparison of the expected output and the real output. The weights and thresholds are redefined until the optimal output values are achieved. The principle of the learning algorithm is shown on Fig. 3.3.

Each weight w_j is updated according to:

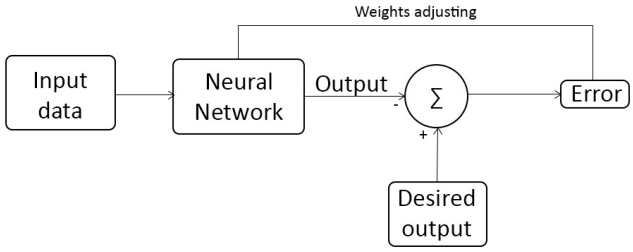

Fig. 3.3: Mulitlayer neural network learning model.

$$w_j^{(k+1)} = w_j^k + \eta(y_i - \hat{y}^{(k)})x_{ij} \tag{3.1}$$

where $w^{(k)}$ is the weight associated with the ith input link after the kth iteration and η is the learning rate, w_{ij} is the value of the jth attribute of the training parameter x_i. The learning rate is between 0 and 1, but typically set to $0.1 < \eta < 0.4$, and in each iteration it is used to control the amount of adjustments. The closer to zero the learning rate is, the more is the new weight influenced by the old weight and if the value of η is closer to 1, the new weight is sensitive to the new adjustment. The learning continues until the error is less than a specified threshold or a predetermined number of iterations has been completed.

Finding the global minimum is calculated as the proportion of misclassified training examples, over a space where all the input values can vary. If one weight is improved too much, then the rest of the weights will lessen. These weights might work for the training data but with other examples the classification might not work as well. When η is too small, the weights will be changed too little in each step and the duration of the algorithm will be quite long. On the other hand, if η is too small the algorithm bounces around the error a diverges. The two marginal cases are shown on Fig. 3.4 [9].

3.2 Decision Trees

Decision tree is an upside-down tree, where root node is located at the very top. The root nose is the starting node and consists of the first question. The branches correspond to possible answers to the preceding question. Other nodes are called decision nodes and other questions are placed on the. In the end, when terminal (leaf) node is reached, the class is assigned. The structure of the tree is not fixed and the tree can grow during the learning process, according to the complexity of the problem.

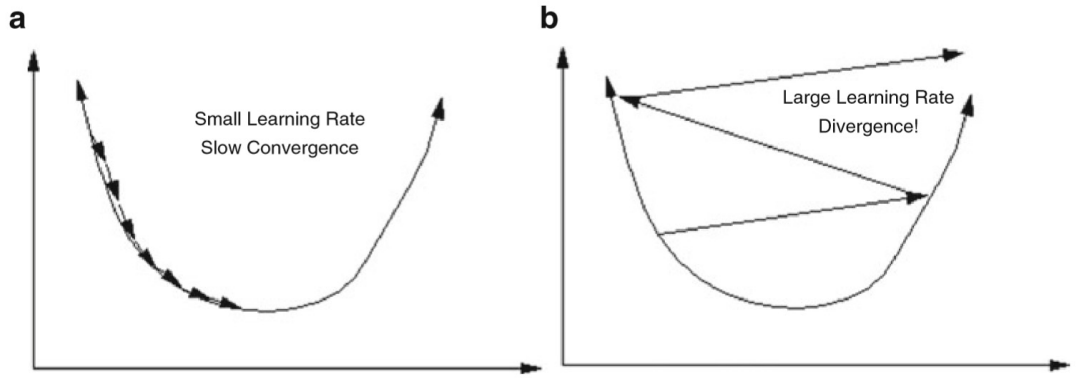

Fig. 3.4: (a) Small value of learning rate resulting in slow convergence and (b) big value of learning rate resulting in divergence [8].

The decision tree itself is not difficult to use, the question of the construction from the training data and the choice of discriminating features [8].

3.3 K- Nearest Neighbors

The k-nearest neighbor (kNN) method is a method that is not fixed on the size. The number of samples k is fixed and the width changes so each region contains the exact number of samples.

The process starts at test point and the region around the test point grows until k samples are enclosed within it. The region is marked after most of its samples is. The kNN method's quality grows with larger k, larger k results in smoother boundaries compared to smaller value of k, where the boundary is convoluted [8].

3.4 Kernel Machines

The kernel machine, originally defined as support vector machines (SVM), is a method of nonparametric supervised learning. Kernel machines can classify both linearly and nonlinearly separable problems. With linearly separable problems the classifier finds the optimal hyperplane separating two classes. To find the hyperplane the machine maximizes the margin - the distance of the closest instances from different classes. The closest instances are called the support vectors (Fig. 3.5), hence the name of the method. The solution of choosing the right hyperplane with nonlinearly separable problems is in the choice of the hyperplane with the least error or the data can be nonlinearly transformed into higher-dimensional space and the right hyperplane is chosen there which is called the kernel trick.

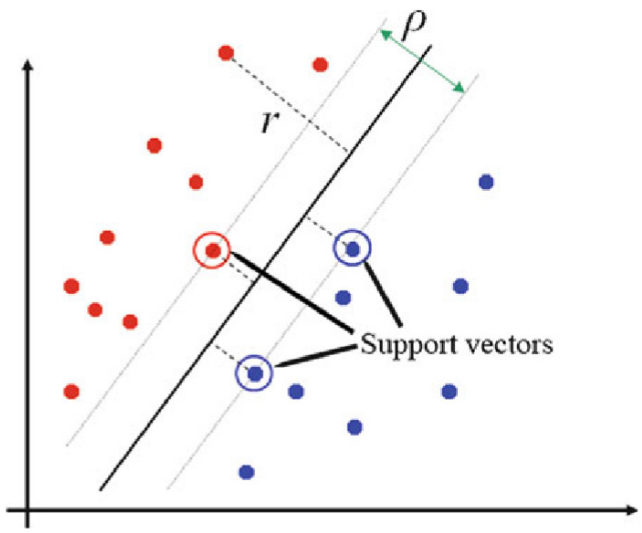

Fig. 3.5: Margin and support vectors [8].

The kernel machines are designed to classify data into two classes only. Anyway, many tasks, sleep scoring as well, require classification into more than two classes. Two approaches are used in multi-class SVM - One-Against-All and One-Against-One approach. The One-Against-All method applies the winner-takes-all strategy, the SVM distinguishes one class from all others. On the other hand, in the one-against-one method more classifiers are used. Every classifier is used for each of all possible pairs of classes [8],[10].

