Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Masterarbeit

Kaled Hamodi

Biometrical Research on Air Traffic Controllers‘ Workload

Fakultät Technik und Informatik
Department Fahrzeugtechnik und Flugzeugbau

Faculty of Engineering and Computer Science
Department of Automotive and Aeronautical Engineering

Kaled Hamodi

Biometrical Research on Air Traffic Controllers' Workload

Masterarbeit eingereicht im Rahmen der Masterprüfung

im Studiengang Flugzeugbau
am Department Fahrzeugtechnik und Flugzeugbau
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

in Zusammenarbeit mit:

Universidad Politécnica de Madrid (UPM)
Escuela Técnica Superior de Ingeniería Aeronáutica y del Espacio (ETSIAE)
Plaza Cardenal Cisneros 3
28040 Madrid (Spanien)

Erstprüfer: Prof. Dr.-Ing. Alexander Piskun, M.Sc.
Zweitprüferin : Prof. Dr.-Ing. Rosa María Arnaldo Valdés

Abgabedatum: 02. August 2018

Zusammenfassung

Name des Studierenden

Kaled Hamodi

Thema der Masterthesis

Biometrische Untersuchung der Arbeitslast von Fluglotsen

Stichworte

Neurologie, Elektroenzephalographie, Biometrie, Arbeitslast, Parameter, Charakteristik, Algorithmus

Kurzzusammenfassung

Diese Arbeit befasst sich mit der Fragestellung, inwiefern kostengünstige biometrische Systeme nützlich sind, um die Arbeitsbelastung von Fluglotsen festzustellen. Hierzu wurden Gerätschaften beschafft, um Parameter zu sammeln und Charakteristiken festzustellen. Diese werden, mit der Absicht Algorithmen zu bilden, mit Hilfe von Berechnungen analysiert und abgebildet. Die erfassten Ergebnisse sollen für die zukünftige Nutzung und Weiterentwicklung dienen.

Name of Student

Kaled Hamodi

Title of the paper

Biometrical Research on Air Traffic Controllers' Workload

Keywords

Neurology, Electroencephalography, Biometrics, Workload, Parameters, Characteristics, Algorithm

Abstract

This paper addresses the question of how low cost biometrical systems are useful for determining the workload of air traffic controllers. For this purpose, equipment has been procured to collect parameters and determine characteristics. These are analysed and displayed with the intention of developing algorithms. The recorded results should serve for future use and further development

Acknowledgements

My explicit acknowledgements have to be directed to many people. Attention has to be drawn to the support in form and content of this thesis of Prof. Dr.-Ing. Rosa María Arnaldo Valdés and Prof. Dr.-Ing. Alexander Piskun, M.Sc. Moreover, I want to thank Fanchao Meng for the practical assistance during the research. Special thanks have to be directed to Prof. Dr.-Ing. Rocío Barragán Montes, Prof. Dr.-Ing. Victor Fernando Gomez Comendador and Julián Martin Pérez for the organization and realization of numerous formal aspects and experiments among the air traffic controllers. Enabling all experiments, I want to thank SkySoft and Emotiv for providing program licenses and enabling this research. Furthermore, the allocation of the workplace was realized with the help of the research environment of CRIDA but especially with the support of Nicholas Suarez Tetzlaff. Regarding the practical results, I want to thank all attendees for supporting this research, but especially to the professional air traffic controllers. Many thanks have to be directed to my dear friend Charles Whitehead for assuring a high standard of expression of this thesis. In addition, I want to express my special thanks to my remarkable Spanish teacher Carmén Peréz García, all of my Spanish friends at the research environment, the student exchange coordinators Aurora García, Hella Rieger, as well as Prof. Dr.-Ing. Gustavo Alonso Rodrigo and Prof. Dr.-Ing. Peter Seyfried.

Besides all mentioned protagonists, I want to thank my family for the extraordinary support they afford.

Table of Contents

Figures

Tables

List of Abbreviations

FAA	Federal Aviation Administration	ACC	Aircraft	EN	European Norm
ATC	Air Traffic Control	TU	Tupolev	IEC	International Electronical Commission
ATM	Air Traffic Management	DHL	Dalsey Hillblom Lynn	AS/NZS	Australian Standards/ New Zealand Standards
SESAR	Single European Sky ATM Research	FL	Flight level	CISPR	Comité International Spécial des Perturbations Radioélectriques
USA	United States of America	EEG	Electroence-phalography	ICA	Independent Component Analysis
DC-7	Douglas Aircraft Company Model 7	PSD	Power Spectral Density	GUI	Graphic User Interface
UAL	United Airlines	DIN	Deutsche Industrie Norm	ML	Machine Learning
SVM	Support Vector Machine	V	Volt	TWA	Trans World Airlines
2D	2 Dimensional	Hz	Hertz	ISO	International Standardisation Organisation
3D	3 Dimensional	dB	Decibel	AUC	Area Under Curve
AF	Anterior Frontal	CMS	Common Mode Sense	txt	Text
F	Frontal	DRL	Driven Right Leg	ASCII	American Standard Code for Information Interchange
FC	Frontal Central	SPS	Sequence Parameter Set	LEMD	Madrid-Barajas ICAO-Code
T	Temporal	SD	Standard Deviation	MAD	Madrid-Barajas IATA-Code
P	Parietial	FIR	Finite Impulse Response	ROC	Receiver Operator Characteristic Graph
O	Occipital	NASA	National Aeronautics and Space Administration	TLX	Task Load Index
ced	JVC Camera Video Data File	CRIDA	Centro de Referencia de Investigación, Desarrollo e Innovacion ATM		

Table 1: List of abbreviations

Deutsche Kurzbeschreibung

Im Rahmen der Untersuchung der Auswirkungen der Arbeitsbelastung auf die menschliche Leistungsfähigkeit wurden viele wichtige Indikatoren für die Interpretation erstellt. Daher sind in den letzten Jahren verschiedene Methoden und Möglichkeiten entwickelt worden. Heutzutage sind medizinische und andere Bereiche aller Art in der Lage, die Vorteile der Erkennung von mentalen Zuständen, wie Erregung, Wachsamkeit, Erschöpfung, mentaler Belastung und deren Ursprung, zu nutzen. Dies ist besonders im Transportwesen immer wichtiger, da Luftfahrt und die Automobilindustrie mehr Aufmerksamkeit auf den mentalen Zustand und die kognitiven Fähigkeiten von Fahrzeugführern und Piloten richten. Dies führt zu einer Verbesserung der verwendeten Systeme in Bezug auf Benutzerfreundlichkeit, Informationswert, Messprozeduren, sowie deren verwendeten Algorithmen. Das Ziel dieser Arbeit ist es, die Verwendbarkeit der Messung von mentalen Zuständen mithilfe von kostengünstigen Systemen durch die Auswertung signifikanter Metriken zu veranschaulichen.

Abstract

Within the research on the impact of workload on human performance a lot of significant indicators for the interpretation have been produced. Therefore different methods have been created in recent years and many new possibilities have been developed. Nowadays, medical and other areas of all kinds are capable of using the benefits of detecting mental states such as arousal, vigilance, fatigue, mental load and their origins. This is especially prominent in transportation, as aviation and automotive industry pay more attention to the mental state and cognitive loads of vehicle drivers and pilots. This leads to the improvement of the systems used in terms of usability, informational value, measurement accuracy, as well as their used algorithms. The aim of this thesis is to illustrate the usability of measuring cognitive states and the methods in biometrical research using low cost biometrics by evaluating recorded metrics.

1. Introduction

1.1 Preliminaries

A common problem in aviation is the occurrence of interfering human performance decrease among air traffic controllers due to excessive workload. According to researchers of the Boeing Company, EUROCONTROL and the US Federal Aviation Administration (FAA), Air Traffic Control (ATC) has a significant influence in air traffic accidents [1]. Therefore a consciousness in the authorities of civil aviation as well as companies has risen in the past years. Resulting from an increased traffic volume or appearance of emergency situations, high workload has the potential to lead to processing errors. Air traffic sectors became more complicated as a result of the named increase of traffic volume, but particularly as a consequence of low manpower. As described in [1], air traffic accidents of particular importance are the mid-air collisions of two aircrafts (Figure 1) above the Grand Canyon, USA in 1956 and near Zurich, Switzerland in 2002. Resulting from several factors, a shortage of manpower and therefore an increased workload per controller led to those accidents.

Date	Aircraft and accident outline	Major cause
1956.6	In the airspace over Grand Canyon, the U.S, DC-7 aircraft of UAL and L-1049 aircraft of TWA (both flying under IFR) had a mid-air collision at 20,000 feet, causing death of all 128 passengers.	Air traffic congestion Shortage of controlling facility Shortage of ATC manpower Insufficient delivery of Traffic information
2002.7	While controlled by the ACC of Zurich, Switzerland, TU-154 aircraft of Russian Bashkirian Airlines and B757 cargo aircraft of the U.S. DHL were flying on a collision course at the same altitude (FL360). Both airplanes descended to avoid each other, then the Bashkirian aircraft collided at a right angle with the Boeing cargo aircraft at FL354, killing all 71 passengers.	ATC instruction error RADAR malfunction (Short Term Conflict Alert) Route congestion Shortage of ATC manpower

Figure 1: ‘Major airplane accidents related to ATC human factor’ Source: [1]

The concept of improving the Air Traffic Management (ATM) by the European Union is operated under the abbreviation ‘SESAR’ financed by the H2020 European Union research program. The term ‘Single European Sky ATM Research’ ‘is the technological pillar of the Single European Sky. It aims to improve Air Traffic Management performance by modernizing and harmonizing ATM systems through the definition, development, validation and deployment of innovative technological and operational ATM solutions. These innovative solutions constitute what is known as the SESAR concept of operations.’ [2]

Many different systems and procedures have been developed to support air traffic controllers. On the one hand those systems have operational properties to support the ATC procedures and the traffic regulations. On the other hand biometrical systems will soon arise in the ATC domain due to the realized benefit. Biometrics are widely used to detect cognitive states resulting from increased workload and therefore an increase, e.g. in brain activity. Those advanced systems allow detecting noticeable changings in human behavior. Hence, human errors can be prohibited in an early stage of demanding situations. Consequently, the safety of air traffic has the potential to improve using withstanding systems. [2; 3]

Among biometrics, many methods showed their respective merits in the past century. A common represented method to detect cognitive states is Electroencephalography (EEG). Based on the measurement of differences between single brain potentials, it allows us to locate and determine brain activities, especially in its responses to certain situational changes. [4]

1.1.1 Task

The task of this research is to use experimental, statistical and analytical methods to identify the usability of low cost EEG in terms of workload detection.

1.1.2 Aim of the Research

According to the foresaid conflicts between rising traffic volume, workload and its detection using biometrics, the aim for this scientific work is to explore if this research is able:

- to quantify how well the subjects (General test persons and ATC) and therefore the measurement device can accomplish with a new task or with high workloads
- to compare the performances of ATC, in terms of cognitive resources
- to significantly differentiate three workload levels
- to determine how the subjective features used for the workload evaluation remain stable in terms of low-cost EEG

1.2 Structure of the Research

The structure of this research is divided into four significant parts:

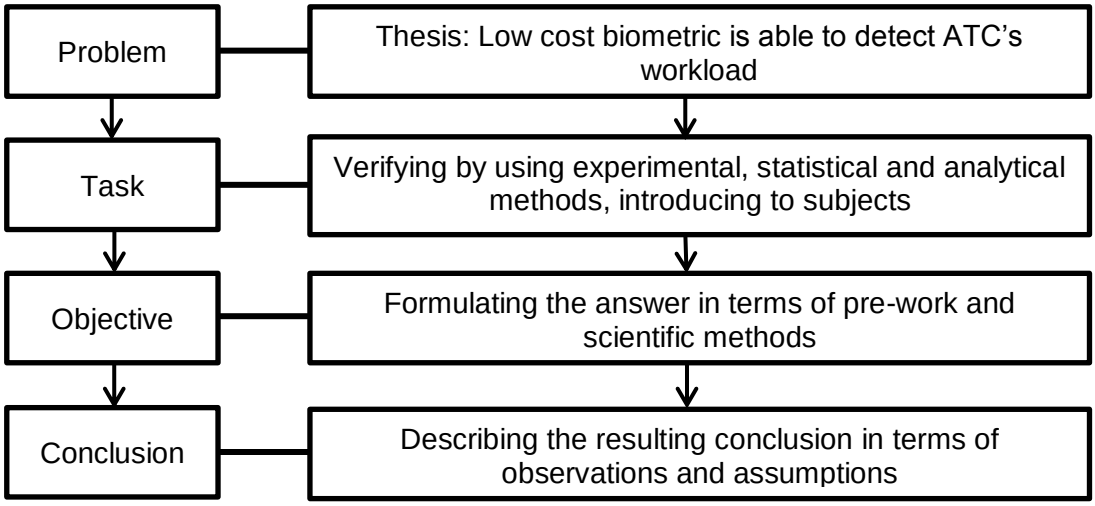


Figure 2: Research structure. Source: Own research

A more detailed explanation is provided in Figure 3:

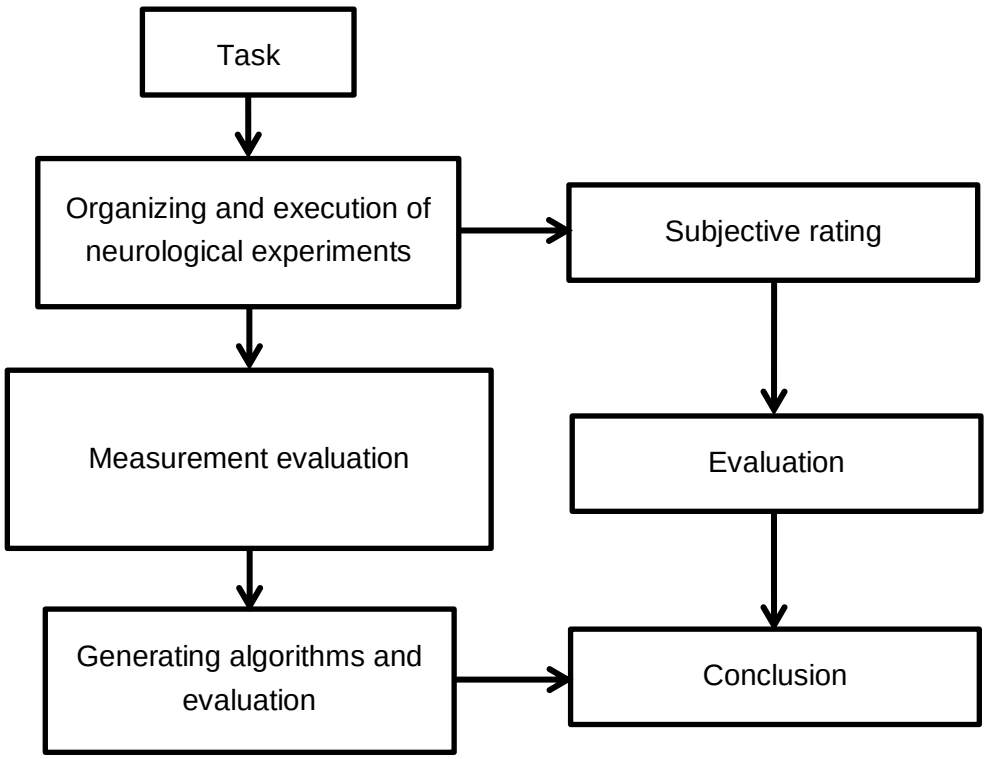


Figure 3: Research task (Step 2 - 4 according to Figure 2). Source: Own research

1.3 Requirements of the Research

A reasonable approach in compliance with the standardization norm DIN EN ISO 10075 contains requirements for this research to explore whether the named scientific methods can:

- collect significant data by representative subjects
    - in their usual working environment containing a computer, screen, keyboard and headset
    - assuring their full awareness of all senses
    - using realistic ATC software
- provide stable and significant data concerning a minimum of noise and impedance in the signal
    - due to unresisting measuring components with no galvanic distortions
    - removing unnecessary distortion such as muscle and eye movements and electrical noise by the net
    - considering procedures according to all referring standardization specifications such as medical or clinical norms
    - estimation of the amount of necessary experiments in terms of quality and informational value
    - generating a reasonable signal quality of at least 70% by preparation of the measuring components with conductive liquid
    - assuring that all measuring components are working for an equal signal distribution
    - advising the subject to follow the procedure according to the preliminary design and execution of the experiment in a laboratory environment
- evaluate and process the data in terms of structured and anomaly detection
    - using MATLAB numerical software with EEGLAB tool box

- predict characteristics according to the aim of the research using Microsoft Azure machine learning software
    - causing demanding situations in the run-up of biometrical states
    - providing valuable information of biometrical states with 90% predictability
    - creating algorithm procedure for biometrical use with predicting properties

1.4 Objective of the research

Applying all defined criteria for the aim of the research, this scientific elaboration shall clarify the aspects of usability of the introduced device to the target group, fulfilling the pre – named requirements.

2. State of the Art in Neuroscience

Invented in 1924 by Hans Berger, the EEG technology allows measuring the characteristics of potential changing based on electrochemical processes. As a non-invasive method, it is fast, simple and precise in comparison to other methods, as it has a resolution of a single millisecond or even less. EEG has a variety of different purposes. Besides quantifying different electrical activities in the brain, it is used to determine a person's workload. As a result of electro-chemical processes in the brain, EEG systems are able to measure extracellular electrical potentials of thousands of brain cells using applied electrodes to the scalp. The potential difference between only two cells is too small to be detected by those devices due to resistance by the scull and scalp [4; 5]. During the past 90 years many methods for analysing EEG have been developed. Using the Fourier Transformation of the EEG signal Power Spectral Density (PSD) leads to the 'strength of variation energy as a function of frequency' concerning cognitive changes [6; 7]. All those applications for neuroimaging allow quantifying characteristics within the EEG and visualizing changings of the brain activity. The EEG of different frequencies generates different waveforms. Besides the frequency, different waveforms differ in the amplitude and shape. Those frequencies are divided by the values shown in Figure 4. [4]

The four basic wave forms are as follows:

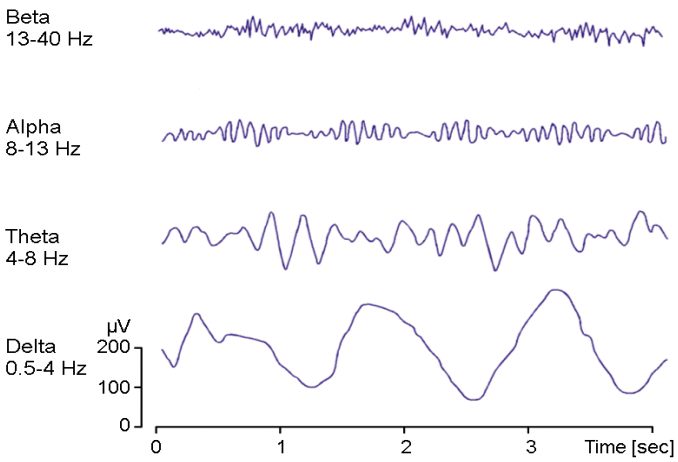


Figure 4: 'Four typical dominant brain rhythms'. Source: [4]

Among EEG systems low-cost alternatives allow to enable usage of this technology for private or non-clinical purposes. Providing two alternatives the research institute makes two low-cost devices available. As [8] concludes a disability of distinguishing workload levels by the 5-channel Emotiv Insight device, the more applicable 16-channel EPOC+ device is chosen for this research. A detailed description of the device will follow in chapter 5.

2.1 Neurological Terminology

Among neurological terminology, important wordings are present. Mental conditions are generally described as mental states relating to functionality and internal contentment. Those mental states are strongly connected to several feelings sensed by humans. [9] refers to a myriad of mental states, which include cognitive state. Related to that, [10] describes cognition as ‘the processing of information gathered by the body's sensory organs, including the processing of knowledge’. Therefore cognitive states conclude to a relation with workload, as it affects human behavior and provokes changes in the mental condition. Furthermore, workload defines the amount of activities to resolve while cognitive load characterizes the amount of the ability rearranging information. As an interconnection, mental effort describes the level and quality of task response according to the task complexity. Moreover, it defines the amount of cognitive resources that have to be devoted to obtaining relevant outputs from processing inputs [10]. In order to understand the behavior of metrics, an explanation follows.