3.5 Clustering

Clustering is a method of unsupervised learning. The class labels are unknown and it is expected that clustering will divide the data into useable classes - clusters. However, created cluster might not resemble human perception of similarity in the data. Cluster should contain all samples that are similar to each other in one way and are different from the rest of the samples. To define the range of similarity, the similarity measure needs to be adopted. The similarity measure is usually represented by measure of proximity - Manhattan distance, Euclidean distance, and others. The criterion function for clustering needs to be defined too. Most typically used criterion function is the sum of the squared error (SSE) where the error is mostly the distance from the closest centroid.

Clusters divide into two basic classes, they can be partitional or hierarchical. In partitional cluster the data is divided into non-overlapping clusters, while in

hierarchical clusters the clusters are nested. Clusters can be also exclusive, where one sample belongs to exactly one cluster or non-exclusive, where one sample belongs to every cluster with different weight between 0 and 1, according to fuzzy logic [8].

K-means clustering

K-means clustering is most widely non-hierarchical clustering algorithm. Every cluster is represented by one prototype object. New data sample is assigned to the nearest prototype and consequently to that cluster.

In the beginning k objects are chosen from the training set to become the prototypes. The rest of the objects are assigned to the nearest objects (according to Euclidean distance or another chosen norm). New prototype of the cluster will become the centroid of all the samples in the given cluster. This process will repeat until the centroids do not change their position or until no data points change clusters.

The results will change according to the initial choice of centroids. On the other hand, due to the algorithm's speed, the process can be repeated more times for different selections of centroids and therefore different results until the optimal result is reached [8].

4 SIGNAL PROCESSING

To begin automatic classification of sleep stages the required signals must undergo pre-processing to dispose of undesired artifacts and as shown on Fig. 4.1, it is beneficial to extract certain features for further classification [11].

Fig. 4.1: Scheme of signal processing.

4.1 Signal Preprocessing

To produce representative signal which can be further worked with, we need to manipulate the original signal. The motivations to do so are various, but usually the reasons behind these motivation are:

- The removal of certain signal components which are corrupt the signal.
- Further extraction of necessary information.
- The prediction of future values.

Biomedical signals, notably PSG signals, are often corrupted by other waves that are not related to the recorded activity. These waves can be caused either by biological or technical issues.

Typical biological artifacts are caused by body's own electrical activities which can be hardly avoided. The most typical artifacts are:

- ECG artifacts – e.g. spike occurrence in the stage of QRS complex.
- EOG artifacts caused by blinking and eye movements.
- EMG high-frequency artifacts.

Technical artifacts are usually caused by avoidable mechanical issues either from the machine, placement of electrodes, movements or source of energy:

- Isoline drift.
- Powerline hum.
- Breathing movements artifacts.
- Pulse artifacts caused by wrong placement of electrodes.

The most common method to remove these artifacts is filtration using low-pass filters, high-pass filters, and band-pass filters, depending on the artifact itself. To

suppress the effects of drift of isoline and powerline hum the most effective is a narrow-band pass filter with linear frequency characteristic. EMG artifacts are repressed by low-pass filter with carefully chosen cut-off frequency. Eye-movements and blinking are usually suppressed by adaptive or correlation filtration [11],[12],[13].

4.2 Extraction of Feature Signal

4.2.1 Time Domain Analysis

Entropy

Entropy (entr) is always computed for one epoch from its histogram.

$$entr_{EEG} = -\sum_{j=1}^{N} \frac{n_j}{n} ln \frac{n_j}{n} \tag{4.1}$$

where n is the number of samples $y(i)$ of the measured signal y in the epoch, N the number of bins used for the calculation of the histogram and nj is the number of samples $y(i)$ which values are within the jth bin [17],[18].

The 75th percentile

The 75th percentile (prc_{75}), also known as upper or one-third quartile, is the value under which is located 75 % of the values.

$$card\{x(i)/x(i) < \text{prc}_{75}\} = \frac{75n}{100} \tag{4.2}$$

where n is the number of samples $y(i)$ of the measured signal y in the epoch and card stands for the number of elements in the set [17],[18].

The Standard Deviation

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})} \tag{4.3}$$

where n is the number of samples $x(i)$ of the measured signal x in the epoch and $\bar{x}$ represents the mean value of the signal x [17],[18]:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{4.4}$$

Skewness

Skewness is an expression of asymmetry of the probability distribution of a random variable about its mean. The values of skewness can be either negative, positive,

or undefined. The value is negative when the values on the left side are more remote than those on the right side. Thus, when the value is positive the values on the right side are more remote and when the value is zero the distribution is symmetric [10],[17],[18].

$$skewness = \frac{M3}{M2\sqrt{M2}} \tag{4.5}$$

where

$$M_k = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^k \tag{4.6}$$

Kurtosis

Similarly to skewness, kurtosis expresses the shape of the probability distribution. It measures the contribution of remote values. Kurtosis of a normal distribution is equal to 3. When the kurtosis value is smaller than 3, the contribution of the outliers is smaller than in normal distribution and this distribution is called platykurtic. On the other hand, when the value is bigger than 3, the contribution of the outliers is bigger than in normal distribution and the distribution is leptokurtic [10],[17],[18].

$$skewness = \frac{M4}{M2M2} \tag{4.7}$$

Hjorth Parameters

Hjorth parameters are three different parameters computed from epochs long 1 s or longer, using variance as a measure of signal activity:

$$var = \frac{1}{N-1}\sum_{i=1}^{n}(x_i - x)^2 \tag{4.8}$$

- Activity

$$A = var(x) \tag{4.9}$$

- Mobility

$$M_x = \frac{var(x')}{var(x)} \tag{4.10}$$

where x' stands for the rst derivative of x.

- Complexity (Form Factor)

$$FF = \frac{M(x')}{M(x)} = \frac{\frac{var(x'')}{var(x')}}{\frac{var(x')}{var(x)}} \tag{4.11}$$

where x'' stands for the second derivative of x [6],[16],[15].

4.2.2 Frequency Domain Analysis

Analysis of signal's spectrum is useful for description of signal using its frequency components. The magnitude and time shift of the frequency components provides the character of the signal and, therefore, the signal can be further classified [13].

Fourier Transform

To express the spectrum of a continuous signal the Fourier transform (FT) is commonly used. However, the definition of FT is useable only for continuous signals and not discrete signals.

$$F(\omega) = F(f(t)) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \tag{4.12}$$

where $F(\omega)$ is the signal's spectrum, $F(f(t))$ is FT of the function $f(t)$, and ω is a frequency.

The discrete signal can be still defined as a borderline case of continuous signals. Thereupon, discrete-time Fourier transform (DTFT) is used as a modification of the FT for discrete signals:

$$DTFT\{f(t)\} = F(\omega) = \sum_{n=-\infty}^{\infty} f_n e^{-j\omega nT} \tag{4.13}$$

where F are complex spectral coefficients, n is a integer index, ω is a sampling frequency, and T is sampling interval.

In real cases the signal has to have final number of discrete points in one period of spectrum. If the frequency axis is divided equally in the half and N different sections are created, then:

$$\Omega = \frac{2\Pi}{NT} \tag{4.14}$$

and

$$DTFT\{f(t)\} = \{F_k = \sum_{n=0}^{N-1} f_n e^{-jk\Omega nT}\} \tag{4.15}$$

Due to the DTFT's time consumption new algorithms began to be used, such as fast Fourier transform (FFT). FFT is not generally different in any other aspect than the speed, even though it doesn't work as well with short duration signals. As a result of these algorithms the use of FT became much wider [13],[17],[18].

Estimate of Power Spectrum

Since EEG is a stochastic signal, its process cannot be predicted. Even though the signal itself has to be available for its analysis, different methods of analysis are used than for deterministic signals. It is important to analyze the stochastic process generating the signal, more than the signal itself. Familiar features of different realizations that are predicted with certain probability help us predict the future realizations of the stochastic process and other procedures can be created to further work with signals generated by this stochastic process. For predictions of power spectrum can be used parametric and nonparametric methods [13],[17],[18].

Nonparametric Methods

Nonparametric methods are defined by usage of received data only, no models of signal generation are set up. The most common nonparametric methods are the method of periodogram and the method of correlogram.

Periodogram is the most commonly used nonparametric method and it is defined as:

$$S_{ff}(\omega) = E\{\frac{1}{N}\left|F_w(\omega)\right|^2\} \approx \frac{1}{M}\sum_{w_i=w_l}^{w_M}\frac{1}{N}\left|F_{wi}(\omega)^2\right| \tag{4.16}$$

where M is the number of realizations and N is the number of samples.

The estimate of power spectrum can be made only from one realization. Despite the fact of the possibility to estimate the power spectrum solely based on one realization, the standard deviation in this case will be quite considerable, as we suppose that only one realization sufficiently represents average characteristics of the process. The given signal of N samples is divided into K overlapping/nonoverlapping segments each of length M. The power spectrum of each segment is after the segmentation evaluated for each segment. The final spectrogram is achieved by averaging the K segments. The assumption of closeness of the neighbouring values of the power spectrum leads to estimation of new values of the power spectrum using weighted averages of the neighbouring values. Thus the power spectrum can be smoothed. If more realizations of one stochastic process are available, the standard deviation using periodogram is rather smaller.

Correlogram is a method based on the periodogram and it is based on discrete version of Wiener-Chinchil relation. The first step is the calculation of the weighted

estimation of the autocorrelation function . Furthermore DTFT is applied to the estimation of the autocorrelation function, usually using FFT algorithms.

Another method is the linear filtering method. The analyzed signal is filtrated through M bandpass filters, every filter with different frequency band, covering the whole frequency spectrum of the signal. The outlet of each filter is squared and time averaged afterwards, to provide the value of estimate of spectral density [13],[17],[18].

Parametric Methods

Parametric methods are based on creation of a model of the origin of the signal. The model itself subsequently characterizes the signal and its spectrum as well. The models reduce the data and represent the spectrum more realistically. On the other hand, these models are quite difficult to choose, along with its order [13],[17],[18].

5 SIGNAL ANALYSIS

The signals used in the study were measured during the DREAMS project. The study was carried out in order to gain data to tune, train and test out automatic sleep classifiers. The whole-night recordings had been collected from 20 healthy subjects in a belgium hospital using a digital 32-channel polygraph (BrainnetTM System of MEDATEC, Brussels, Belgium). At least two EOG signals, three EEG signals, and one EMG signals were recorded and the sampling frequency of these signals is 200 Hz. These recordings were chosen for this study for their clarity, as the subjects were healthy and free of any medication. The recordings are annoted according to AASM standards and one epoch lasts 30 seconds [19].

5.1 Statistical Analysis

To determine which obtained features of the signal are useful for the classification, the data needed to be tested. As the features describe the signal, it is required for the features to differ in between stages. Therefore the variance of the data was tested, using the Kruskal-Wallis test.

The Kruskal-Wallis test is a basic test designed to decide if sets of data originate from the same distribution. As the Kruskal-Wallis test is non-parametric, the test does not assume normal distribution. Due to this advantage the test results are less stringent which can be replaced by bigger size of the sample. The null hypothesis of the Kruskal-Wallis test is that the medians of all groups are equal. The null hypothesis is accepted if the p-value is bigger than α. α is the lowest significance level of the given test and usually is α equal to 0.05. If p is lower than α, the null hypothesis is rejected and the alternative hypothesis is accepted. In the case of the Kruskal-Wallis test the alternative hypothesis says that at least one population median of one group is different from the population median of at least one other group. This test was applied to our data using the Matlab implemented function *kruskalwallis*.

The box plot method can be further used to verify the results of the analysis of the variance and to graphically depict the data. Box plots are also non-parametric tests.

The box plot consists of main box, which represents the lower quartile - 25 % and the upper quartile - 50 %. A small dot, square, or line in the box might represent 50 % quantile - median. The box plot also might have whiskers, representing the complete range of the values [20].

5.2 Results of Statistical Analysis

To analyze the PSG data certain features were selected in the time and the frequency domain. The chosen features are: entropy, the 75th percentile, the standard deviation, skewness, kurtosis, Hjorth parameters (activity, mobility, complexity), EMG analysis, EOG analysis, analysis of occurrence of k complexes in the EEG signal and EEG power spectrum.

The features and their statistical importance is described below.

5.2.1 Time domain features

Most of the features gained in the time domain have already been described in the chapter 4.2.1. Aside from those features, computed for all the signals (EEG, EOG, EMG), two other features were obtained from EOG signals and one more from EMG signals.

First EOG time domain feature is the average distance between eye movements estimated in samples, and the second feature is the number of movements in one epoch. Both of these features are statistically significant for every pair, except for stages W-REM.

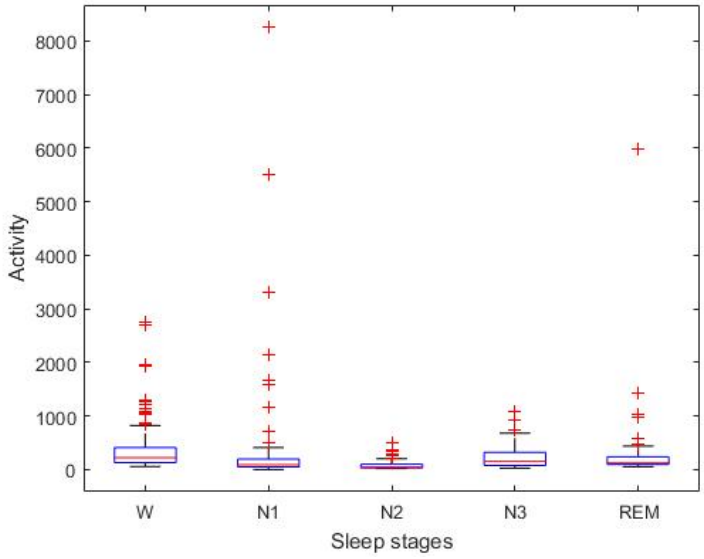

Fig. 5.1: Boxplot of EOG feature activity.