2.2 Neurological Metrics

Different workload states are detectable using EEG. As workload changes are dependent of the individual action and reaction as well as the environmental changes according to [11; 12; 13; 14], the brain activity changes compulsory.

According to several studies, EEG is able to detect different neural activities across the brain. Besides, [15] and [16] hypotheses an increasing workload induces theta (θ: over frontal and occipital sites) and alpha (α: over parietal) activities. [4] refers that an ‘increasing workload leads to an increase in theta-frequency band activity (synchronization) over frontal-midline electrodes and a decrease in alpha-frequency band power (desynchronization) over parietal-occipital electrodes’. Furthermore, alpha and theta behavior indicate other mental states such as fatigue and mental effort according to [17]. It is known that increasing alpha frequency indicates resting states [17] and ‘transitions from single to dual task’ performance results in decreasing theta activity in accordance with [18]. Considering biometrical aspects in [19], it is known that ‘Amplitude distributions and power spectral density functions of the human operator outputs may be regarded as indications of effort expended.’ As metrics are an essential part of this research the structure will be explained in the following chapter.

2.3 Experimental Design and Theory

This part involves the holistic description of the task. It explains how the experiment needs to be processed from the idea to the final obtaining of the data, as defined in chapter 1.2. According to DIN EN ISO 10075-3 an experiment can be physiological, subjective scaled, performance measured or task and procedure analysed. In compliance with workload detection the performance measurement is chosen in Figure 5.

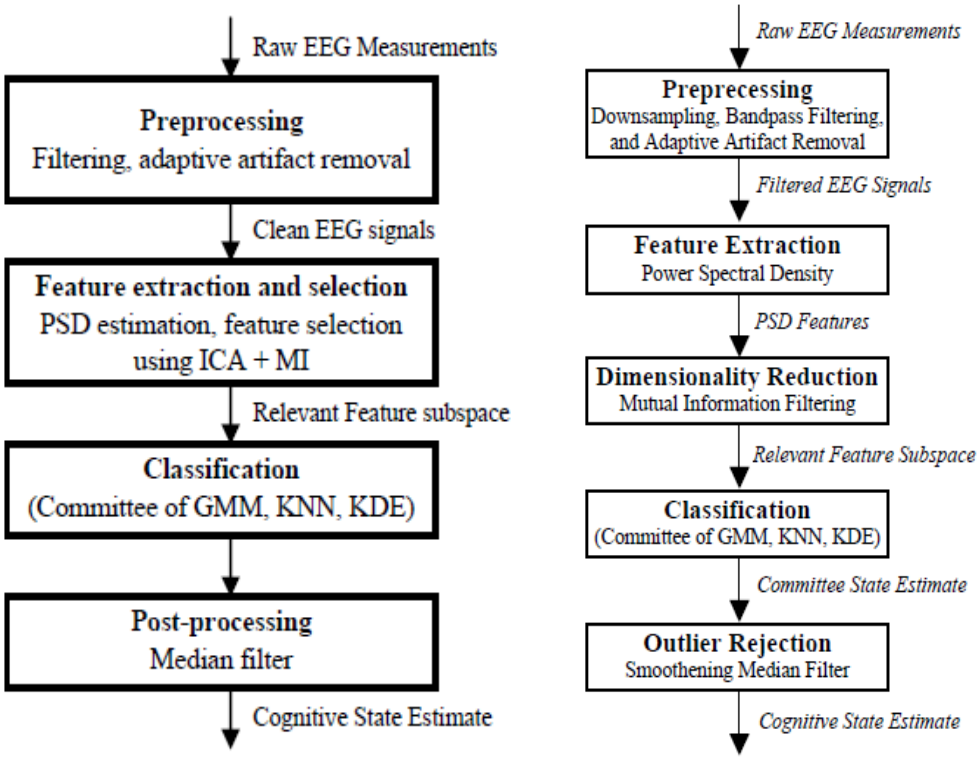


Figure 5: Typical procedure sequence of the EEG in two experiments. Source: [21; 22; 24]

DIN EN ISO 10075-3 names the proposed setup to be probabilistic due to the necessary interconnection between measured results and medical observations. Before the experimental setup subjective questions of equipment, experimental sequence and environment, definition of the workload questionnaire, but also objective aspects such as feasibility and usability of the results have to be answered. Moreover, [20] describes in the realized experiment a preliminary test period for detecting certain issues concerning the software or hardware, as well as optimization of the experimental duration.

The experimental design for EEG contains eight crucial points according to [20] and [21]:

- Definition of the equipment distinguishing between amount of measurement components and type of used instruments
- Definition of the experiment considering duration, amount and sequence of the realized procedures
- Possibility of a prior experiment to detect issues on hardware and software
- Adjusting of the equipment and subsequently following calibration
- Following of customized procedures by the subjects holding a defined attitude
- Planned brakes to avoid additional fatigue with subsequent re-calibration
- Rating of the experience on own workload profile
- Evaluating of the results

Achieving results requires an evaluative process introduced as follows.

2.3.1 EEG Evaluation

As described in [21; 22; 23; 24] and much other literature, EEG contains a certain procedure for extracting necessary information out of electric activities of the brain. All sources name four stages of getting from the measurement to the cognitive state. Those steps, as shown in Figure 5, contain the measurement of the raw EEG metrics, pre-processing of the measured values with filtering, down sampling and adaptive artifact removal, e.g. eye-blinks, to generate clean EEG signals. Moreover, feature extraction and selection of e.g. Power Spectral Density (PSD) figures have to be done. A classification of those features according to a scheme containing different classifiers as probability-related statistical functions need to be taken into account concluding post-processing of the detected values [21; 22]. Examination of those methodological aspects provides estimated cognitive states as a result of cognitive load factors.

2.3.2 Data Pre-processing

To process raw EEG signal distortions, artifacts have to be removed from the signal. According to [23] and [25] pre-processing consists of 6 crucial steps for a clean EEG signal. Essential steps are the acquisition of the signal, removal of artifacts, signal averaging, defining of the threshold of the output signal, amplification of the resulting signal and edge detection.

2.3.3 Feature Extraction

The basic reason for feature extraction is to narrow the amount of relevant data of the EEG. Three different sources from which the feature extraction is possible are:

- Spatial information: Describes the location of the signal by referring to single EEG channels

- Spectral (frequently) information: This approach names the variability of the power in the frequency bands

- Temporal information: describes the variability of a signal as function of the time

Those three sources are, according [26], known as the most common and most used among the feature extraction domain. [27] describes a variety of features with two common types, summarized as follows:

- Amplitude values of EEG signals

- Power Spectral Density (PSD)

$$S_{ff}(\omega) = 2\int_0^{\infty} r_{ff}(t)\cos\omega t\, dt \qquad 1.1$$

$$S_{ff}(\omega) = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} f_T(r)\int_{-\infty}^{\infty} F_T(x)e^{-j\omega(x-r)}dxdr \qquad 1.2$$

As the PSD uses the Fast Fourier Transformation (FFT) to transform the signal from the time (Equation 1.1) into the frequency domain (Equation 1.2), it estimates the signal variance (energy) as a function of the frequency. Visualized in a graph, the area in a distinctive range under the plotted curve defines the distributed power. Therefore, the dimension of PSD is a density shown as power/frequency defined as dB/Hz (Decibel/Hertz) or V^2/Hz (Volt²/Hertz). For estimating the power spectrum, diverse methods can be applied, as named in [23].

2.3.4 Classification

Different types of classifiers and its properties are described in the following lines as part of the machine learning in the following chapter. Algorithms can be designed and engineered in different ways using diverse methods. In accordance with [28] four categories of algorithms were mentioned: Generative-discriminative, static-dynamic, stable-unstable and regularized. Apparently every named denomination consists of subcategories that qualify themselves and their opposites. Among a myriad of classifiers, common exponents own statistical properties, such as artificial learning as machine learning, vector classification or different types of performance capacity. The final size of classifiers and therefore the algorithm is dependent of the dimensionality concerning the ‘amount of data needed’ (Feature) [28]. On the other hand this magnitude is dependent of the ‘dimensionality of feature vectors’ [28] by an exponential increase that describes their quality. For predicting analysis linear, binary or multiclass classifier are common. In this endeavour representative classifiers will be introduced in chapter 3 and 4.

2.3.5 Machine Learning

‘Machine learning (ML) is a data science technique that allows computers to use existing data to forecast future behaviors, outcomes, and trends. Using machine learning, computers learn without being explicitly programmed’ [29]. Using the before mentioned classifier (chapter 4) Azure Machine Learning, as a cloud based system, is able to evaluate them in terms of this thesis. Doing so, a model containing all logical aspects has to be built in the graphical user interface (GUI) according to Figure 6. Evaluating this model leads to a number of results containing statistical comparisons.

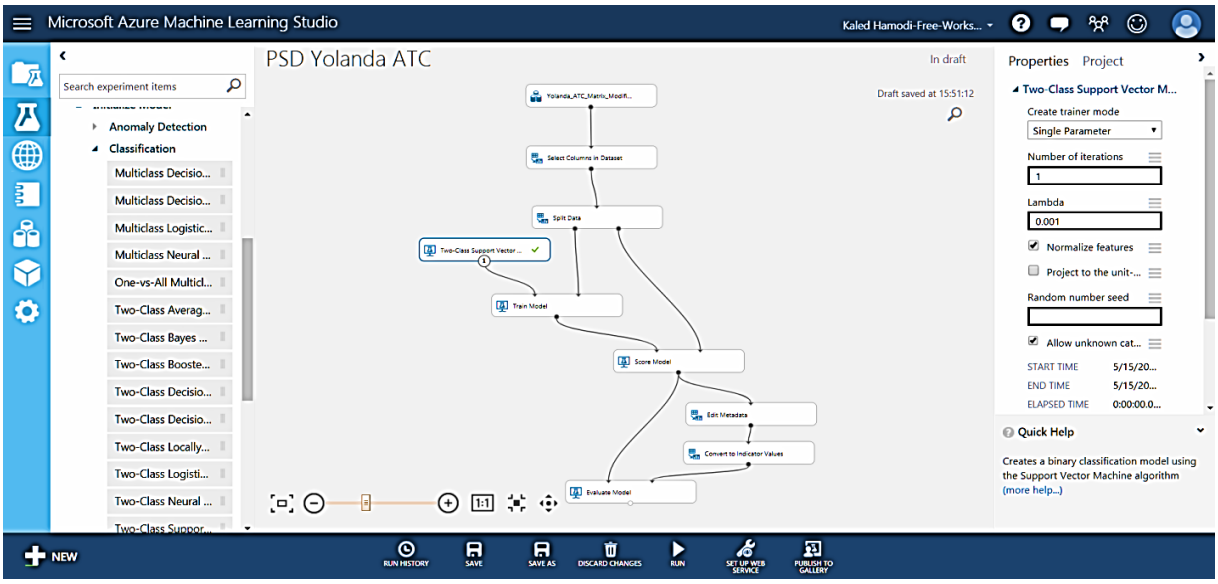


Figure 6: Azure Machine Learning GUI. Source: Own research

3. Hierarchical Concept

3.1 Morphological Box

In order to choose the right classifier for the ML, different possible solutions have to be taken into account. Therefore, several options are presented in the following morphological box:

	Classification	Regression	Anomaly Detection
Classifier	Linear	Binary	Multiclass
Type	Data rank	Predicting categories	Finding unusual data points
Informational value	Low	Medium	High
Ensemble	Yes	No	-
Iteration	1	n	-
Processing time	Short	Medium	Long
Accuracy	Low	Medium	High

Table 2: Morphological Box

According to the morphological analysis and the needs of the research, the assessed classifier options have to be scored in a functional analysis as follows.

3.2 Functional Analysis

Concept evaluation matrix	Linear classifier	Binary classifier	Multiclass classifier
	Score	Score	Score
Accuracy	1	9	3
Score for amount of variables	5	10	10
Type	3	10	10
Informational value	4	10	7
Ensemble	0	8	6
Iteration	0	10	10
Processing time	10	6	8
Mean value	3.3	9	8

Table 3: Functional analysis. Source: Own research,

The functional analysis requires binary and multiclass classifiers to be optimal for the research. Due to that a common classifier [4; 30] will be introduced. In endeavoring to optimize predictability in the results, a second classifier will be introduced [30]. This motivation shall provide a contrast between both classifiers.

4. Classifier Properties

4.1 Support Vector Machine

SVM is known as the maximization of the distance between the features of different classes and the optimal hyperplane. As a SVM learns in a range between two features it is a discriminative (Chapter 2.3.4) classifier [4; 27; 30]. According to Figure 7 the optimal hyperplane has equal margins between the nearest support vectors that represent the features [28]. As a learning algorithm SVM is capable of being applied to machine learning. It is used to map low dimension data on a vector with a high dimension. This high dimension vector is applied to the computer as learning data and allows using linear classification methods rather than more complicated non-linear methods and therefore a reduction of the error risk. The classification of features is in SVM defined by the hyperplane.

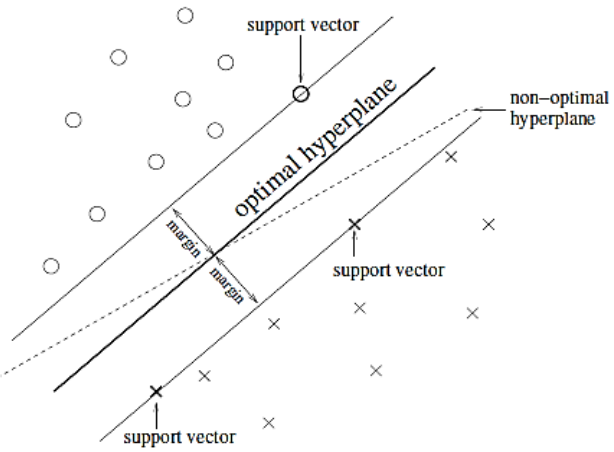


Figure 7: Difference between optimal and non-optimal hyperplane. Source [28]

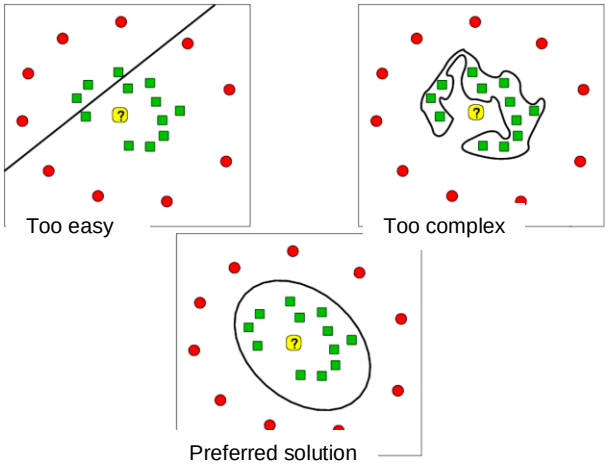


Figure 8: Fitting classes. Source: [38]

A disadvantage is that linear bounding may lead to under- or overfitting as seen in Figure 8. According to [28] SVM uses a Kernel function (Equation 1.3) in order to transform the feature space from 2D into 3D. This Kernel function uses a similarity procedure to quantify feature characteristics, allocating them into the predicted group and using the standard deviation σ, as it is a Gaussian function.

$$K(x,y) = e^{\left(\frac{-\|x-y\|^2}{2\sigma^2}\right)} \qquad 1.3$$

4.2 Boosted Decision Tree

Boosted decision trees are ensembles of several classifiers. Due to different types of processing, the terms ‘boosting’ and allocating data in a ’tree’ form have to be explained separately.

Decision trees contain sequential nodes describing ‘particular properties’ or questions [30]. This classifier begins with the root and ends with the leaf node visualizing the prediction/answer. Every node contains allocated features, from the feature matrix, as seen in Figure 9. Those features need to be relocated (discriminative, chapter 2.3.4) (Split) in a bin for further procedure. This ongoing procedure is repeated with randomly selected data from the feature matrix until the optimal result is achieved.

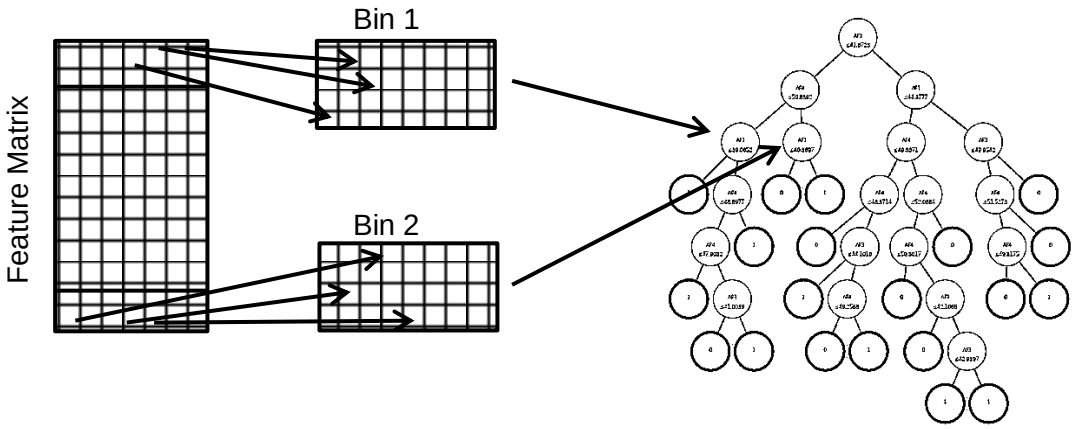


Figure 9: Sample of decision tree and procedure. Source: Own research

Boosting refers to an optimization technique (Figure 10), as several randomly chosen sub-sets D_n of the complete training set D are introduced to different classifiers C_n within the ensemble C. Due to that this classifier has the ability to reduce the general bias by combining and correcting the errors, results and therefore properties of every single classifier C_n. As a result the boosting achieves preferable results compared to other classifiers. [30]
The boosted decision tree is the result of both procedures combined.

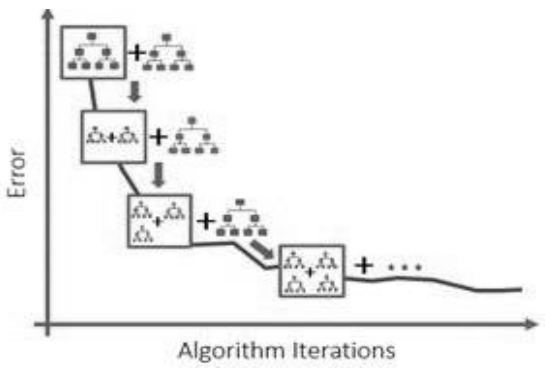


Figure 10: Error correction due to boosting. Source: [38]

5. Experimental Characteristics

5.1 Biometrical Background

As described in chapter 2, brain activity is measurable due to the conductivity of the skin. According to the 10-20 international location norms, all of the device electrodes are arranged corresponding to the electrode placement system, either 10% or 20% away from each other around the scalp. This electrode placement system is important due to location of different signals of the brain. Thus, the scalp is fragmented into fields of interest to measure the brain activity [4]. As the norm divides the brain into five segments, Figure 11 and Figure 12 show a combination of letters and numbers for the electrode placement of the Emotiv EPOC+ system. 'The letters F, T, C, P and O stand for Frontal, Temporal, Central, Parietal and Occipital sections respectively, (the letter C identifies the horizontal central line and it does not reference any lobe)' [31]. The two electrodes AF 3 & 4 refer to the 'frontal" functional regions of the frontal sections, whereas 'A' stands for 'anterior' (lat. Frontal). The numbers are allocated according to the side of the brain; however even numbers are on the right side and respectively odd numbers on the left side. When looking at the electrodes, every electrode has to be considered as a combination of either two measuring electrodes, or a measuring electrode and a common mode sense (CMS) and driven right leg (DRL) as reference electrodes. For the electrodes to detect electrical signals, a saline fluid, similar to contact lens fluid has to be dropped on the electrodes first to ensure a higher conductive quality between the scalp and the electrode.