For the EMG signal analysis the signal EMG1 was chosen from the three available EMG signals. The usual features of time domain, such as entropy, the 75th percentile, the standard deviation, skewness, kurtosis and Hjorth parameters were evaluated. The average duration of one movement was estimated, but only from

movements lasting more than one half of a second. According to this it is possible to differ mainly between REM and W stages, as W is characteristic by generally higher muscle activity and REM is characteristic by complete atonia or very short (up to 1 second = 200 samples in our case) and rarely occurring (usually one per one sleep epoch) short movements.

As the most statistically significant feature appears to be Hjorth parameter activity (Fig. 5.1), showing significant results for every pair of sleep epochs.

5.2.2 Frequency domain features

One of the most important features estimated from the EEG signal is the power spectrum estimated from the frequency domain.

To estimate the power spectrum nonparametric method of Welch periodogram was used. Using this method the 30s long epoch (Fig. 5.2) is windowed, in our case by 3s long windows with 50 % overlay.

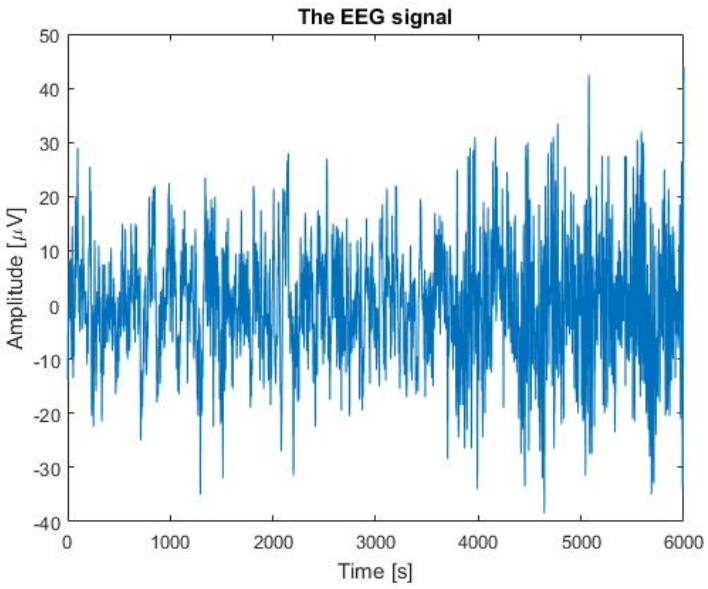

Fig. 5.2: EEG signal.

Each window was weighted by Hamming window. The result is the estimate of power spectrum density for each desired frequency (Fig. 5.3).

To analyze the power spectrum in the EEG, the representation of certain frequencies within the signal is carried out. The frequencies important for the sleep scoring, therefore also the frequencies chosen for the estimate, are α (8-12 Hz), $\beta 1$ (12-22 Hz), $\beta 2$ (22-35 Hz), δ (0.3-4 Hz), θ (4-81 Hz). The final features are the relative estimates of the power spectrum within the mentioned frequencies.

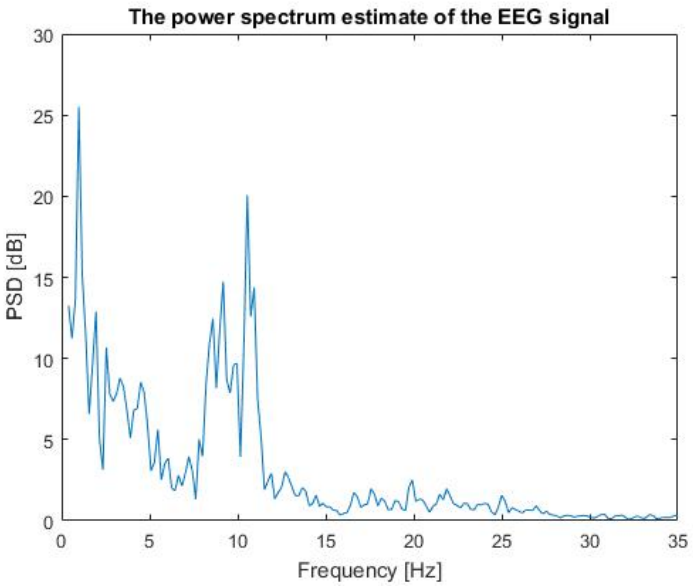

Fig. 5.3: Power spectrum of the EEG signal.

Another important feature estimated from the EEG signal is the detection of k complexes. To improve the differentiation between the stages N1 and N2 it is almost necessary to analyze the occurrence of these waves in the epoch, as they are specific for stage N2. The detection of the k complexes was in this case carried out using the Teager-Kaiser energy operator (TKEO), defined as:

$$\Psi[x(n)] = x^2(n) - x(n+1)x(n-1) \tag{5.1}$$

where x is the signal value and n is the sample number. The operator was used to smoothen the signal and to enhance the k complex in the signal.

5.3 The Selected Features

According to Kruskal-Wallis test, the number of statistically significant features is too high. Therefore the need for more thorough selection of the features arises. The selection of the appropriate features for the following sleep scoring, was based mainly on the subjective analysis of the Kruskal-Wallis test for certain of sleep stages, and also according to other articles concerning the issue of feature selection for ANN.

The feature activity, one of the Hjorth parameters, was selected due to its statistical significance for every PSG signal. Other Hjorth parameter, complexity, was selected from EOG and EMG signals only, as the statistical tests for EEG did not prove as much significance. Mobility was selected only for EMG signals as in EOG and EEG the number of insignificant pairs was one of the highest.

Tab. 5.1: The features selected for sleep scoring ANN.

Signal	**Selected Features**
EOG	Activity, complexity, entropy, kurtosis, the average distance between eye movements, the standard deviation, the 75th percentile
EEG	Activity, entropy, the detection of k complexes, the power spectrum density estimate
EMG	Activity, complexity, mobility, the duration of higher muscle activity, the standard deviation

Entropy was selected for EOG signals, as for EMG signals. The only pair of sleep stages proved statistically insignificant was W-REM, but the mentioned pair was statistically insignificant in most of the features. One of the features, where W-REM was recognized as statistically significant was the duration of higher muscle activity, also selected for EMG signals.

For EOG two other parameters, kurtosis and the average distance between eye movements, were selected, even though they are not statistically significant for W-REM.

The remaining parameters are EEG features detection of k complexes and the power spectrum density estimate. The detection of k complexes is not significant for certain pairs of sleep stages but the feature was selected as it is important for determination of N2 stage. The power spectrum estimate (for frequencies $\beta 1$) was statistically significant for every pair of sleep stages and therefore the feature was selected too.

6 SLEEP SCORING USING NEURAL NETWORKS

The features selected in the previous chapter are going to be used as an input for classification of sleep epochs according to their sleep stages. The ANN is used as the classifier, regarding the positive results of other studies on this topic. The ANN classifier was realized using Matlab 2017. The range of used sleep epochs, 4365 to 7276, differs according to combination of sleep stages that were being classified, as 1455 epochs for each sleep stage were always used for training and testing. The data was divided into the training set, consisting of 60 % of the data, 20 % of the data was used as the validating set, and the remaining 20 % was used to test the network. Three different classifications of the data were approached. The first classification differs only between the W stage, N1 stage, and REM stage, while the second alternative classifies W, N1, N2, and REM sleep stage, and the third option scores all the sleep stages.

6.1 Realization of Neural Networks in Matlab

The essential part of the neural network is the input matrix and the target matrix. The input matrix is the training set used to train the ANN and the target set is to adjust the neural network according to differences between the output and the target of the neural network. In our case the training set consists of 21 inputs, features computed for 4365-7276 sleep epochs. The input data were divided into training, validating and testing set using the function *dividerand*, in ratios 60 % training set and remaining 40 % the testing set.

The target set, necessary to train the neural network, was created as matrix with 5 columns and 4365-7276 rows, columns representing the expected sleep stage and rows representing the sleep epochs. The target matrix consists of zeroes and ones, as one is only in the column representing the correct stage marked by a sleep expert. To create the ANN itself, the function *nftool* opened the ANN generator in Matlab. The input variables, target and input, were selected along with the distribution of the training and testing set. The training function of the network is the algorithm of Levenberg-Marquardt optimization of weights and bias update. Even though the algorithm requires more memory than other algorithms, it was chosen due to its speed. To create the ANN structure (Fig. 6.1) the number of hidden neurons had to be chosen. The output was is the variable *net* with all the networks parameters in it.

The output arguments, *net* and *tr*, represent the adjusted network (*net*), the

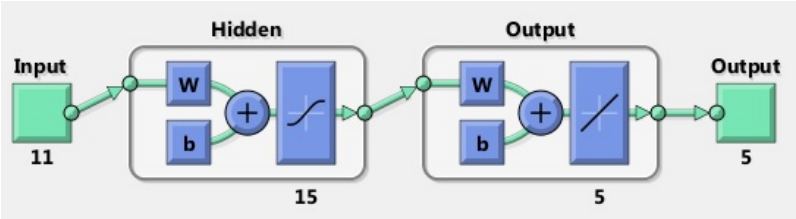

Fig. 6.1: The structure of the ANN.

number of the learning epochs (*epoch*), and the momentary performance (*perf*). The networks' performance was set using the minimum square error method.

To statistically evaluate the accuracy and performance of the ANN the confusion matrix (Tab 6.1) provides general overview of results compared to targets in percents. The rows of this matrix represent the actual class and the columns usually represent the predicted class.

Tab. 6.1: An example of simple confusion matrix

		Target class	
		1	0
Output class	1	True positive	False positive
	0	False negative	True negative

In Matlab the confusion matrix (Tab. 6.2) is plotted with function *plotconfusion*. The true positive is the number of correct classifications into class 1, false positive is the number of classifications mistakenly assigned to class 1, the false negative represents number of cases when the sleep epoch is inaccurately predicted to belong to class 0, and true negative is the number of times when sleep epoch was correctly not assigned to class 1.

In order to statistically asses the accuracy of the neural network, three important rates can be estimated - specificity, sensitivity, and accuracy of the neural network. Specificity,

$$Specificity = TN/(TN + FP), \tag{6.1}$$

measures the rate of positives that are correctly identified and on the other hand, sensitivity,

$$Sensitivity = TP/(TP + FN), \tag{6.2}$$

measures the rate of negatives correctly identified. Accuracy is, simply said, the closeness to the true values, defined as:

$$Accuracy = TP + TN/(TP + FP + TN + FN). \tag{6.3}$$

Tab. 6.2: An explanation of confusion matrix used further in the chapter

	W	Correctly classified	Incorrectly classified	Incorrectly classified	Positive predictive value False discovery rate
Output class	N1	Incorrectly classified	Correctly classified	Incorrectly classified	Positive predictive value False discovery rate
	REM	Incorrectly classified	Incorrectly classified	Correctly classified	Positive predictive value False discovery value
		True positive rate False negative rate	True positive rate False negative rate	True positive rate False negative rate	Overall accuracy Overall inaccuracy
		W	N1	REM	
		Target class			

6.2 Classification of the sleep stages W, N1, and REM

For the first classification stages W, N1, and REM have been chosen. The particular combination of classes is the combination with the highest odds to result in a highly accurate ANN. The reason behind this is that the stages N2 and N3, the stages most mistaken with the rest of the stages, are not included in the classification. The target matrix consists of three columns and 4365 rows, representing three sleep stages and 4365 analyzed epochs. The input matrix had 21 columns (inputs) and 4365 rows too and while 60 % of the data was used as the learning matrix, the other 40 % was used for testing the ANN.

The structure of the ANN was one hidden layer, containing 8-20 neurons. All these neural networks were created and the accuracy is displayed below on Fig. 6.2.

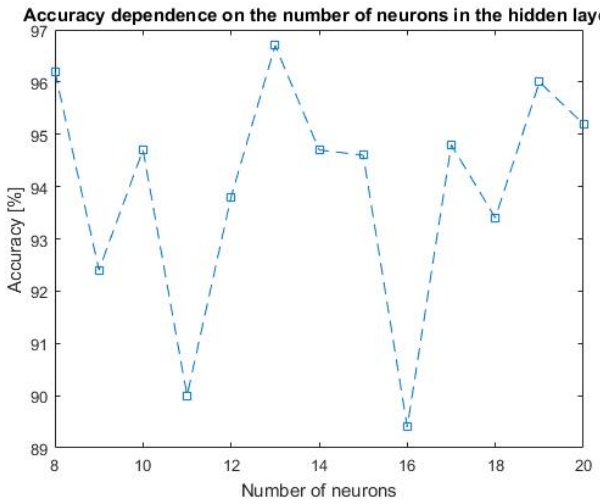

Fig. 6.2: Accuracy dependence on the number of neurons in the hidden layer for classification of W, N1, and REM.

As the most accurate ANN structure appears to be the ANN with 13 neurons in the hidden layer with accuracy of 96.7 %. The confusion matrix of the aforementioned ANN is on Fig. 6.3.