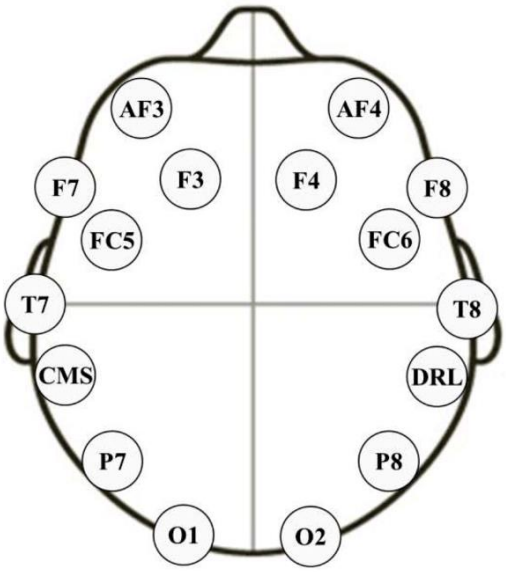


Figure 11: Electrode placement. Source: [38]

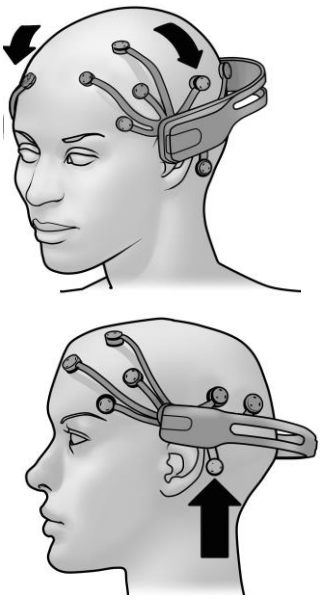


Figure 12: Headset Placement. Source: [32]

5.2 Experimental Device

The Emotiv EPOC+ system consist of the headset (Figure 13), 16 electrodes, the USB receiver and the foresaid saline solution. Emotiv names a headset specification for the data acquisition, electromagnetic compatibility, safety and radio disturbance in accordance with EN 300 440-2, EN 301 489-1, -3, EN 60950-1:2006, IEC 60950-1:2005 and AS/NZS CISPR 22:2009. Those standard norms define amongst the system requirements a strong variety enabling the recording of data with a variable resolution according to the researchers needs as shown in Table 4.

Figure 13: Emotiv EPOC+. Source: [38]

Moreover, the headset consists of a 3-axis Accelerometer, Gyroscope and a Magnetometer to detect the heads position relative to the earth gravity. In addition to that the hardware is delivered with the EmotivPRO Software for data recording and processing.

	EEG HEADSET
Number of Channels	14 (plus CMS/DRL reference, P3/P4 locations)
Channel names (International 10-20 locations)	AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4
Sampling Method	Sequential sampling, Single ADC
Sampling Rate	128 SPS (2048 Hz internal)
Resolution	14 bits 1 LSB = 0.5 μV (16 bit ADC, 2 bit instrumental noise floor discarded)
Bandwidth	0.2 – 45 Hz, digital notch filter at 50 Hz and 60 Hz
Filtering	Built in digital 5th order Sinc filter
Dynamic Range (input referred	8400 μV (pp)
Coupling mode	AC coupled
Connectivity	Proprietary wireless, 2.4 GHz band
Power	Lipoly
Battery life (typical)	12 hours
Impedance	Real-time contact quality using patented system

Table 4: System specification. Source: [32]

5.3 Experiment Evaluation

In order to evaluate EEG data the procedure from chapter 2.3 has to be done in MATLAB according to the following steps:

- running EEGLAB extension in MATLAB
- introducing channel locations on the scalp using the predefined .ced file

Pre-processing

- Artifacts removal due to:
  - offset of the alternate current signal in the brain
  - range discrimination between 0.1 and 60 Hz using the finite impulse response filter (FIR)
  - noise at 50 Hz of the electrical net using a notch filter between 49 and 51 Hz
  - eye-blinking with activity at anterior frontal and occipital lobe using independent component analysis (ICA) according to Figure 14

Feature extraction

- Extracting spectral (chapter 2.3.3) data by filtering events:
  - with introducing three additional columns for latency, type and duration of the features
  - saved as an ASCII-File in .txt format

- Extracting epochs:
  - by allocating epoch ranges from -1 to +2 seconds to get a frequency time window
  - choosing epochs randomly or structured from different difficulty levels according to the task

- Generating the power spectral density:
  - by plotting the PSD graph as seen in Figure 15
  - extracting data points within desired frequency range

.

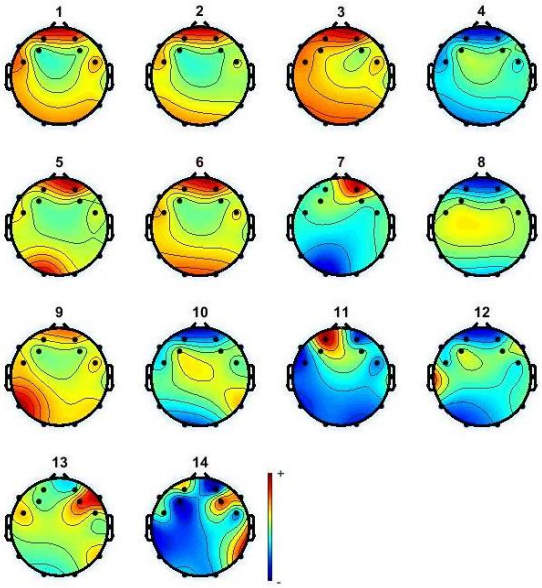


Figure 14: Topography of the brain activity. Source: Own research

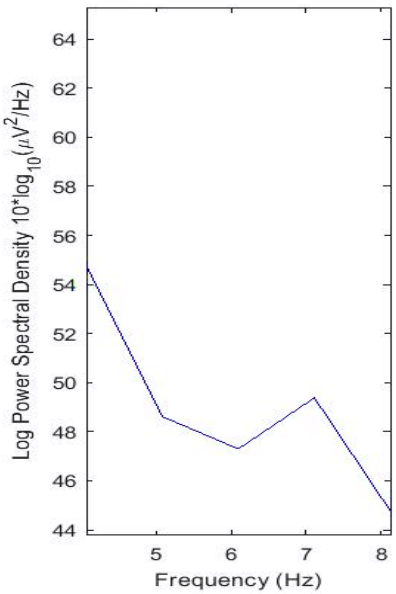


Figure 15: PSD plotting. Source: Own research

The extracted data is now ready to be classified using e.g. a SVM described in chapter 4. For this purpose machine learning, as a statistical data science, is used in terms of characteristical detection with further application. The described procedure is now applied to the experiments in chapter 6.

6. Preliminary Physiological Research

6.1 Electroencephalography

To understand the possibilities for measuring cognitive states using low cost EEG, those preliminary experiments shall give information about its use, advantages and disadvantages. Moreover, its purpose is to get familiar to the system, as well as procedures for a more professional experimental execution in the ATC experiments.

6.1.1 Preliminary Experiment Design

The experiment has been made in Microsoft Power Point (Figure 16) and it consists of three types of difficulty levels that the attendee has to pass. To ensure a comparison of the recorded data with a state of no activity 'EmotivPRO allows taking a baseline recording for each session. The baseline recording helps to compare the recording data with a baseline session and is recorded as part of the overall recording timeline. The baseline protocol involves…individual states where the subject is asked to open or close their eyes. An audio cue is used to mark the start and end of the eye open and eye close phases' [32]. In the beginning of the experiment the test person has to read instructions. Those instructions consist of four points: mental preparation, task definition, hints on solvation and the aim of the test. Every difficulty level has to be answered in a predefined time. The estimated time for the first level is 4 seconds, level 2 is set to 7 seconds and the third level to 10 seconds regarding the difficulty of the tasks.

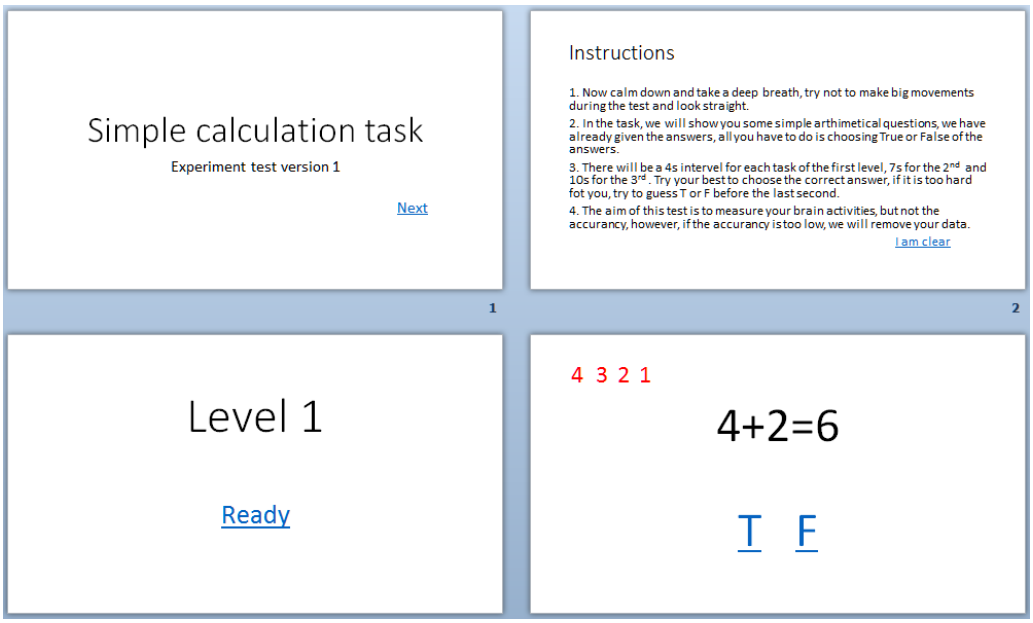


Figure 16: Preliminary arithmetical test procedure. Source: Own research

The attendees have to solve several math subjects containing different types of calculation, in which the solution is already given. The person has to answer those questions by a mouse click with true or false as shown in Figure 16. The slide will change automatically to the next page with the next task when answering the question or the time period running out.

This experiment design ensures a minimum of artifacts due to eye or respectively muscle movement as the attendee is not dependent of changing the slide by his own. They are also instructed not to make intense movements, affecting the EEG signals with artifacts. Such a procedure allows the obtaining of a cleaner EEG signal from the beginning, without a high amount of filtering the signal. Every subject is asked to do the experiment twice with a different set of questions for better comparability and anomaly detection.

6.1.2 Preliminary Experiment Execution

The preliminary experiment execution participates in the preparation for the main experiment (chapter 6.1.3) involving professionals in ATC (chapter 7). As the preliminary experiments are carried out by students and university employees, the initial task is superficial according to the chapter before. 11 persons have been chosen to carry out the experiment; four female and seven male test persons between an age of 20 to 40 years. The male students could start the experiment after small adjustments of the hair. The adjustment of the headset and assuring of a sufficient signal quality for the female subjects took more effort due to thick or long hair. Also a higher amount of saline solution has been applied. During the experiment all subjects have been healthy and fully aware of their senses. The experiment has been executed with no miscellaneous distortion by the environment and no interruption by the subjects. Depending on the individual performance, the experiments had a duration between two and more than three minutes, which corresponds to the maximum predefined experiment duration of 3.5 minutes.

6.1.3 Preliminary Experiment Result & Interpretation

The quantification and assessment of the usability for the applied device requires a high amount of data. In this endeavour a number of experiments were executed. High workload is described in chapter 2.1 as an increase of theta and decrease of alpha frequency. The measurements show a general characteristical behavior, correlating with literature pattern. The preliminary attempts have been done twice, for comparability, and evaluated according to chapter 2.4, using processed PSD data as a function of the frequency (Appendix A). Following an averaging and normalizing, mean PSD band powers demonstrate the progression as a function of the workload level (L=Low, M=Medium, H= High) containing 5 successive epochs (Table 5). All measurements show an overall positive tendency, achieving results between 50 and 100%

according to Table 6 and Appendix A. Moreover, the individual characteristics are symbolized as arrows. Those values have been calculated by considering the observed tendencies, as linear trend lines, over the literature observations.

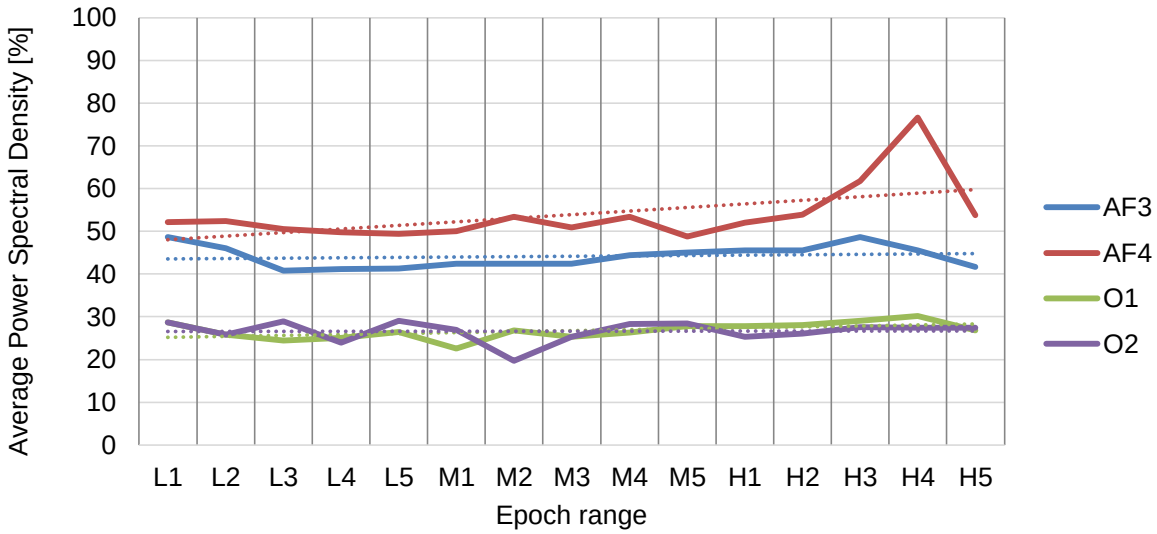


Table 5: Average PSD pattern. Source: Own research

The measured values in Appendix A oscillate in an average range between 36 and 85 $10*\log_{10}(\mu V^2/Hz)$ which equals 42 to 100% on the measured scale. The positive tendency refers to EEG pattern characteristics in agreement with the literature, introduced in chapter 2.2. Despite the individual EEG results, the general mean PSD level tends to be higher during the second attempt as shown in Table 6 and Appendix A, while single EEG activities differ from literature specifications. Resulting from the fact, that the phenomenon of increasing PSD band power, due to increasing workload is described in [33], this aspect shall be proposed as an assumption for the executed experiments. Furthermore, the appearance of the increased mean power level during the second attempt is postulated in this research as a result of increased mental effort, referring to chapter 2.2.

Subject	AF3		AF4		O1		O2		Mean PSD band power [%]		Correct tendency [%]	Average Standard deviation [%]	
Attempt	1	2	1	2	1	2	1	2	1	2	-	1	2
#1	↑	↑	↑	↓	↓	↑	↓	↑	47	49	62.5	2.550	5.447
#2	↑	↓	↑	↓	↑	↓	↓	↓	47	65	62.5	3.110	3.554
#3	↑	↑	↑	↑	↓	↓	↓	↓	64	60	100	2.161	2.342
#4	↑	↓	↑	↓	↑	↓	↓	↓	47	53	62.5	2.582	2.702
#5	↑	↑	↑	↑	↓	↓	↓	↓	55	56	100	2.538	1.755
#6	↑	↑	↑	↑	↑	↓	↓	↓	44	85	87,5	2.506	1.953
#7	↑	↑	↑	↑	↑	↑	↑	↑	62	60	50	2.161	2.163
#8	↑	↑	↑	↑	↑	↑	↑	↑	61	63	50	2.004	1.562
#9	↑	↓	↑	↑	↑	↓	↓	↓	48	52	75	5.357	4.836
#10	↑	↓	↑	↓	↑	↓	↑	↓	45	47	50	5.417	4.004
#11	↑	↑	↑	↑	↓	↓	↓	↑	47	49	87.5	5.604	2.553
Literature	↑	↑	↑	↑	↓	↓	↓	↓	-	-	100	-	

Table 6: Preliminary experiment results. Source: Own research

Considering all analyzed measurements, the author postulates the development of fatigue and an ongoing higher mental effort due to a general increasing band power level, as the majority of the subjects mentioned an increase of mental effort during the second attempt. This shall lead to a better understanding of following experiments. Moreover, the used device for EEG measurement provides a sufficient level of stable data, according to the correct tendencies in Table 6 and provided that enough saline solution is applied to the felt and no intense movements are made by the subject. Those characteristics have been observed by evaluating brain activity referring to chapter 6.1.3, as the average trend and PSD band power correlate with observations in the literature.

7. ATC Experiment

By introducing the Skysoft-ATM software in Figure 17, the south eastern as well as the north eastern sectors of Madrid-Barajas International Airport (LEMD/MAD) have been simulated and controlled by three professional female air traffic controllers with duration of one hour. Depending on the simulation 37 – 53 in- and outbound aircraft with several proximity warnings and no-fly zones due to military activity with up to 120 instructions have been delegated.

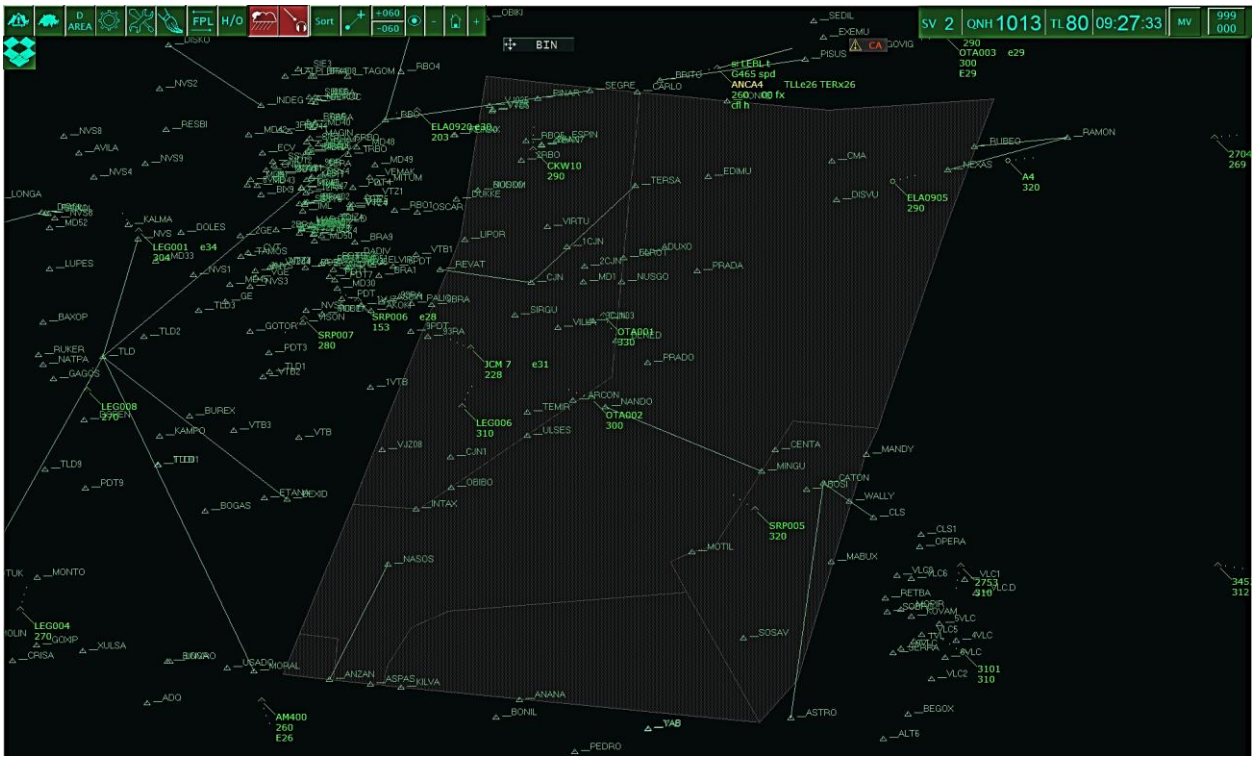


Figure 17: SkySoft-ATC Software. Source: Own research

The signal quality reached 70% – 85% due to long hair and the subject's motion behavior during communication with other controllers. The signal quality has been improved before the measurement by applying saline solution directly between the electrode and the hair of the subject.