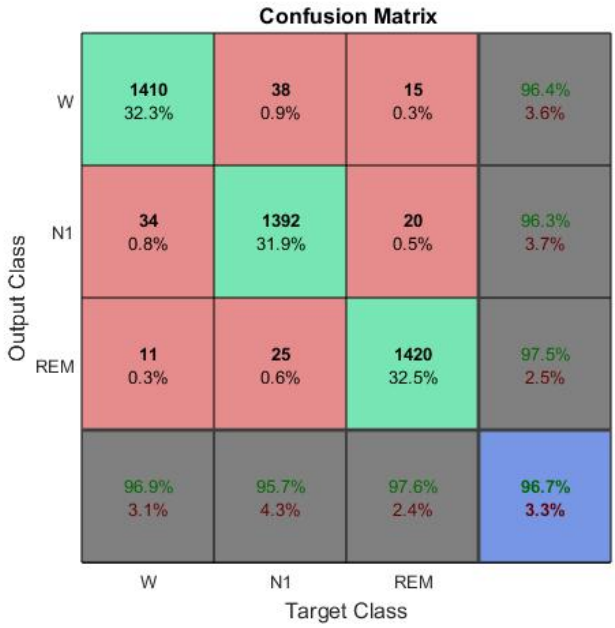

Fig. 6.3: Confusion matrix of the ANN with one hidden layer and 13 neurons.

As seen on the confusion matrix, the best accuracy has the sleep stage N1. Sleep stage REM was often confused with stage W as the stage W is the most similar to stage REM. On the other hand, stage W is most similar to stage N1 which is confirmed by the confusion matrix.

6.3 Classification of the sleep stages W, N1, N3, and REM

In the second classification sleep stage N3 was added to W, N1, and REM. It is likely that the accuracy of the network (Fig. 6.4) will be lower than in the previous section, as N1 and N3 are quite similar stages that are harder to distinguish.

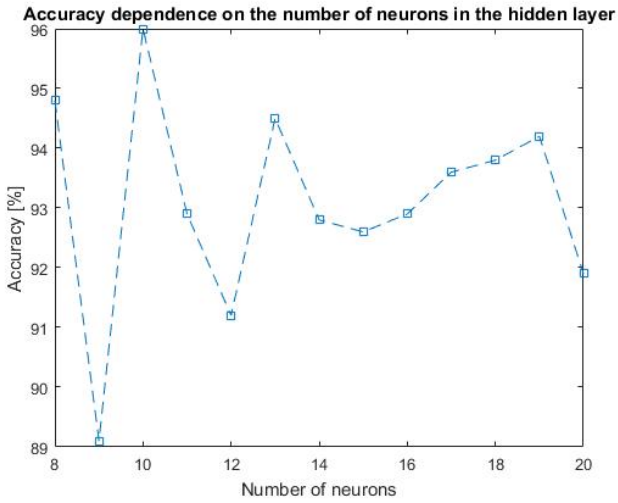

Fig. 6.4: Accuracy dependence on the number of neurons in the hidden layer for classification of W, N1, N3, and REM.

The input data consists of 21 columns and 5821 rows, therefore 21 inputs and 5821 sleep epochs. The learning set consists of 60 % of the data and 40 % of the data forms the testing set. The ANN had one hidden layer with 8-20 neurons. The dependence of the network size and the networks accuracy is displayed on the graph below.

The ANN with 10 neurons, according to the given data the best ANN structure, has the accuracy of 96.0 %. The confusion matrix (Fig. 6.5) supports the theoretical presumption that the N3 will be mostly confused with REM due to the low EEG frequencies, and N1 will be the most comparable stage to stage W.

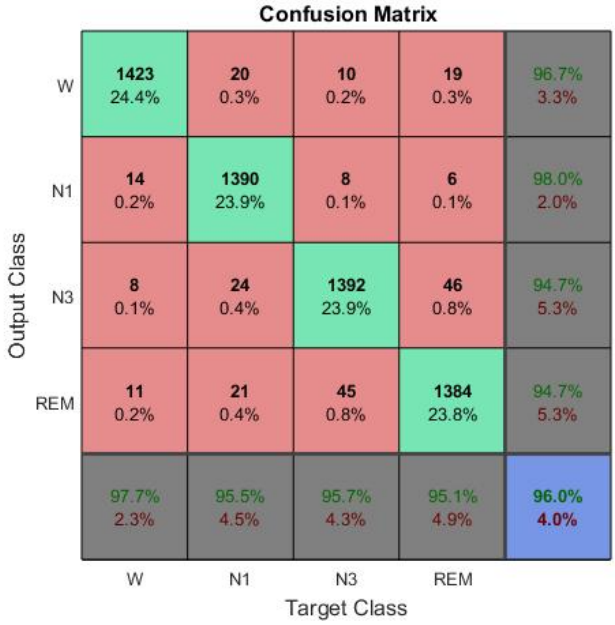

Fig. 6.5: Confusion matrix of the ANN with one hidden layer and 10 neurons.

6.4 Classification of all the sleep stages

The target matrix in this case consisted out of 5 columns representing all the sleep stages - W, N1, N2, N3, and REM stage. The input matrix has 21 columns, inputs computed for 7276 30s long sleep epochs. N stages have similar properties and it is expected for the accuracy of the ANN to be lower than in the classification of the sleep stages W, N1, N3, and REM. The testing set is made of 60 % of the input data and the remaining 40 % was used as the test matrix.

The actual ANN used for this classification is using the back-propagation algorithm described in chapter 3.1. Only one hidden layer covered 8-20 neurons. All of the layouts of neurons were tested. According to the results of accuracy, pictured for every layout below on Fig. 6.6, the ANN with one hidden layer with 15 neurons appears to be the most notable.

After regarding the confusion matrix (Fig. 6.7) for the ANN with 15 neurons, the most difficult stage to classify is the stage N2. The stage N2 was 205 times marked as stage REM, while only 4 times marked as stage N1.

Another notable ANN layouts are ANNs with 13 and 20 neurons. Also this ANNs seem to have problem with distinguishing sleep stage N2 (Fig. 6.8, Fig. 6.9).

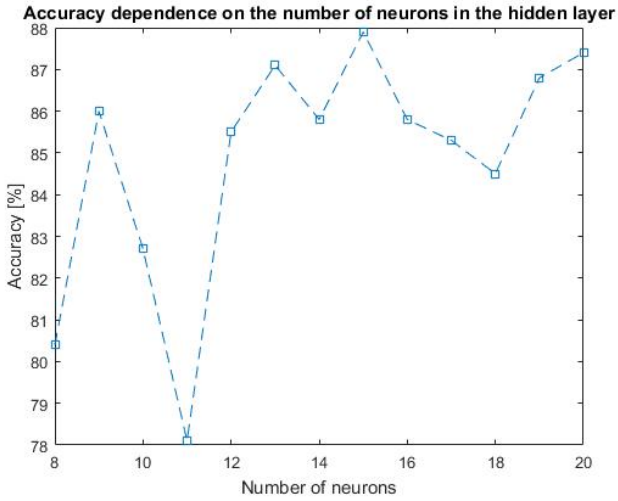

Fig. 6.6: Accuracy dependence on the number of neurons in the hidden layer for classification of W, N1, N2, N3, and REM.