All controllers were challenged during two continuous levels of air traffic controlling with resulting transitioning low and high workload. It is observed in the preliminary experiment, that data from the medium level tends to overlap with low as well as high level. Therefore this data has been reassigned expecting a better comparability between both levels.

7.1 ATC Experiment Result #1

The first measurement shows marginally differing trend lines between frontal and occipital lobe with approximately 12% gradual difference with rising workload in Table 7. Besides, the average band power levels indicate an oscillating frontal lobe activity in both hemispheres between 64 and approximately 73% of the measured PSD range. This intensive difference with noticeable temporal peaks indicates a partial characteristic for high workload during the simulation, as well as considerable increased mental effort due to rising alpha PSD resulting from highly predominant fatigue declared by the subject, due to an intensive working day.

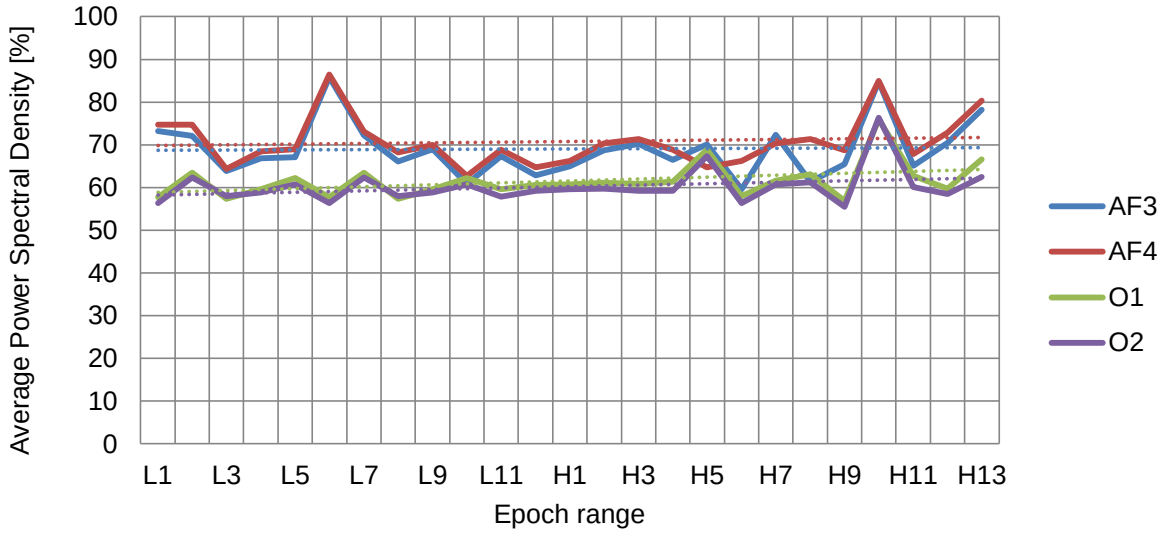


Table 7: ATC Experiment #1. Source: Own research

7.2 ATC Experiment Result #2

The second measurement indicates a characteristic workload form with noticeable increasing theta and marginally decreasing alpha band power with an average gradual difference of 43%. The same subject as in the first experiment declared a less state of fatigue on that day. It is observed, that theta is characteristically higher than the alpha band, primarily on the right hemisphere, due to optical impressions and spatial thinking [34] experienced by the ATC. Although the first experiment shows the same phenomenon, it is more visible in the second attempt. The oscillation range is equal to the previous attempt.

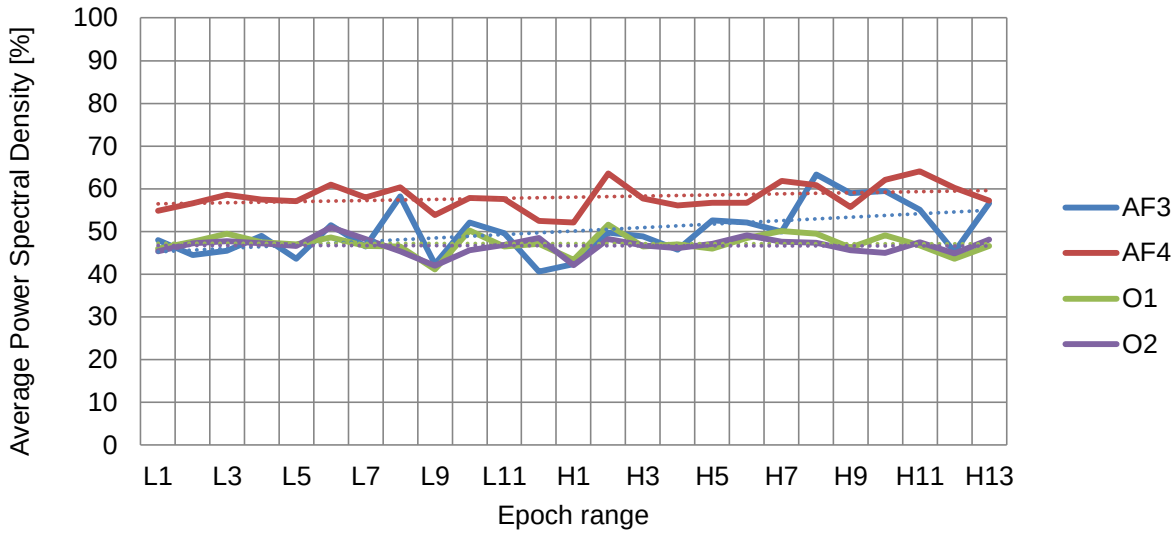


Table 8: ATC Experiment #2. Source: Own research

7.3 ATC Experiment Result #3

During the third experiment the subject shows average PSD activity, gradually differing between theta and alpha frequency by 33.5%. The subject induced a large amount of artifacts to the measurement due to intensive head and arm movement. As a result, partial peaks are visible.

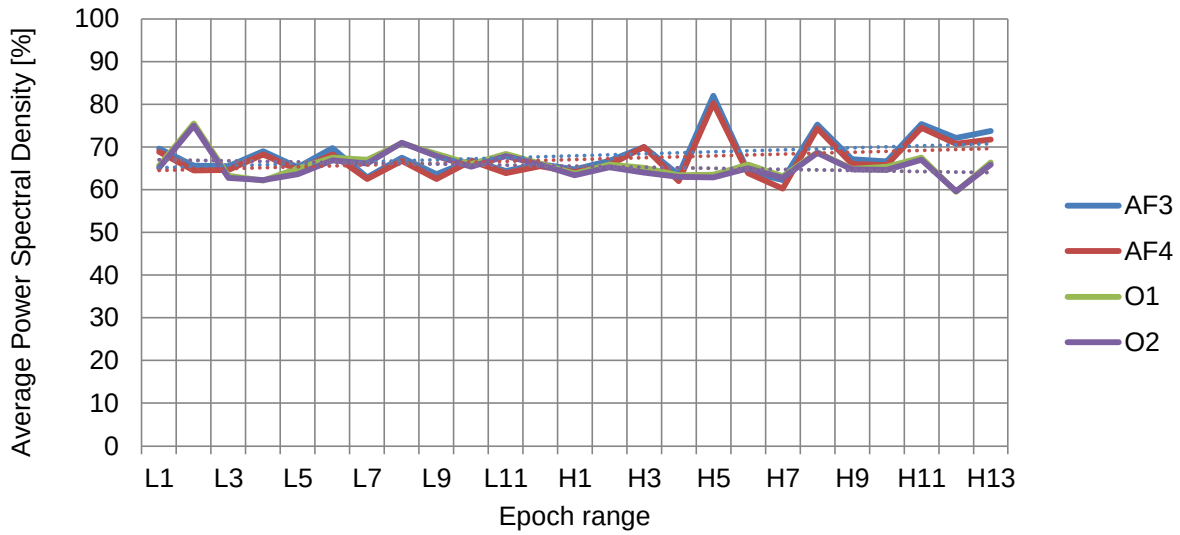


Table 9: ATC Experiment #3. Source: Own research

7.4 ATC Experiment Interpretation

Subject	AF3	AF4	O1	O2	Mean PSD band power [%]		Correct tendency [%]	Average standard deviation [%]	
ATC #1	↑	↑	↑	↑	AF	O	50	AF	O
					70	61		7.930	6.351
ATC #2	↑	↑	↓	↓	54	46	100	7.108	3.327
ATC #3	↑	↑	↓	↓	67	66	100	4.623	3.109

Table 10: ATC Experiment

Although the state of the art argues that excessive workload develops due to decreasing theta PSD and attenuated alpha band, the author hypotheses that there is a stronger influence in increasing pattern due to present fatigue in ATC experiment #1. Furthermore, the intense oscillation frequency suggests a high predictability by the ML, due to an intense SD in the form. The second and third ATC experiments have a correct tendency according to literature pattern. Moreover, the mentioned predominance of the first attempt PSD band level over the second is visible in the ATC experiment during a measurement of one hour. As an interconnection, this is shown in the average standard deviation.

8. Machine Learning

8.1 Setting

The setting for the Machine Learning (ML) contains of 9 components, as visualized in Figure 18, chapter 2.3.5 and Appendix B. In order to achieve a high resolution the data has been spread artificially. 99 linear trend values have been added between every two features to enable the ML evaluating the same data pattern with a higher data density. The uploaded data has to be saved in Microsoft Excel as a .csv file in one column. This column has to contain all data separated by a comma. As Microsoft Azure splits the data according to the placed comma, the required columns have to be chosen. Moreover, ML requires splitting the dataset into training and a test set which leads to a separation of the dataset with 75% for the training and the remaining 25% for the comparing test set. For training the 75% of data, a classifier has to be chosen. As mentioned in chapter 4 one of the most common classifiers is the Support Vector Machine. In this endeavour the 'Two-Class-Support Vector Machine' has been chosen in order to classify low and high workload. After scoring the model and extracting the data of interest, the model evaluation follows. The result is presented in a Receiver Operator Characteristic (ROC) graph with true positive and false positive rate in contrast and relating values such as accuracy, area under curve (AUC) and others shown in Figure 19.

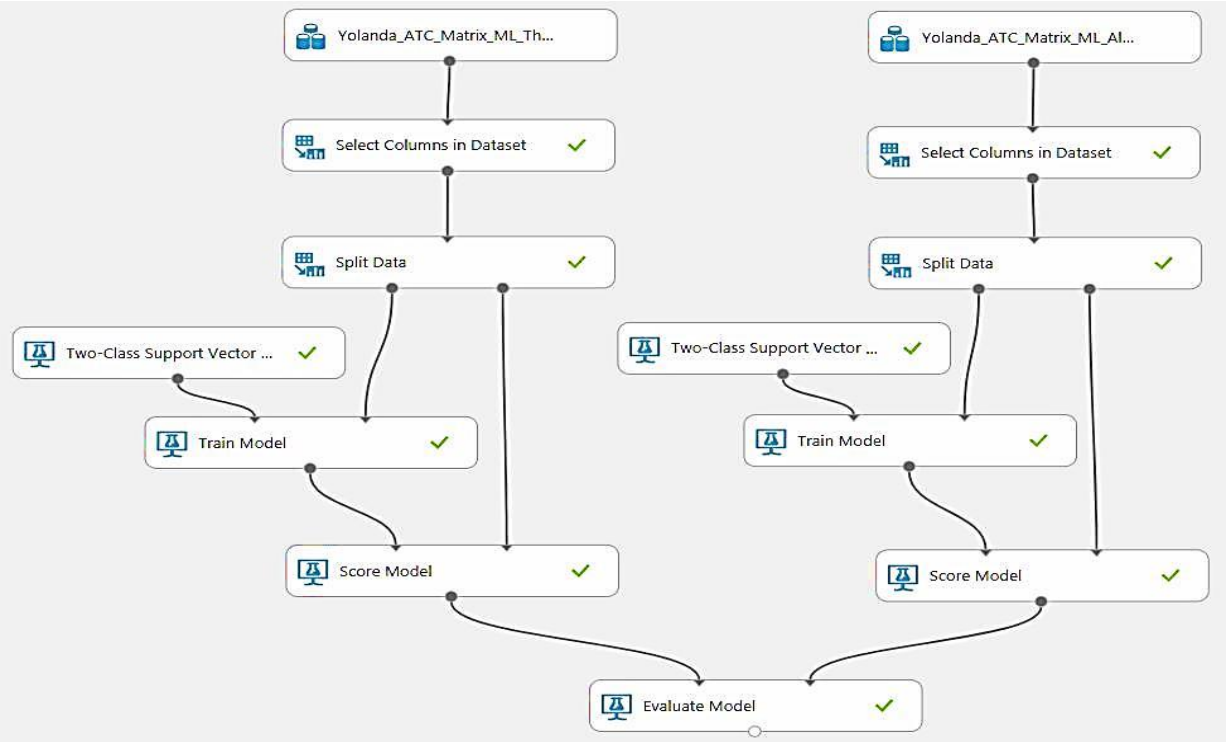


Figure 18: ML setting. Source: Own research

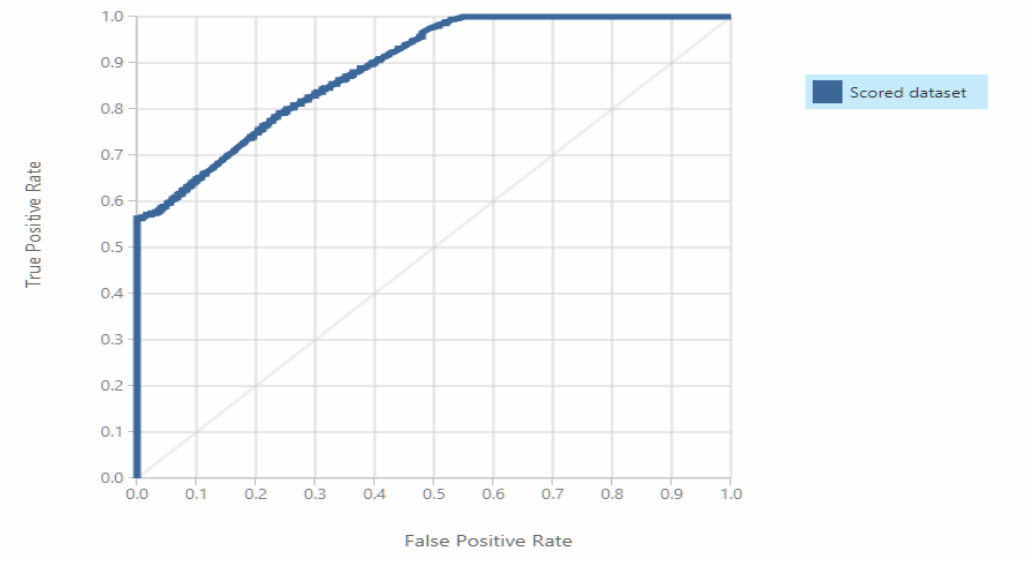


Figure 19: Sample ROC diagram of workload predicition. Source: own research

To understand those values a brief description follows:

- True Positive

True data that was correctly predicted to be true.

- True Negative

False data that was correctly predicted to be false.

- False Positive

True data that was incorrectly predicted to be false.

- False Negative

False data that was incorrectly predicted to be true.

- Accuracy

Describes the ratio between the accurate values and the total value. It is an indicator for the performance of the algorithm.

- AUC

Describes the area under the curve (AUC) and the predictive quality of the algorithm.

Figure 19 shows a linear recess at approximately 55% resulting in a predictability of 89%. Introducing a basic rule high predictability is given, when the scored dataset has a high distance to the median 50% - line. As the AUC has the highest informational value amongst the mentioned parameters, the results of the following chapter are referred to it.

8.2 ML Experimental Results and Interpretation

The ROC – diagrams (Figure 20 - 22) of the ATC experiments show similar results between the all attempts:

	AUC [%]		Accuracy [%]	
	Theta (Blue)	Alpha (Red)	Theta (Blue)	Alpha (Red)
ATC #1	54.9	74.7	57.6	73.0
ATC #2	82.1	54.4	75.0	55.4
ATC #3	71.9	79.2	66.7	78.5

Table 11: AUC values of ATC experiments. Source: Own research

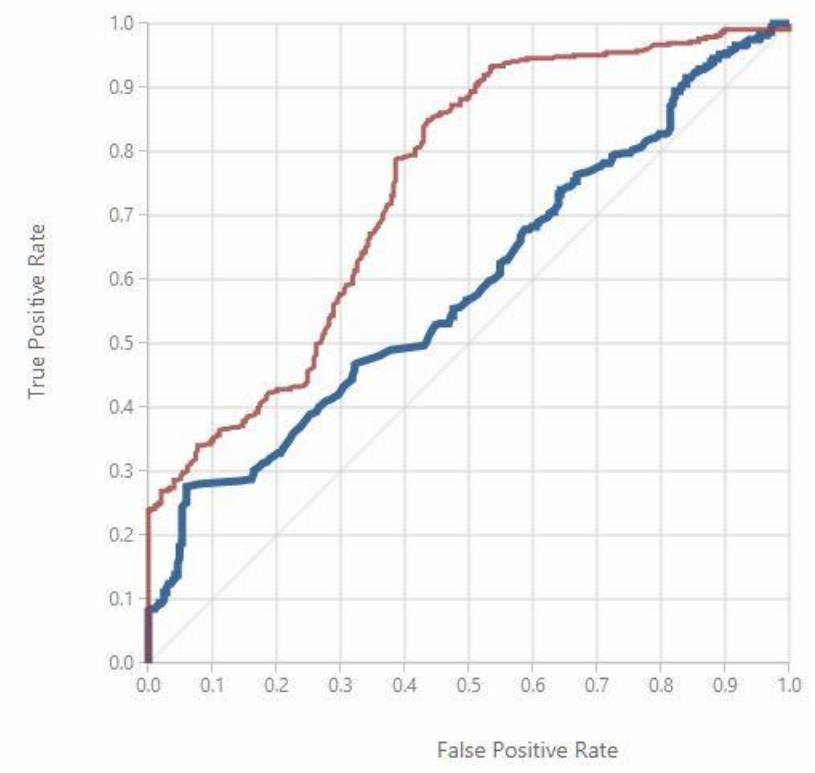


Figure 20: ATC Experiment #1 ROC - diagram. Source: Own research

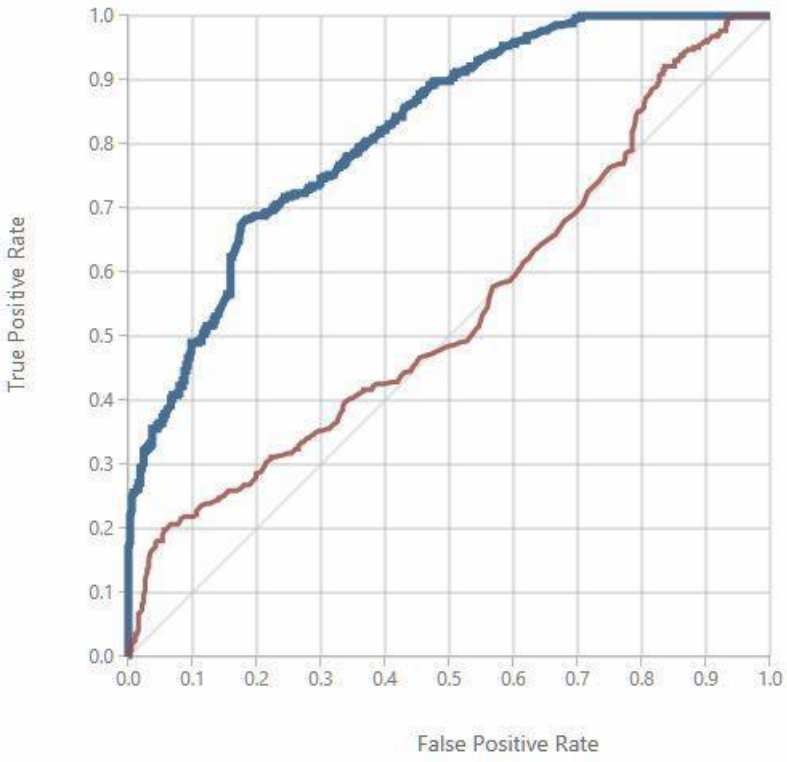


Figure 21: ATC Experiment #2 ROC – diagram. Source: Own research

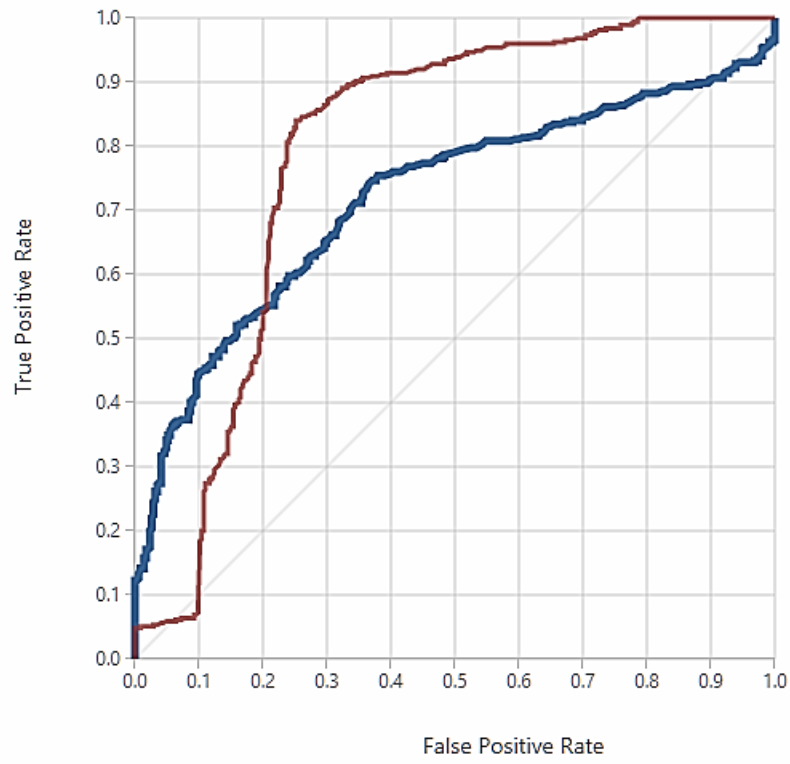


Figure 22: ATC Experiment #3 ROC - diagram. Source: Own research

According to the experimental setting and the measured feature characteristic in chapter 6.1.3 and chapter 7.1, different training results have been achieved in evaluating all 25 experiments. Those results differ in ranges between 54.9 with low and 96.6% with high predicting quality of the AUC value (Table 12) and have been calculated by averaging the value of theta and alpha. Furthermore, the accumulation of all values in the right upper quarter is noticeable.