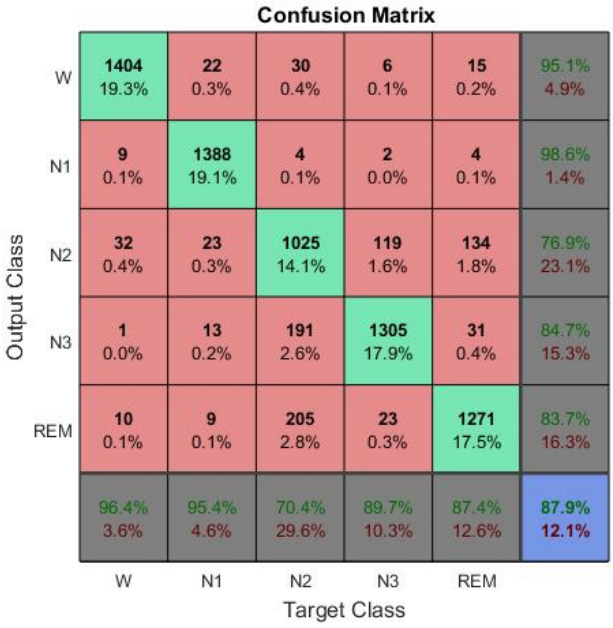

Fig. 6.7: Confusion matrix of the ANN with one hidden layer and 15 neurons.

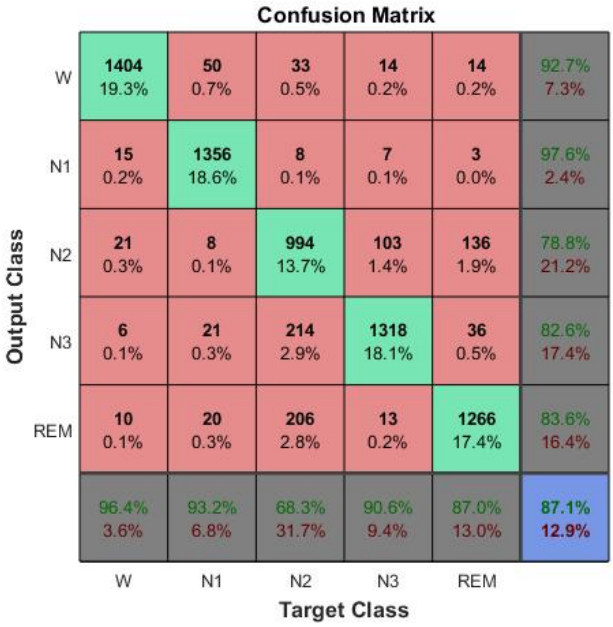

Fig. 6.8: Confusion matrix of the ANN with one hidden layer and 13 neurons.

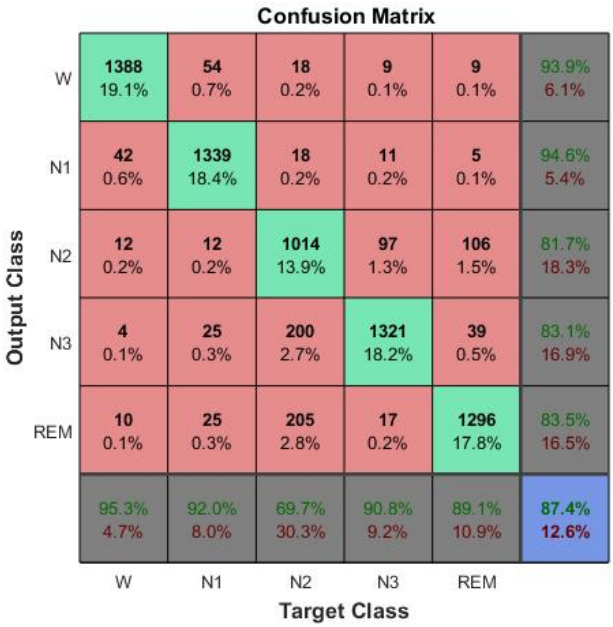

Fig. 6.9: Confusion matrix of the ANN with one hidden layer and 20 neurons.

6.5 Comparison with other publications

To thoroughly evaluate the results of the classification, the results (Fig. 6.7) are further compared with other professional publications on Tab. 6.3.

Tab. 6.3: The overview of the results for different publications.

Author	**Year of publication**	**Classified stages**	**Method of classification**	**Accuracy**
Zoubek [14]	2007	W, S1, S2, S3, S4, REM	ANN	80 % 84.75 % (W) 64.56 % (S1) 85.55 % (S2) 92.9 % (S3-S4) 72.81 % (REM)
Schaltenbrand [21]	1996	W, S1, S2, S3, S4, REM	ANN	82.34 % 86.84 % (W) 21.37 % (S1) 87 % (S2) 58.13 % (S3) 80.13 % (S4) 83.57 % (REM)
Oropensa [22]	1999	W, S1, S2, REM	Wavelet transform, ANN	97.5 % 97 % (W) 98 % (S1) 99 % (S2) 95 % (REM)
Lajnef [10]	2015	Awake, S1, S2, SWS, REM	Multi-class decision based SVM framework (DVSM)	90.55 % 87 % (Awake) 82 % (S1) 85 % (S2) 88 % (SWS) 98 % (REM)
Agarwal [23]	2001	W, S1, S2, S3, S4, REM	Clustering	76.8 %

7 CONCLUSION

The main focus of the bachelor's thesis is to create theoretical overview of disciplines associated with automatic sleep scoring.

The first chapter describes sleep, sleep stages and cycles, and sleep disorders caused by irregularities of the sleep cycle. Furthermore, in the second chapter, PSG, method of sleep recording and the practice of recording the most valuable signals are described as well. Various methods of classification most useful for sleep scoring are mentioned in the third chapter of the thesis. Signal processing and extraction of features needed as an input of the chosen classifier are analyzed in the fourth chapter. The fifth chapter of the thesis discusses the design of the sleep classifier itself and summarizes information from previously mentioned chapters.

Based on this analysis it was possible to further proceed to practical part of the bachelor's thesis. The practical part is centred on the signal processing, as this part is the most fundamental in the sleep-stage classification, being most influential to the classification results. Furthermore, suitable features were used as in input to an ANN, described more thoroughly in the chapter 3.1. The features were chosen according to the results of Kruskal-Wallis test. Crucial part of the thesis was evaluation of obtained results of the classification using ANN with back-propagation algorithm. The effectiveness of the classification method chosen by us can be determined according to results of statistical tests of our method compared to others and to the sleep stages set by sleep experts.

Based on the results of the classification (Chapter 6) the features proposed as appropriate for an input, and the structure of neural network too, seem to be successful for sleep scoring.

To improve the results, especially the results of N2 scoring, additional features might have been computed and included as the input of the ANN, such as sleep spindles detection, which are typical for stage N2. Another proposal to improve the accuracy is to choose different learning algorithm or to add hidden layers to the neural network.

BIBLIOGRAPHY

[1] Editor Alon Y. AVIDAN, editor Teri J. BARKOUKIS. *Review of Sleep Medicine.3rd Ed.* Philadelphia: Elsevier Saunders, 2011. ISBN 9781455703197.

[2] SMITH, Harold R., COMELLA, Cynthia L., HOGL, Birgit. *Sleep Medicine.* New York: Cambridge University Press, 2008. ISBN 9780511545085.

[3] Ebrahimi, F., Mikaeili, M., Estrada, E., & Nazeran, H. (2008). *Automatic sleep stage classification based on EEG signals by using neural networks and wavelet packet coefficients.* 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, (May 2014), 1151–1154. https://doi.org/10.1109/IEMBS.2008.4649365

[4] IBER, Conrad. *The AASM manual for the scoring of sleep and associated events: rules,terminology, and technical specifications.* Westchester, IL: American Academy of Sleep Medicine, c2007.

[5] KANIUSAS, Eugenijus (2012). *Biomedical Signals and Sensors I, Linking Physiological Phenomena and Biosignals.* Springer-Verlag Berlin Heidelberg, 2012. ISBN 978-3-642-24843-6.

[6] RANGAYYAN, Rangaraj M. *Biomedical signal analysis. Second edition.* Hoboken, New Jersey: John Wiley & Sons, 2015. ISBN 978-0-470-91139-6.

[7] Becq, G., Charbonnier, S., Chapotot, F., Buguet, A., Bourdon, L., & Baconnier, P. (2005). *Comparison Between Five Classifiers for Automatic Scoring of Human Sleep Recordings.* Classification and Clustering for Knowledge Discovery, 127, 113–127. https://doi.org/10.1007/11011620_8

[8] DOUGHERTY, Geoff. *Pattern recognition and classification: an introduction.* New York: Springer Science+Business Media, 2012. ISBN 9781461453222.

[9] Farooq, H., & Jain, A. (2017). *A Novel Computer Assisted Automatic Sleep Scoring System by Employing Artificial Neural Network – A review, 5(3), 43–47.* http://doi.org/10.11648/j.ijssn.20170503.12

[10] Lajnef, T., Chaibi, S., Ruby, P., Aguera, P. E., Eichenlaub, J. B., Samet, M., ... Jerbi, K. (2015). *Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines.* Journal of Neuroscience Methods, 250, 94–105. http://doi.org/10.1016/j.jneumeth.2015.01.022

[11] Bruce, E. N. (2001). *Biomedical Signal Processing and Signal Modeling.* New York: Wiley, c2001. Wiley series in telecommunications and signal processing. ISBN 0-471-34540-7.

[12] KOZUMPLÍK, J. *Č: Zpracování biologických signálů. Přednáška.* Brno: VUT v Brně, Fakulta elektrotechniky a komunikačních technologií, 2017

[13] JAN, Jiří. *Číslicová filtrace, analýza a restaurace signálů. 2. upr. a rozš.* vyd. Brno: VUTIUM, 2002. ISBN 80-214-1558-4.

[14] Zoubek, LUKÁŠ, Charbonnier, SYLVIE, Lesecqb, SUZANNE, Buguete, ALAIN, Chapototcd, FLORIAN. Feature selection for sleep/wake stages classification using data driven methods *Biomedical Signal Processing and Control. Volume 2, Issue 3, July 2007, Pages 171-179.* DOI:10.1.1.576.7501. Available at:http://www.sciencedirect.com/science/article/pii/S174680940700036

[15] Estrada, E., Nazeran, H., Ebrahimi, F., & Mikaeili, M. (2009). Feature selection for sleep/wake stages classification using data driven methods *EEG signal features for computer-aided sleep stage detection.* 2009 4th International IEEE/EMBS Conference on Neural Engineering, NER '09, (June), 669–672. https://doi.org/10.1109/NER.2009.5109385

[16] Simões, H., Pires, G., Nunes, U., & Silva, V. (2010). *Feature Extraction and Selection for Automatic Sleep Staging Using EEG. Icinco, 128–133.*

[17] KUTZ, Myer. *Standard handbook of biomedical engineering and design.* New York: McGraw-Hill, c2003. ISBN 0071356371.

[18] FREEMAN, Walter J. and Rodrigo QUIAN QUIROGA. *Imaging brain function with EEG: advanced temporal and spatial analysis of electroencephalographic signals.* New York: Springer, 2013. ISBN 978-1-4614-4983-6.

[19] *http://www.tcts.fpms.ac.be/ devuyst/Databases/DatabaseSubjects/*

[20] PAVLÍK, Tomáš a Ladislav DUŠEK. *Biostatistika.* Vyd. 1. Brno: Akademické nakladatelství CERM, 2012, 131 s. ISBN 978-80-7204-782-6.

[21] Schaltenbrand, N., Toussaint, M., Luthringer, R. (1996). *Automatic Analysis and Sleep Scoring Sleep Stage Scoring Using the Neural Network Model : Comparison Between Visual and Automatic Analysis In Normal Subjects and Patients, 19(1), 26–35.*

[22] Oropesa, E., Cycon, H., Jobert, M. (1999). *Sleep stage classification using wavelet transform and neural network.* International Computer Science Institute, (510).

[23] Agarwal, R., Gotman, J. (2001). *Computer-assisted sleep staging. IEEE Transactions on Biomedical Engineering, 48(12), 1412–1423.* https://doi.org/10.1109/10.966600

LIST OF SYMBOLS, PHYSICAL CONSTANTS AND ABBREVIATIONS

AASM	American Academy of Sleep Medicine
REM	Rapid Eye Movements
NREM	Non-rapid Eye Movements
EEG	electroencephalogram
EOG	electrooculogram
EMG	electromyogram
ECG	electrocardiogram
MBM	Major Body Movement
PSG	polysomnography
entr	Entropy
prc_{75}	75th percentile
std	standart deviation
FT	Fourier transform
DTFT	discrete-time Fourier transform
FFT	fast Fourier transform
STFT	short-time Fourier transform
WT	Wavelet transform
CWT	Continuous wavelet transform
DWT	discrete wavelet transform
ANN	Artificial neural networks
kNN	k-nearest neighbor
SSE	sum of the squared error
SVM	support vector machines