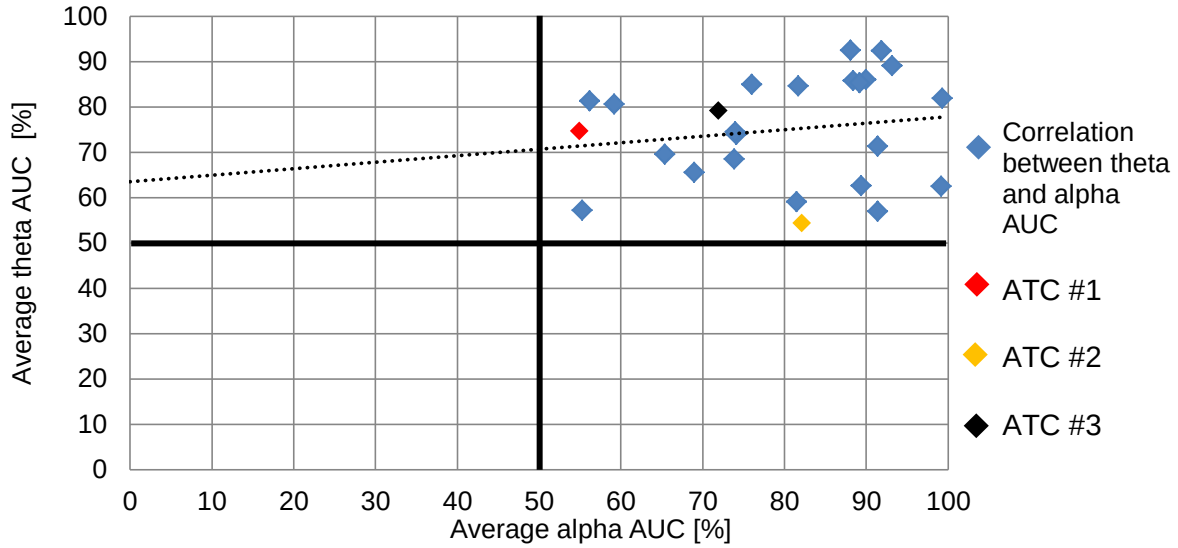


Table 12: Correlation between theta and alpha AUC. Source: Own research

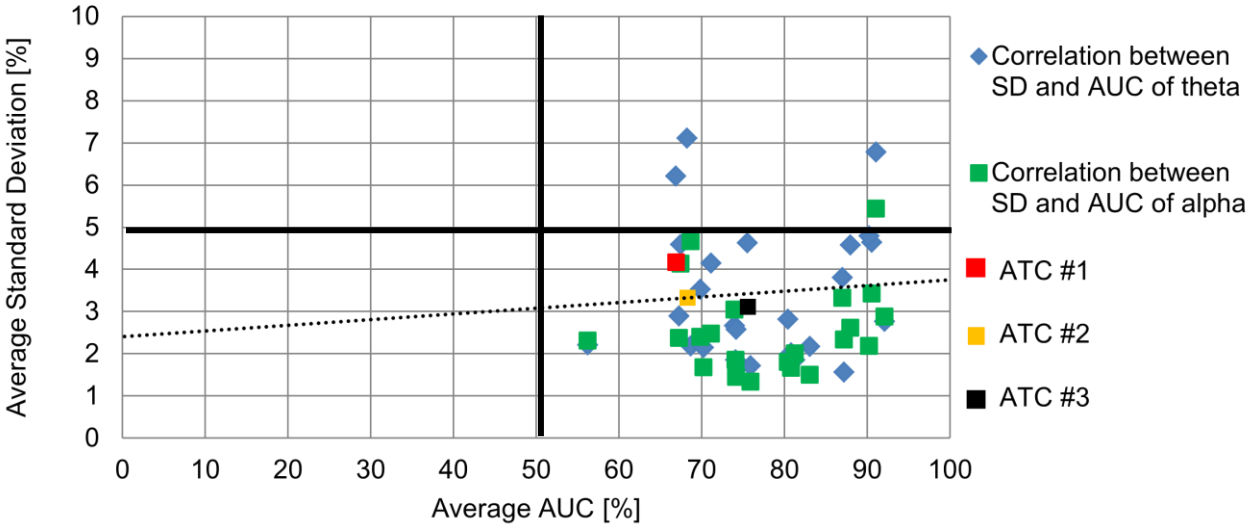


Table 13: Correlation between SD and AUC. Source: Own research

Seen in Table 13, AUC values correlate in the 65 – 95% area with PSD quality, illustrated as the SD with two statistical outliers. It shows a U-shaped accumulation deviating between 1 – 6% of the vast majority, which leads to the assumption of the possibility of high measurement quality, relating to the conclusion in [19] that increased amplitude indicates higher mental effort. As the SD specifies the mean amplitude of the measured brain activity and increasing workload results in increased PSD band amplitude, the author concludes that neurological informational value and therefore measurement quality are related to the workload predictability using the Support Vector Machine with an increasing trend.

8.3 ML Optimization

As the support vector machine achieves ML results of approximately 77% in average, according to the previous chapter, the optimization of the algorithm is preferable. Therefore, the boosted decision tree, introduced in chapter 4.2, is applied in this chapter. Achieving AUC scores between 77.8 and 100% for the 25 experiments (Appendix C) in total, this algorithm shows optimal results for the AUC, accuracy and therefore predictability of workload states.

According to Appendix A different measurement results have been introduced to this classifier showing optimal results. This observation leads to the assumption that this classifier is less dependent on a high measurement quality. This hypothesis was disproved by testing low quality PSD band signal. Increasing the iteration rate, amount of decision trees and leaves, provokes an improvement in a certain range. Exceeding this range, results in self-compensation of ML qualities, leading to a reversion of the AUC up to 50% predictability.

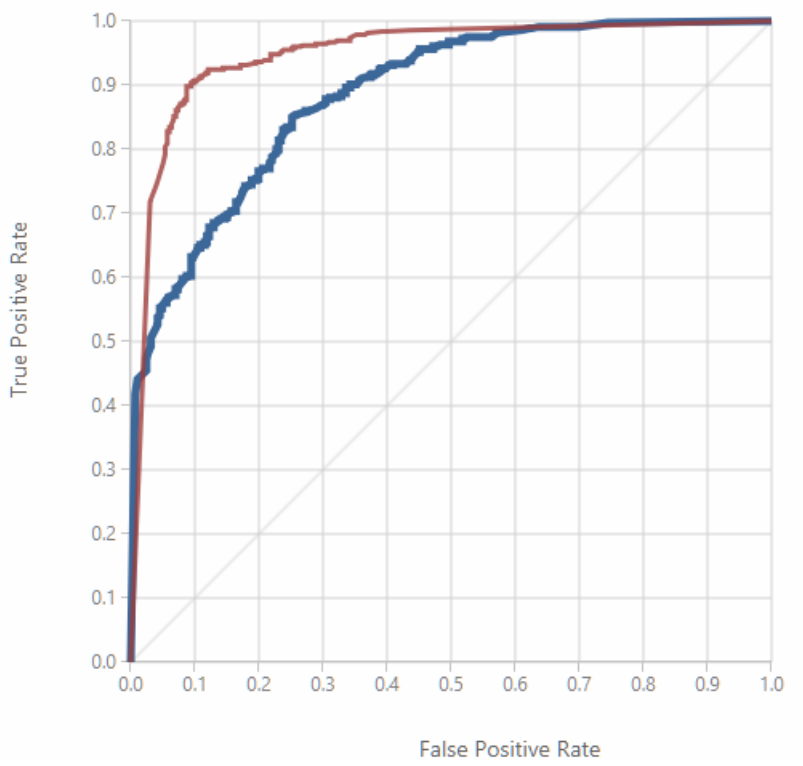


Figure 23: ATC Experiment #1 Improved ROC - diagram. Source: Own research

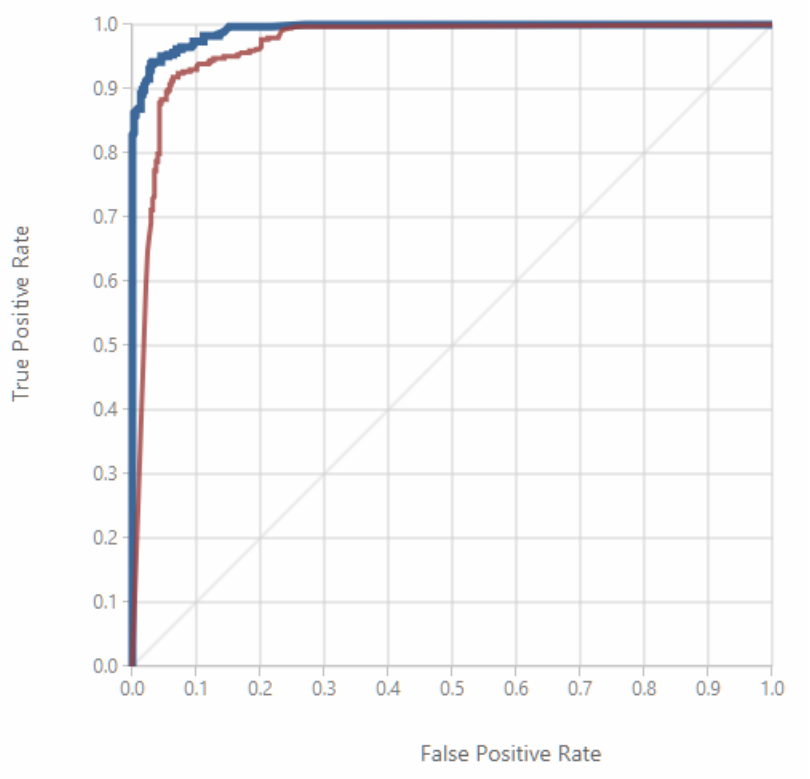


Figure 24: ATC Experiment #2 Improved ROC - diagram. Source: Own research

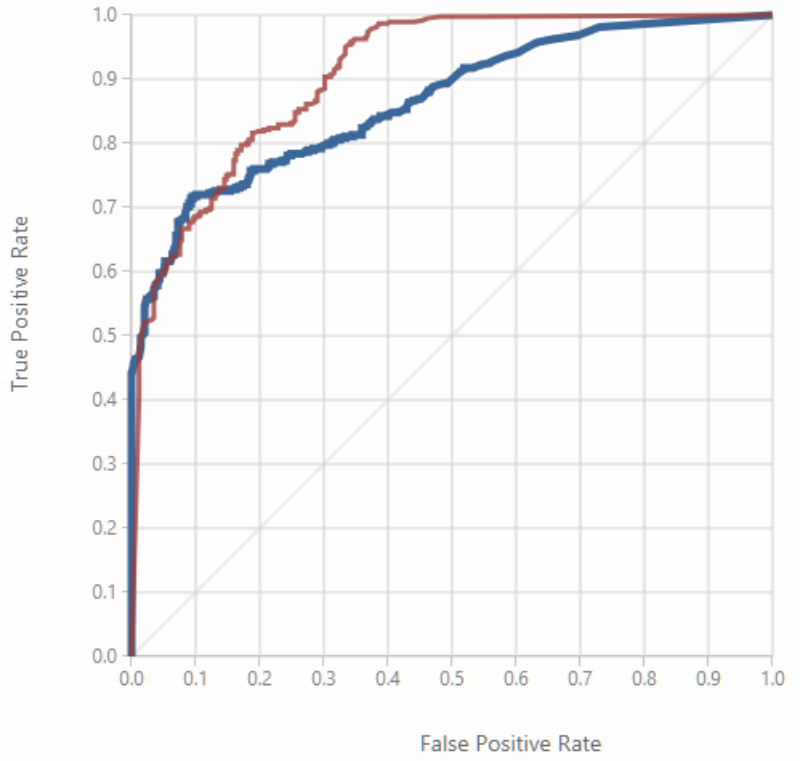


Figure 25: ATC Experiment #3 Improved ROC - diagram. Source: Own research

	AUC [%]		Accuracy [%]	
	Theta (Blue)	Alpha (Red)	Theta (Blue)	Alpha (Red)
ATC #1	88.7	95.7	80.2	89.9
ATC #2	99.3	97.2	94.9	91.7
ATC #3	86.5	90.9	77.8	81.0

Table 14: AUC values of ATC experiments (optimized). Source: Own research

Table 14 and Table 15 show the improved correlation between theta and alpha AUC and accuracy, accumulating the vast majority of chapter 6.1.3 to 95 and 100%. ATC measurements, as single outliers, are characterized by high predicting quality as they achieve AUC and accuracy values higher than 80% irrespective of the fact, that ATC experiments have more intensive artifact affections, than the arithmetical task. Compared to Table 12 and Table 13 of the previous chapter, the averaged values of Table 15 and Table 16, show a focused accumulation in the vicinity of 100%, due to the advantages of the algorithm, described in chapter 4.2. According to that, a low SD is preferable, to achieve high predictability.

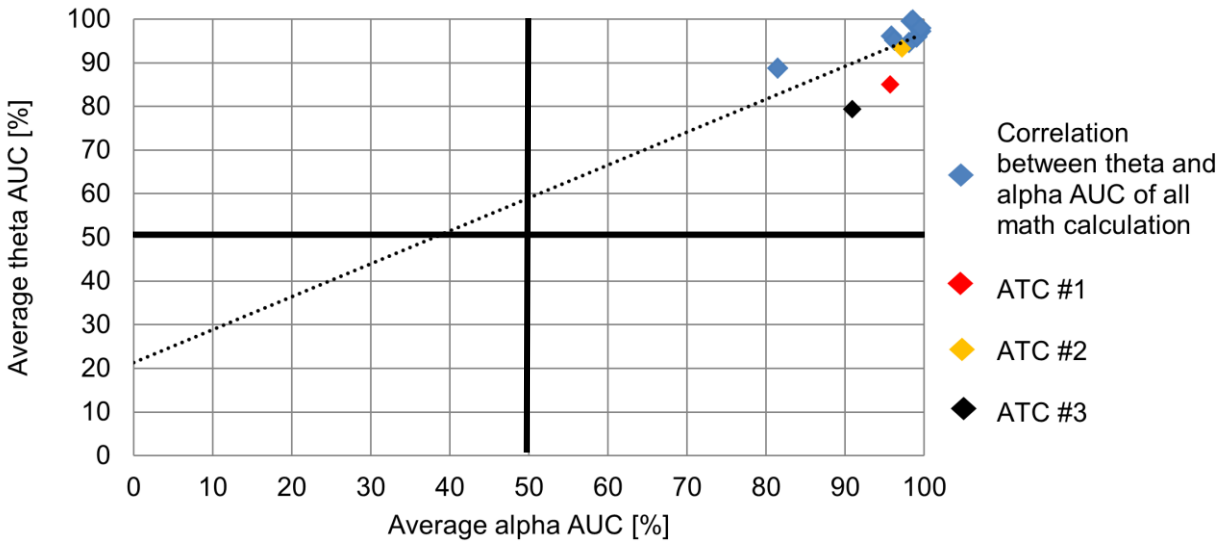


Table 15: Improved correlation between theta and alpha AUC. Source: Own research

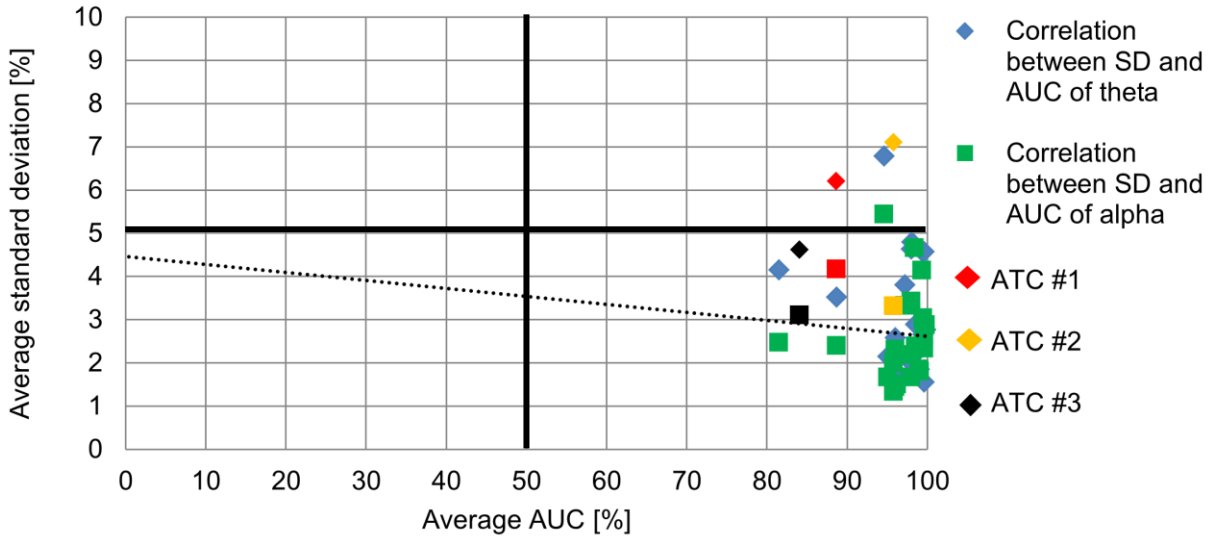


Table 16: Correlation between SD and AUC. Source: Own research

The executed measurements have proven a reliability of the measurement procedure. Still, those results have to be verified independently from statistical calculations in the following chapter.

9. Self-assessment

In order to compare the usability and informational value of the measurement in chapter 6.1.3, a self-assessment has been done after the experiments using the NASA TLX test. This test allows scoring the subjective self-rating containing mental demand and performance effort amongst other categories, on a 20-point scale. As a result the PSD bands of all 25 experiments have been averaged to three mean values for all workload levels. Following that, the averaged PSD band levels and the 20-point TLX are corresponding to two different scales. The comparison shows a visible difference between the self-assessment levels with increasing trend. A slight increase appears between the measured PSD band levels, as noticed in previous chapters. Regardless of the fact, that both scales are different, the tendencies are the similar.

Table 17 shows the similarity between PSD band and the self-assessment is an indicator for significance of the measured values and predicting quality.

According to the perceptions in this chapter and the aim of this work in chapter 1.1.2, the shown comparison quantifies the ability of the subjects and the measurement device to accomplish with new tasks and higher workload.

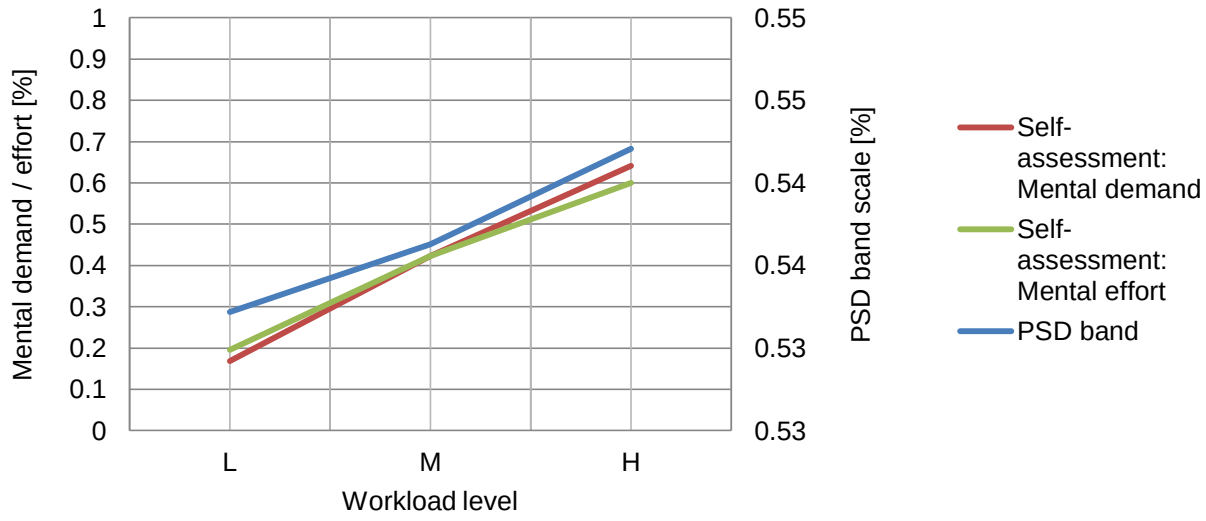


Table 17: Comparison between PSD Band and NASA TLX Test. Source: Own research

10. Summary and Conclusion

With focus on the aim of the research in chapter 1.1.2 and the associated requirements in chapter 1.3, several steps have been taken into account. After familiarizing with the state of the art and defining the experimental design, 25 experiments (22 preliminary and 3 ATC experiments) have been done, extracting brain activity using pre-processing, feature extracting and classifying. The preliminary, as well as the ATC experiments show an overall correct tendency with single outliers. The more intensive artifact affection of the ATC PSD band due to communication and more intense movement is noticeable. For this purpose saline fluid has to be applied, as a limiting factor. AUC values for both theta and alpha PSD band accumulate between 50 and 100% and enhance with improved algorithm. This observation is also valid for the accumulation between SD and AUC value. Furthermore, the application of the subjective test in the previous chapter, confirms the thesis introduced in chapter 1.2 and perceptions in 2.1. Regarding the applied methods in this work, the aim of the research has been fulfilled, meeting the requirements according to chapter 1.3.

An increased focusing on the accumulation in Table 16 is mirrored by improved ML and therefore predictability of the device. Regarding all applied methods this leads to the observation, that the Emotiv EPOC+ device is suitable to measure workload of ATC, as well as arithmetical tasks.

11. Perspective

As the research was executed within a simulated environment, many factors such as:

Operational and Organizational aspects as:

- ATC situation authenticity
- authenticity of the hardware
- working hour rotation
- simultaneous dual-task procedures
- standard flight routes, levels and rate of climb conformity
- coordination with collateral sectors
- time deficit

could not be reproduced. Moreover, other industrial areas such as automotives, have the potential to use this technology to monitor the performance of the drivers e.g. for Truck drivers or in the Formula 1 domain, integrating the EEG system directly into the helmet of the driver.

Therefore, additional research during the named authentical working environments has to be aspired.

Reference

[1]. **cf. Moon, W. C. et al.** Air Traffic Volume and Air Traffic Control Human Errors. *Journal of Transportation Technologies.* 2011.

[2]. **Directorate General for Mobility and Transport.** Mobility and Transport. *Welcome to the SESAR project.* [Online] European Commission, August 3, 2017. [Cited: March 2, 2018.] https://bit.ly/2x0qLxM.

[3]. **cf. International Air Transport Organisation.** European Assocation for Biometrics. *BIOMETRICS IN AVIATION.* [Online] European Association for Biometrics (EAB) in cooperation with IATA, June 12, 2017. [Cited: March 02, 2018.] https://bit.ly/2tykxAu.

[4]. **cf. Walter, C.** EEG workload prediction in a closed-loop learning environment. [Online] 2015. [Cited: February 27, 2018.] http://hdl.handle.net/10900/66395.

[5]. **cf. Berger, H.** *Über das Elektrenkephalogramm des Menschen.* Jena : s.n., 1924. pp. 1-44. ISBN: 0003-9373.

[6]. **cf. Cygnus Research International.** Power Spectral Density. [Online] [Cited: February 14, 2018.] https://bit.ly/2tqHOFe.

[7]. **cf. Heisz, J. J. et al.** Applications of EEG Neuroimaging Data: Event-related Potentials, Spectral Event-related Potentials, Spectral. [Online] June 27, 2013. [Cited: February 17, 2018.] https://bit.ly/2IlBBiB.

[8]. **cf. Wiersma, M.** *Identiying workload levels with a low-cost EEG device using an arithmetic task.* Gronongen : University of Gronongen, 2016. pp. 13-14.

[9]. **cf. Vocabulary.com.** Mental State. [Online] [Cited: February 23, 2018.] https://bit.ly/2lwkrWG.

[10]. **cf. MedRevise.** *Cognitive state examination.* [Online] September 05, 2008. [Cited: March 24, 2018.] https://bit.ly/2Mjn89o.

[11]. **cf. Parasuraman, R. et al.** Neuroergonomics. [book auth.] A. Gevins and M.E. Smith. *Electroencephalography (EEG) in neuroergonomics.* Oxford : Oxford University Press, 2007, pp. 15-31.

[12]. **cf. Klimesch, W. et al.** *A short review of slow phase synchronization and memory: evidence for control processes in different memory systems?* s.l. : Brain Res., 2008. pp. 31-44. ISBN: 1235:31–44.

[13]. **cf. Sauseng, P. et al.** Control mechanisms in working memory: a possible function of EEG theta oscillations. [book auth.] P. et al. Sauseng. *Neuroscience & Biobehavioral Reviews.* s.l. : Neurosci Biobehav Rev., 2010, pp. 1015-1022.

[14]. **cf. Freunberger, R. et al.** Brain oscillatory correlates of working memory. [book auth.] M.J. LaVoie. *Brain Research.* s.l. : Brain Research, 2011, p. 95.

[15]. *Towards a multimodal bioelectrical framework for the online mental workload evaluation.* **cf. Aricò, P. et al.** s.l. : 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2014. pp. 3001-3004.

[16]. **cf. Borghini, G. et al.** *Quantitative Assessment of the Training Improvement in a Motor-Cognitive Task by Using EEG, ECG and EOG Signals.* s.l. : Springer US, 2015. pp. 3001-3004. ISBN: 1573-6792.

[17]. **cf. Basu, A. et al.** *Intelligent Human Computer Interaction.* Pilani, India : Springer, 2016. pp. 122-136. ISBN 978-3-319-52502-0.

[18]. **cf. Damos, D. L.** *Multiple Task Performance.* London : Talor & Francis, 1991. pp. 1-25. ISBN: 0-80566-757-7.

[19]. **cf. Moray, N. et al.** Workload and Workload Measurement. [book auth.] Nevill Moray. *Mental Workload: Its Theory and Measurement.* Stirling, Scotland : Springer US, 1979, pp. 3-11.

[20]. **cf. Dijkstra, K.** Radboud Universiteit Nijmegen. [Online] August 29, 2011. [Cited: February 20, 2018.] https://bit.ly/2IntKBi.

[21]. *Cognitive State Estimation Based on EEG for Augmented Cognition.* **cf. Erdogmus, D. et al.** Arlington : Neural Engineering, 2005. pp. 1-4. ISBN: 0-7803-8710-4.

[22]. **cf. Lan, T. et al.** Semantic Scholar. *Estimating cognitive state using EEG signals.* [Online] 2005. [Cited: February 19, 2018.] https://bit.ly/2IlBqnf.

[23]. **cf. Al-Fahoum, A. S. et al.** Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. 2014, Vol. 2014, p. 2.

[24]. **cf. Bird, S. et al.** Supervised Classification. [book auth.] Steven et al. Bird. *Natural Language Processing with Python.* Sebastopol, USA : O'Reilly, 2014, pp. 221-231.

[25]. *Pre-processing and Feature Extraction Techniques for EEGBCI Applications- A Review of Recent Research.* **cf. Sarma, P. et al.** 1, Guwahati, India. : Prof. Shakuntala Laskar. et al., 2016, Vol. 5, p. 2. ISBN: 2348-7305.

[26]. **cf. Lotte, F. et al.** A Tutorial on EEG Signal Processing Techniques for Mental State Recognition in Brain-Computer Interfaces. [book auth.] E.M Prof. Miranda. *Guide to Brain-Computer Music Interfacing.* London : Sprinter, 2014, pp. 133-151.

[27]. *A review of classification algorithms for EEG-based brain–computer interfaces.* **cf. Lotte, F. et al.** s.l. : IOP Publishing, 2007, pp. 1-15.

[28]. **cf. Lotte, F. et al.** Journal of Neural Engineering. *A review of classification algorithms for EEG-based brain-computer interfaces.* [Online] 2007. [Cited: February 22, 2018.] https://bit.ly/2KftdWY.

[29]. **cf. Microsoft.** Microsoft Azure. *What is Machine Learning?* [Online] September 21, 2017. [Cited: May 15, 2018.] https://bit.ly/2lyiWax.

[30]. **cf. Duda, R. O. et al.** Linear Discriminant Functions, Nonmetric Methods. [book auth.] R. O. et. al. Duda. *Pattern Classification.* Canada : John Wiley & Sons, Inc., 2001, pp. 259-265, 395-410.

[31]. **cf. Kim, J. H. et al.** *Robot Intelligence Technology and Applications 2: Results from the 2nd International Conference on Robot Intelligence Technology and Applications.* Switzerland : Springer, 2014. p. 299. ISBN: 978-3319055817.

[32]. **Emotiv.** User Manual. [Online] 2014. [Cited: March 12, 2018.] https://bit.ly/2lvCqws.

[33]. **cf. Dan, A. et al.** Real time based measurements of cognitive load indicates mental states during learning. [book auth.] Andrew Olney. *Journal of Educational Data Mining.* Memphis, USA : University of Memphis, 2017, Vol. 9, pp. 31-44.

[34]. **cf. United States Army.** Chapter 1, Lesson 5: Left-rain/Right-Brain. [book auth.] US Air Force. *Know Yourself — Socrates:.* s.l. : Publishing directorate, 2016.

[35]. **cf. Emotiv.** [Online] [Cited: April 15, 2018.] www.emotiv.com.

[36]. **cf. Kunze, K.** Hauptseminar Machine Learning: Support Vector Machine, Kernels. [Online] January 13, 2004. [Cited: February 26, 2018.] https://bit.ly/2MfZt9S.

[37]. **cf. Piatetsky, G.** KDNuggets. *Understanding Machine Learning Algorithms.* [Online] Octobre 2017. [Cited: June 14, 2018.] https://bit.ly/2gvamdH.

[38]. *Robot Intelligence Technology and Applications 2 - Results from the 2nd International Conference on Robot Intelligence Technology and Applications.* **cf. Kim, J.-H. et al.** Denver, USA : Springer, 2014. p. 299. ISBN 978-3-319-05582-4.

Appendix

A. Average PSD

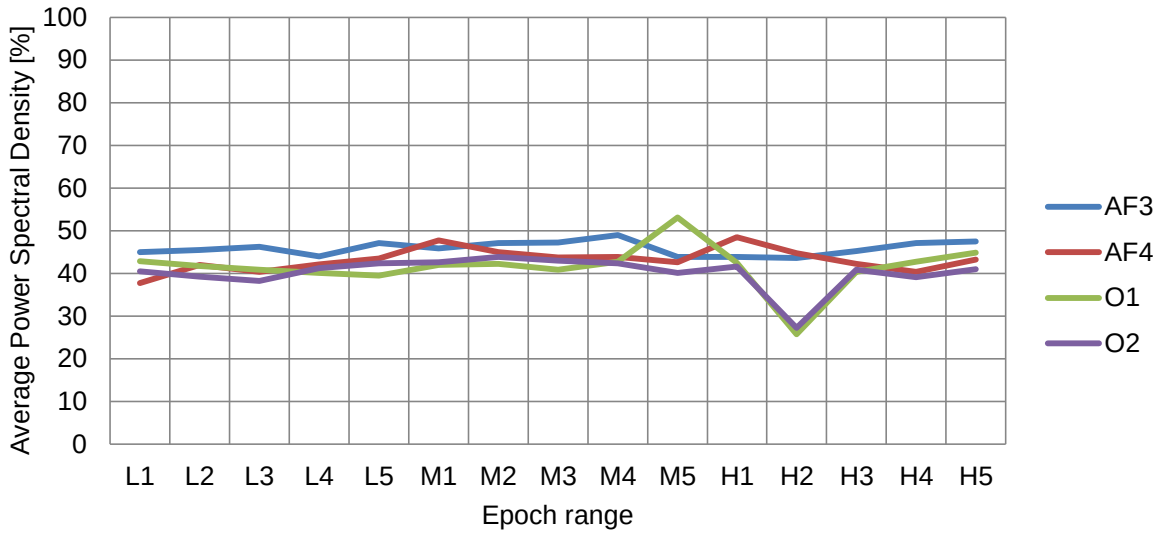


Table 18: Subject #1 - Preliminary Experiment #1

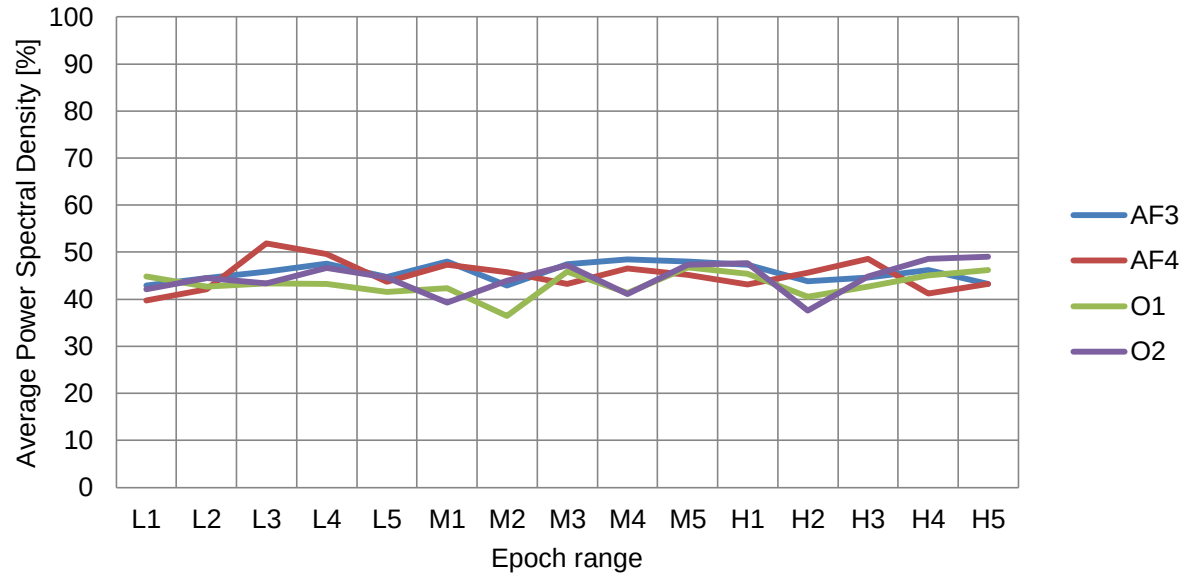


Table 19: Subject #1 – Preliminary Experiment #2

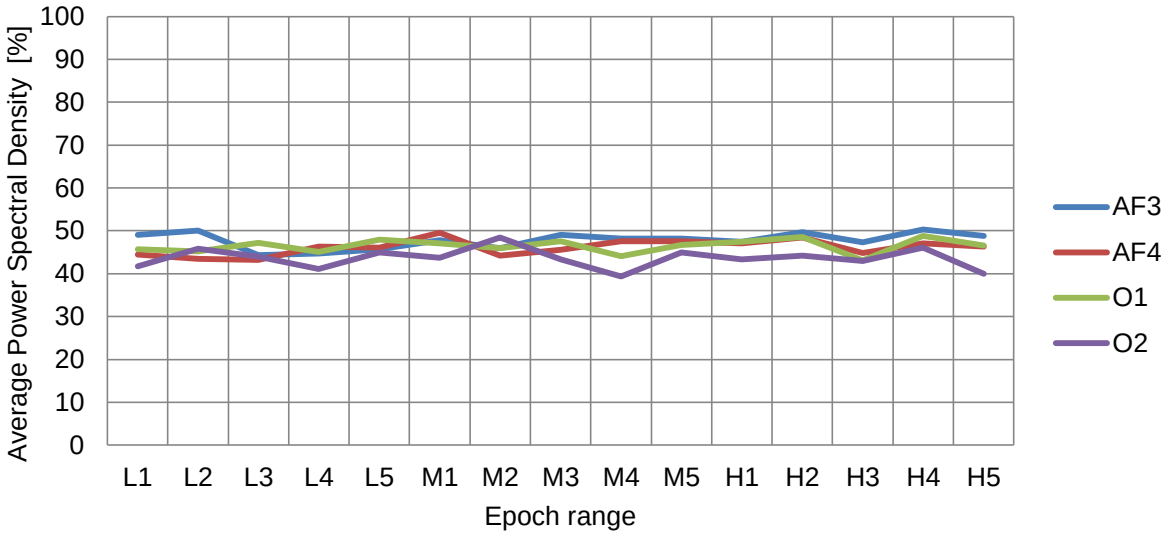


Table 20: Subject #2 – Preliminary Experiment #1

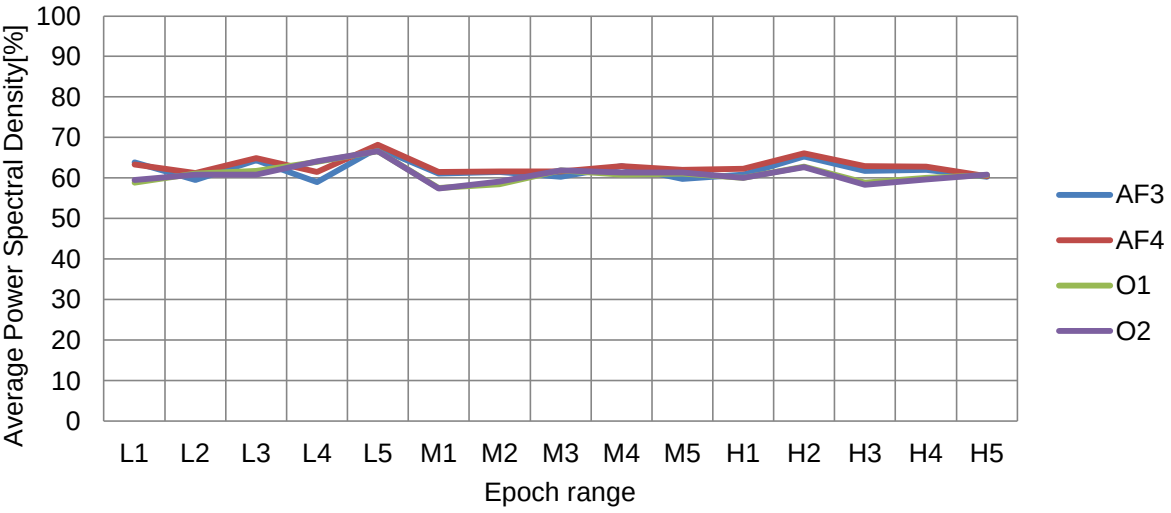


Table 21: Subject #2 - Preliminary Experiment #2

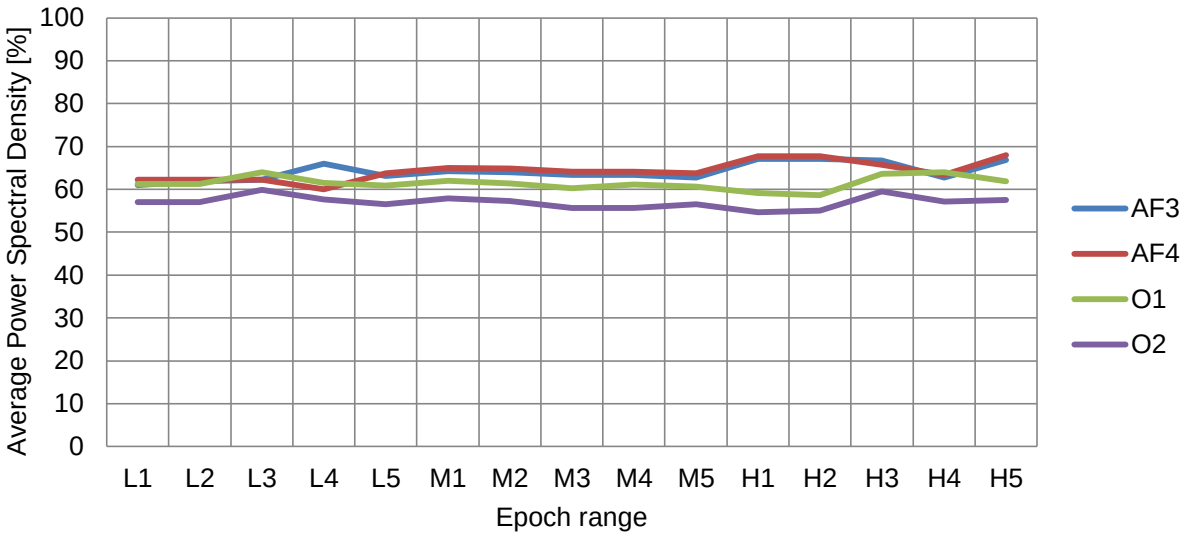


Table 22: Subject #3 - Preliminary Experiment #1

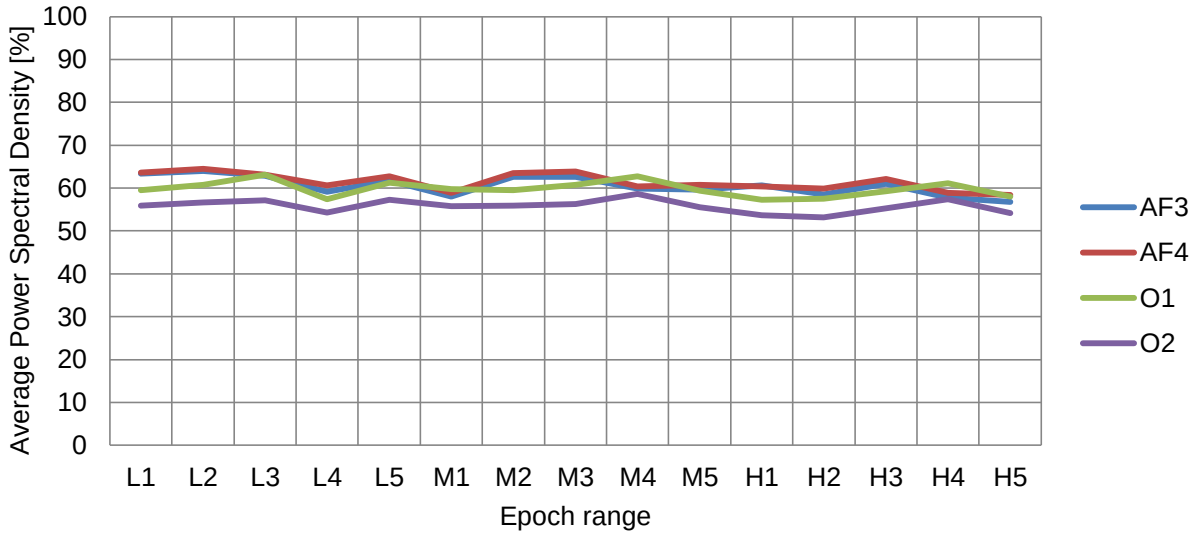


Table 23: Subject #3 - Preliminary Experiment #2

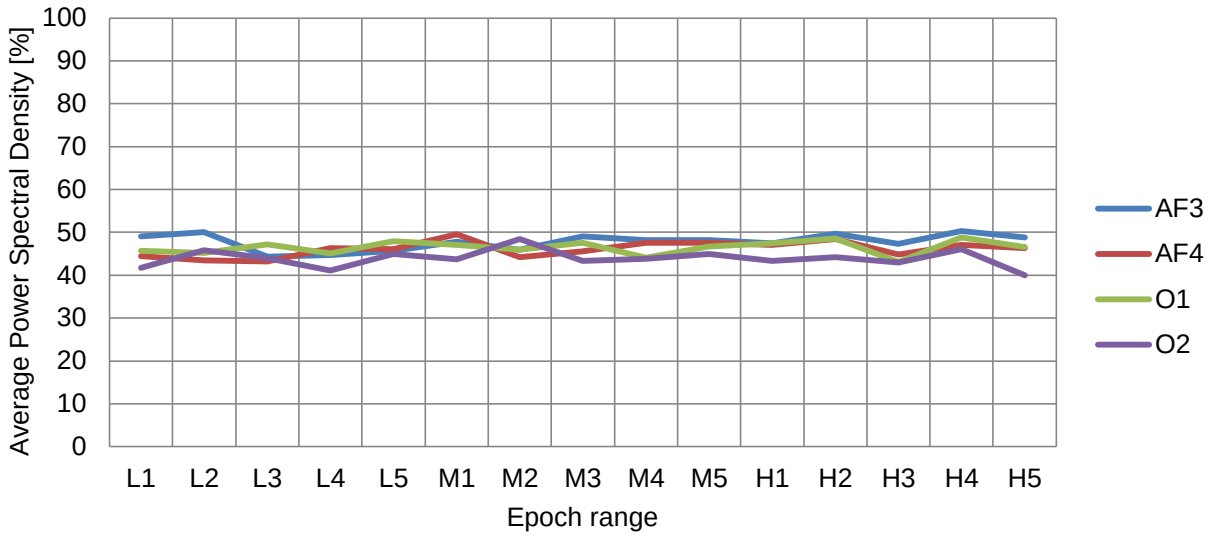


Table 24: Subject #4 - Preliminary Experiment #1

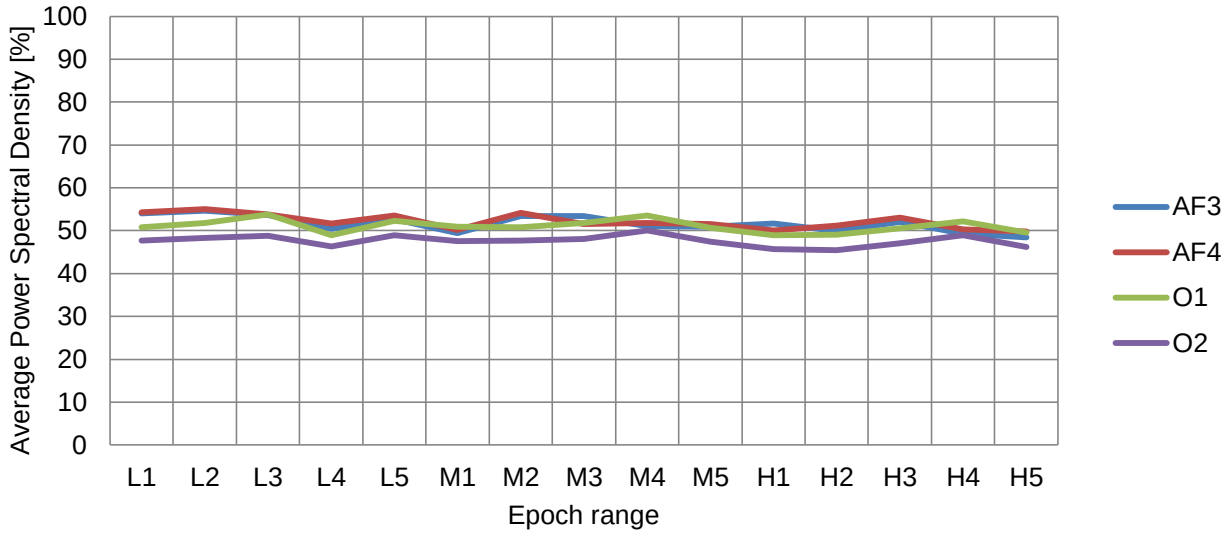


Table 25: Subject #4 - Preliminary Experiment #2

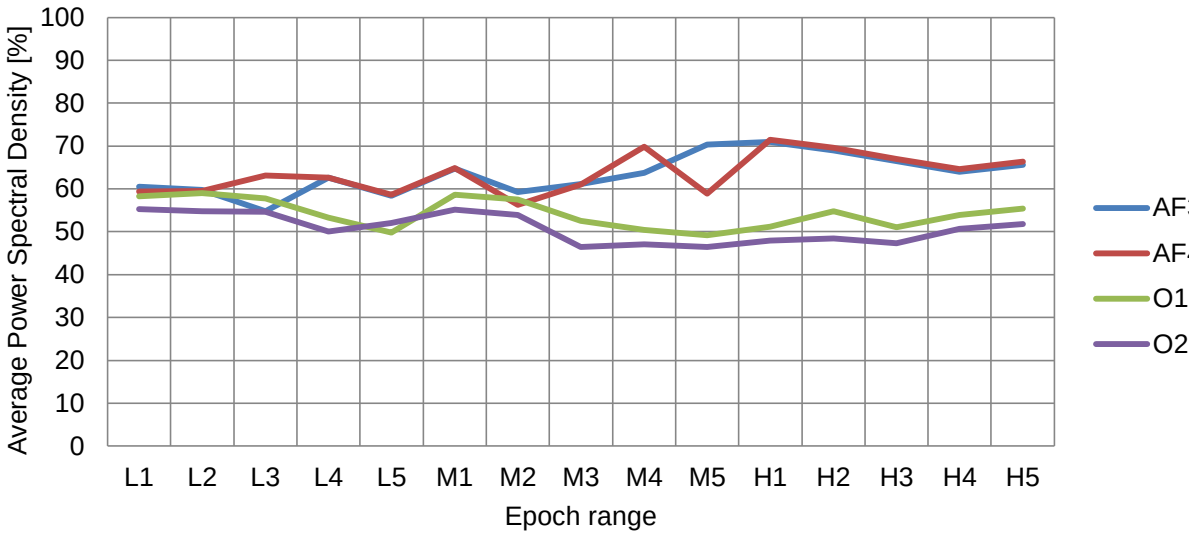


Table 26: Subject #5 - Preliminary Experiment 1

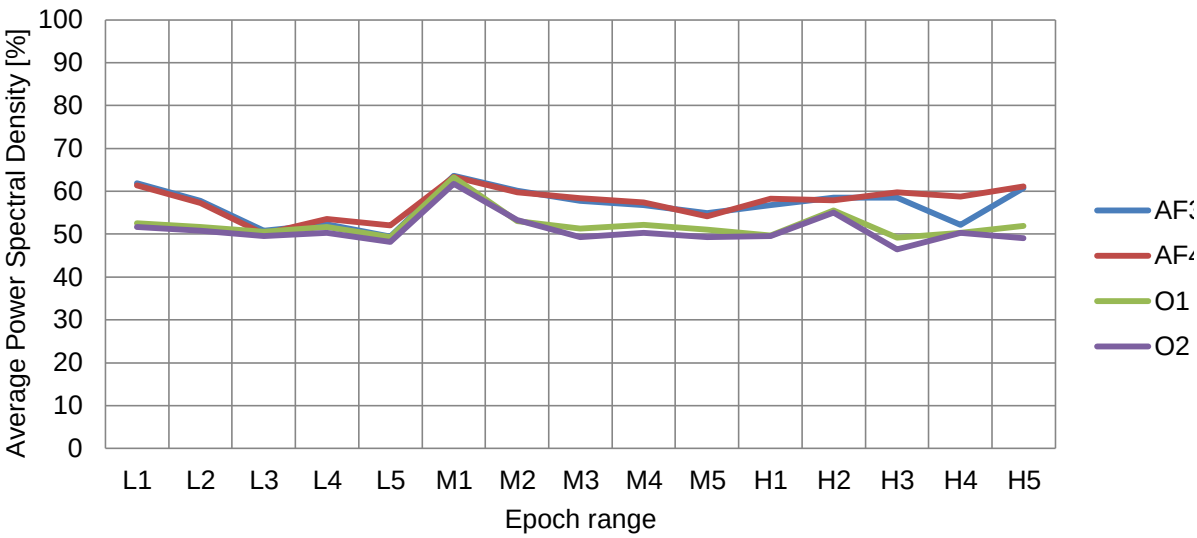


Table 27: Subject #5 - Preliminary Experiment 2

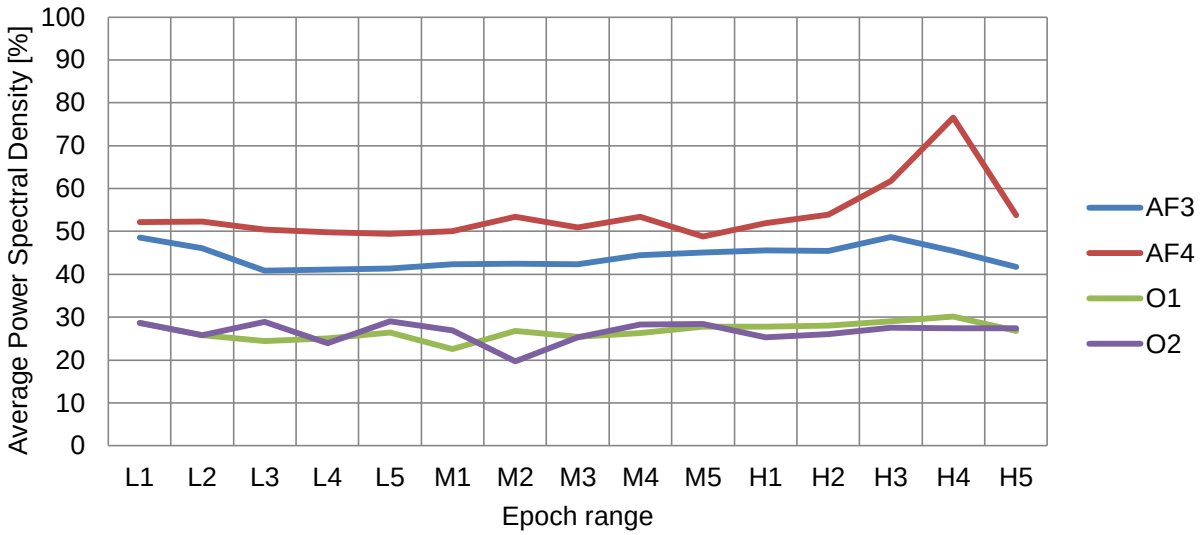


Table 28: Subject #6 - Preliminary Experiment #1

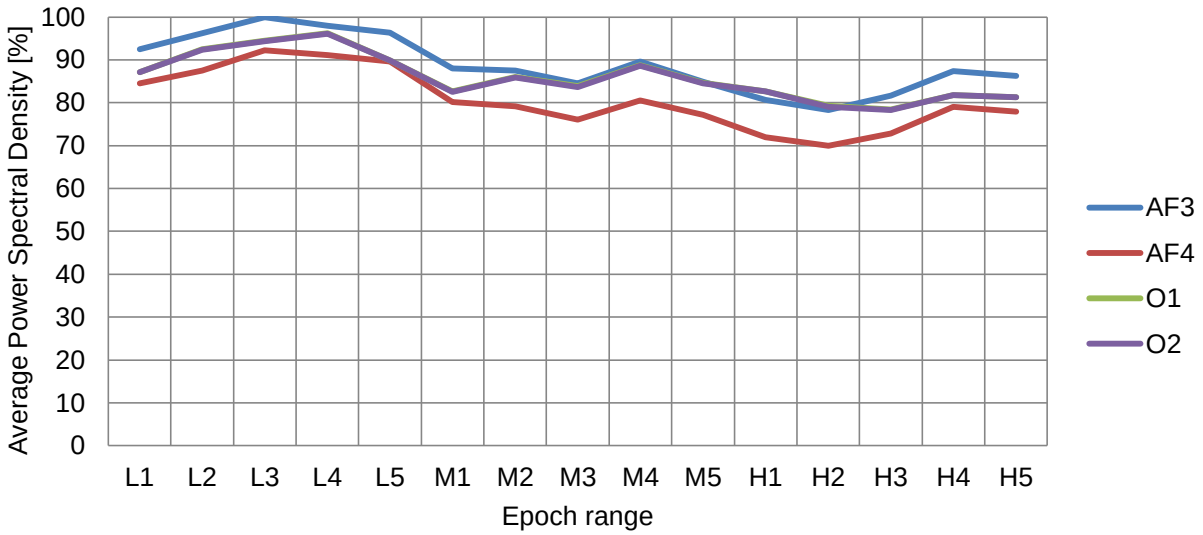


Table 29: Subject #6 - Preliminary Experiment #2

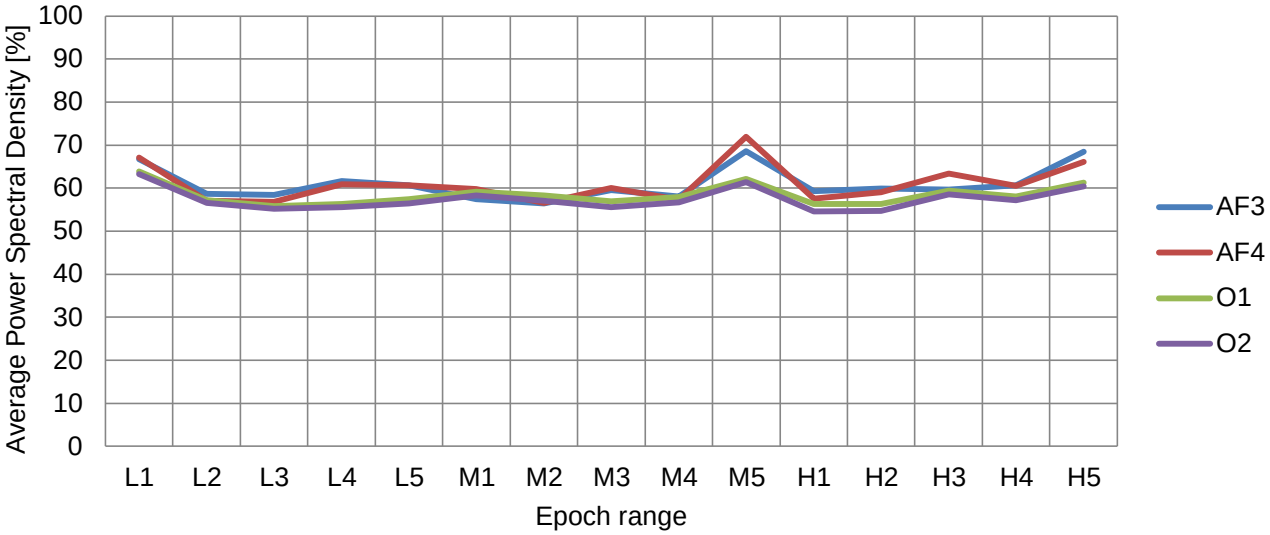


Table 30: Subject #7 - Preliminary Experiment #1

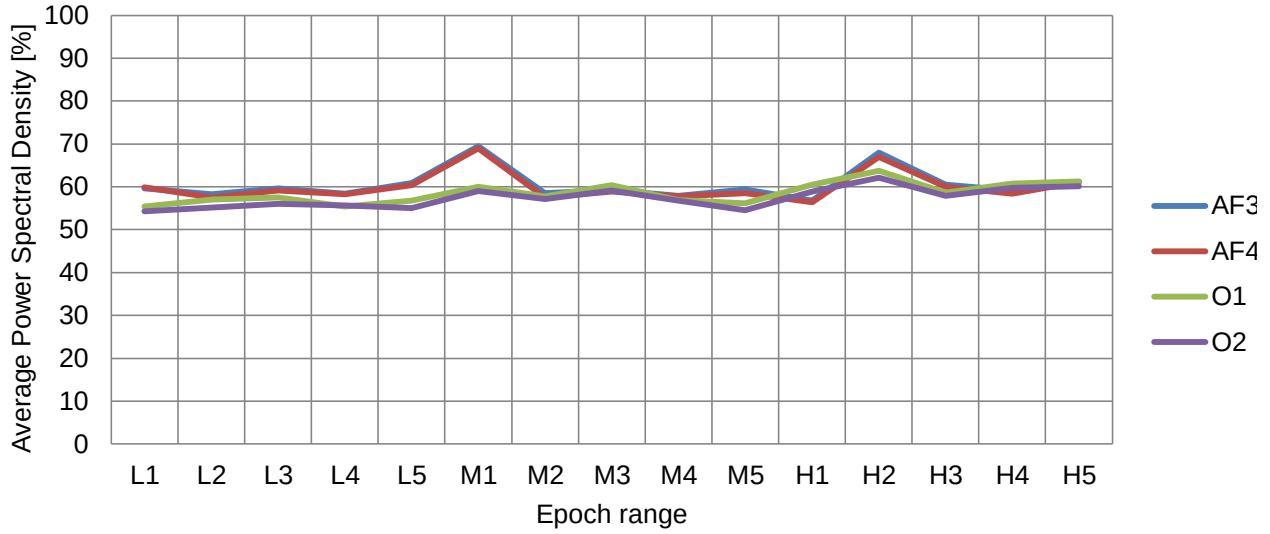


Table 31: Subject #7 - Preliminary Experiment #2

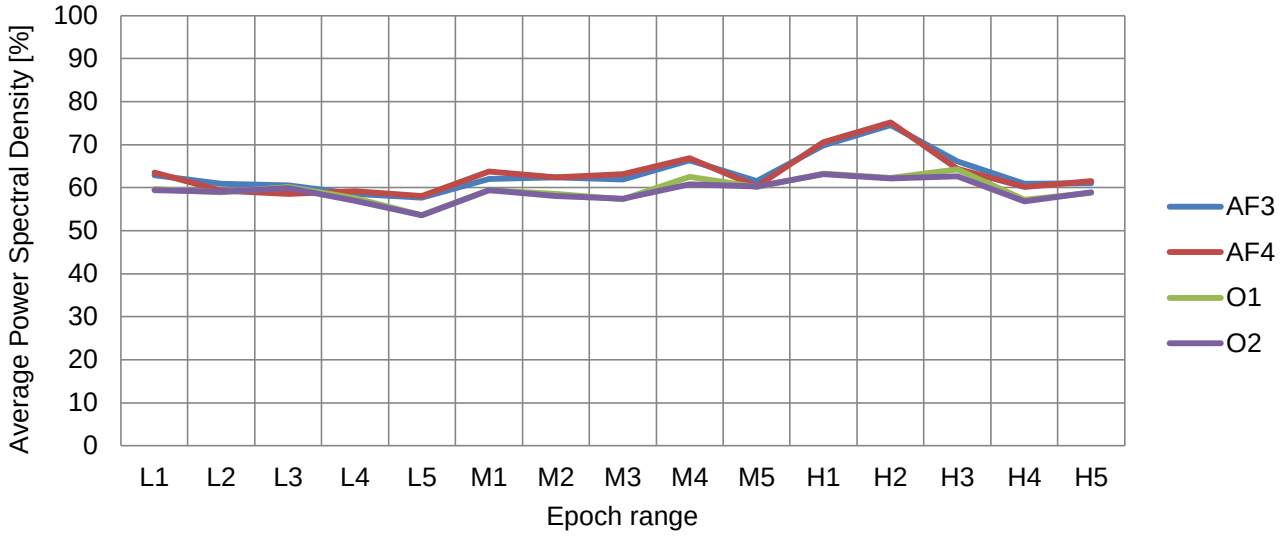


Table 32: Subject #8 - Preliminary Experiment #1

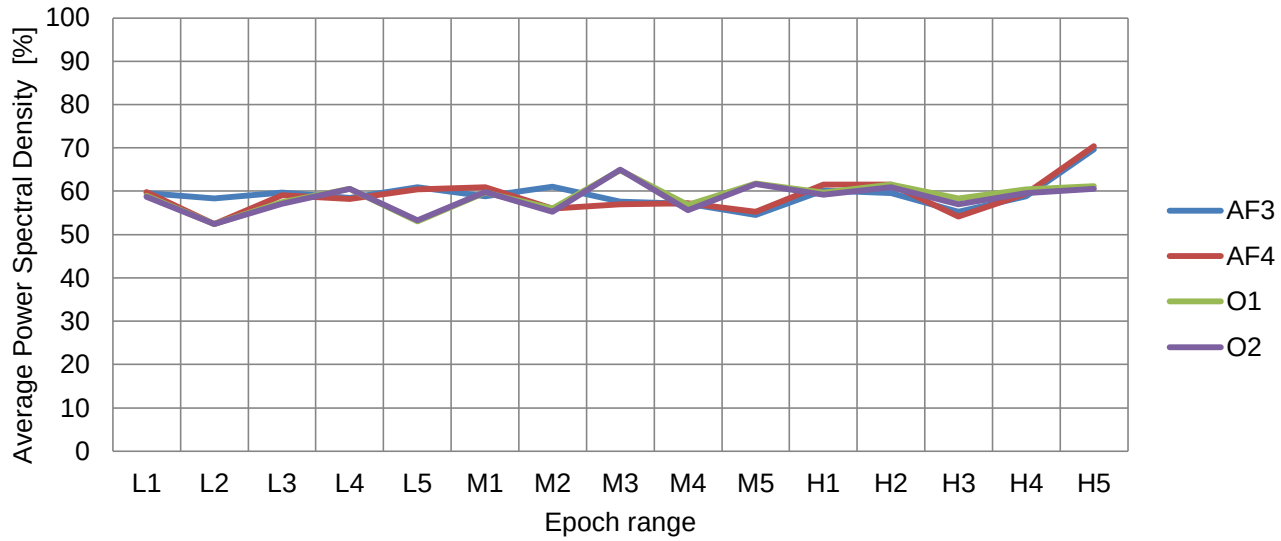


Table 33: Subject #8 - Preliminary Experiment #2

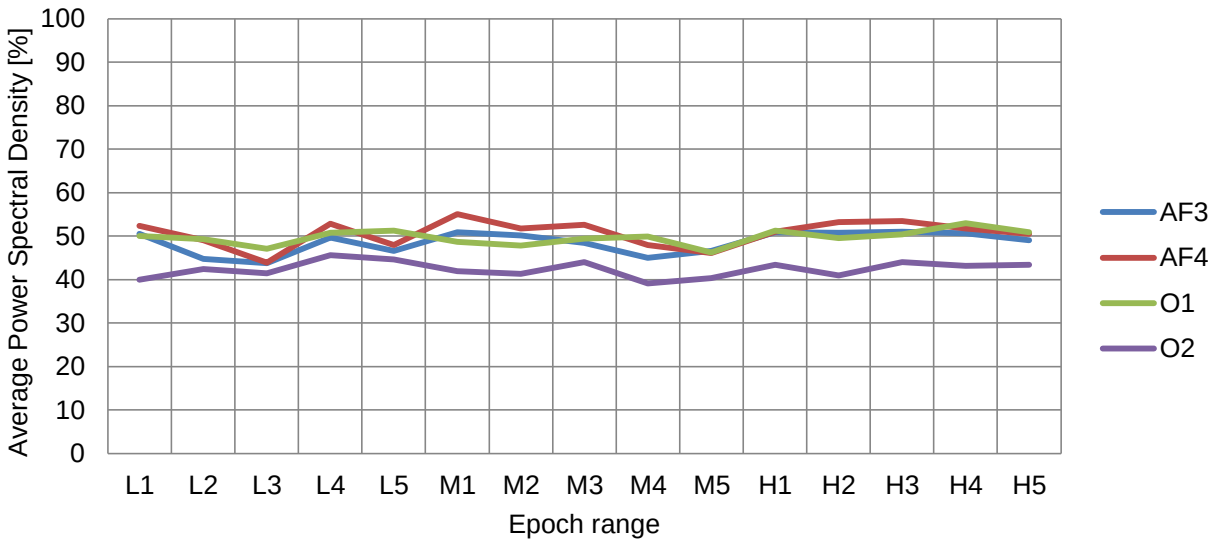


Table 34: Subject #9 - Preliminary Experiment #1

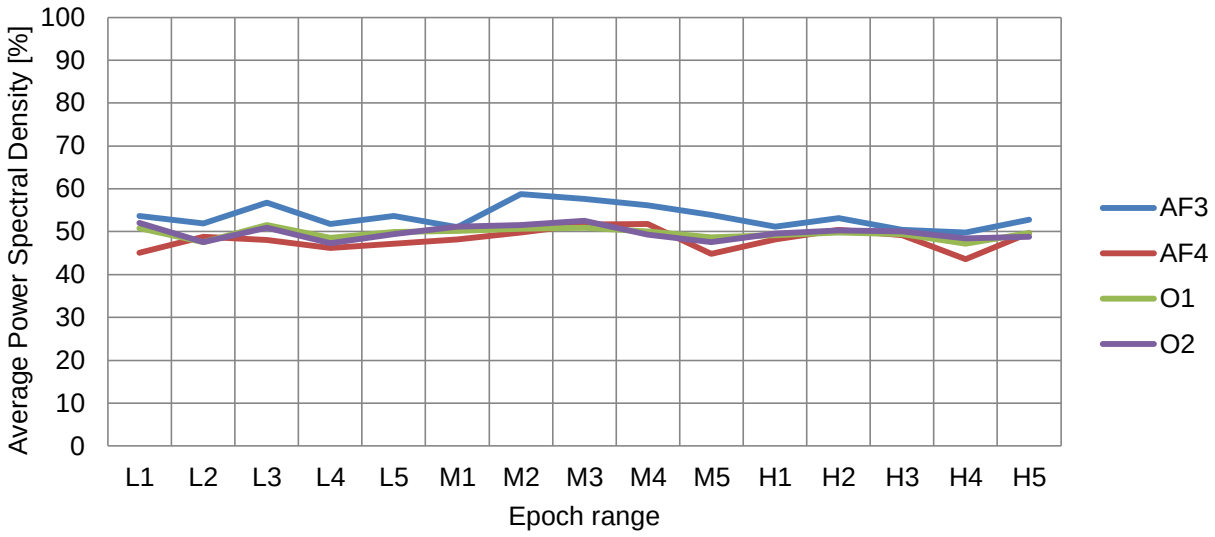


Table 35: Subject #9 - Preliminary Experiment #2

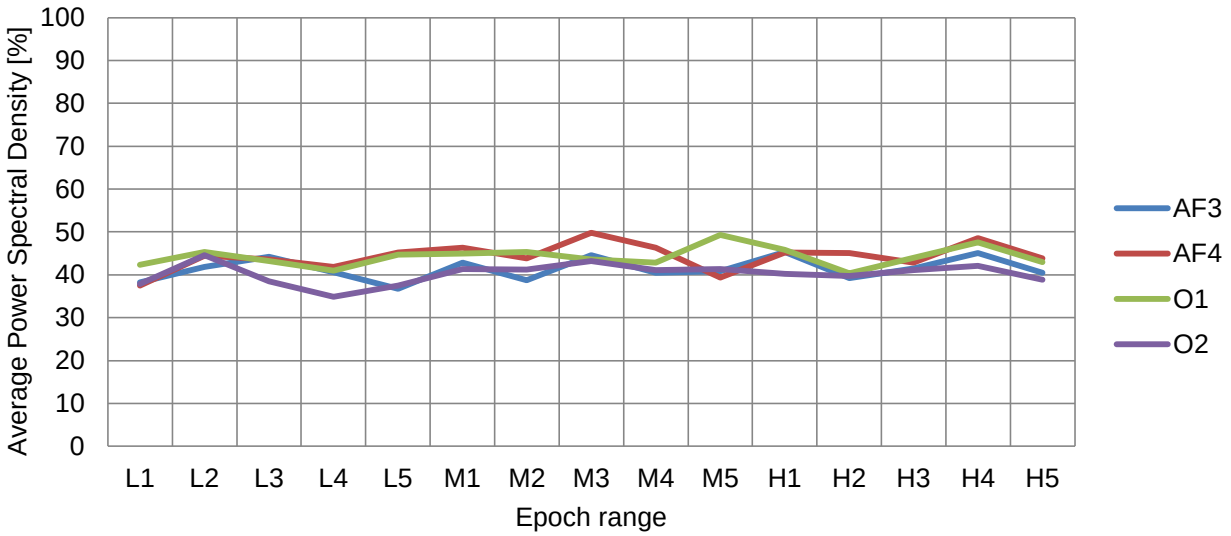


Table 36: Subject #10 - Preliminary Experiment #1

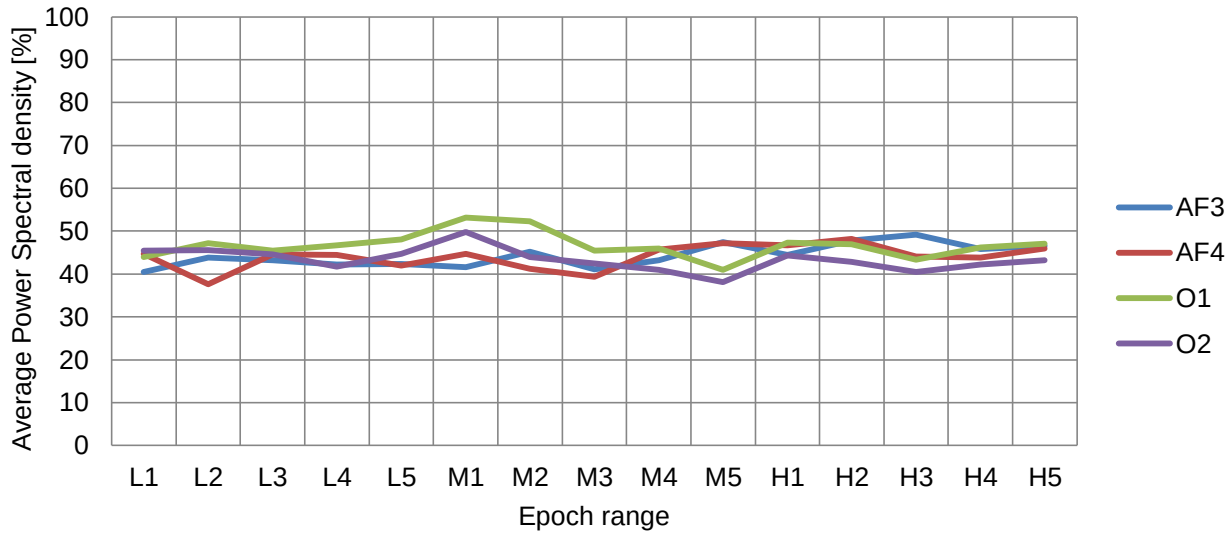


Table 37: Subject #10 - Preliminary Experiment #2

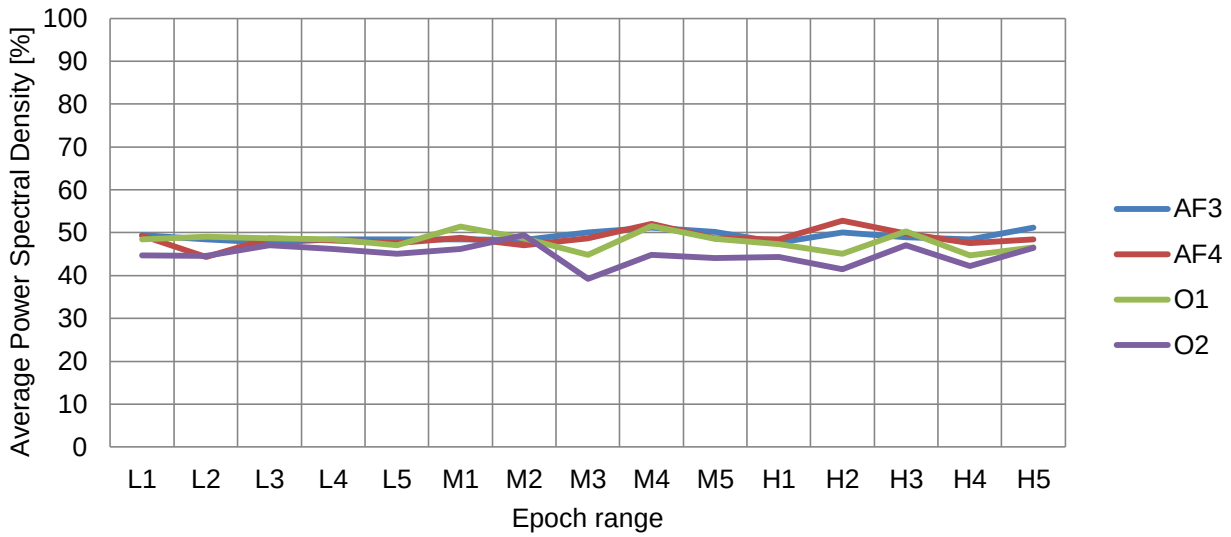


Table 38: Subject 11 - Preliminary Experiement #1

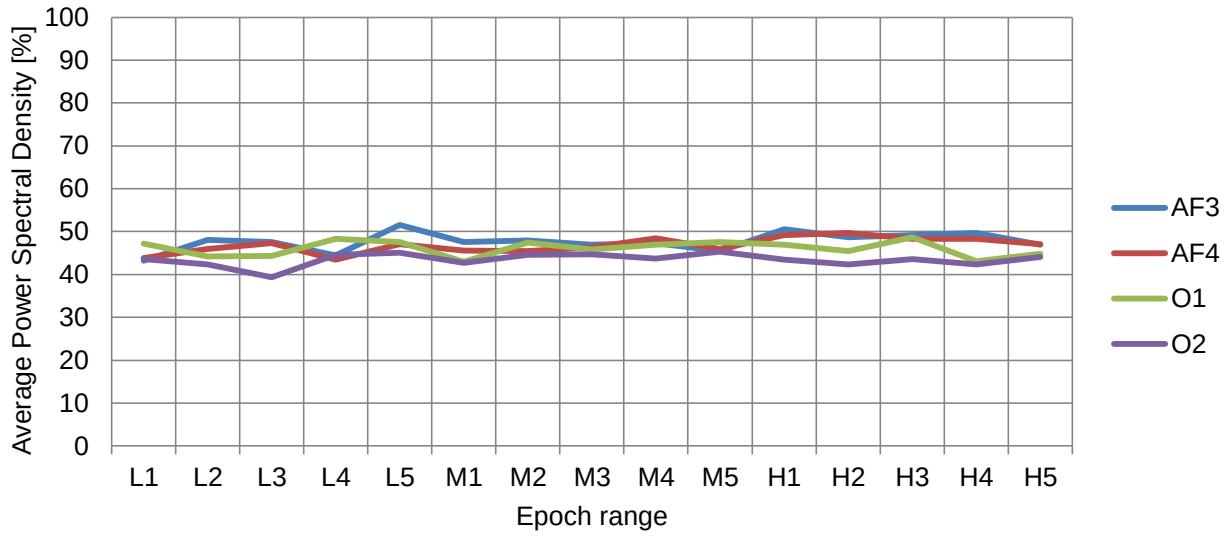
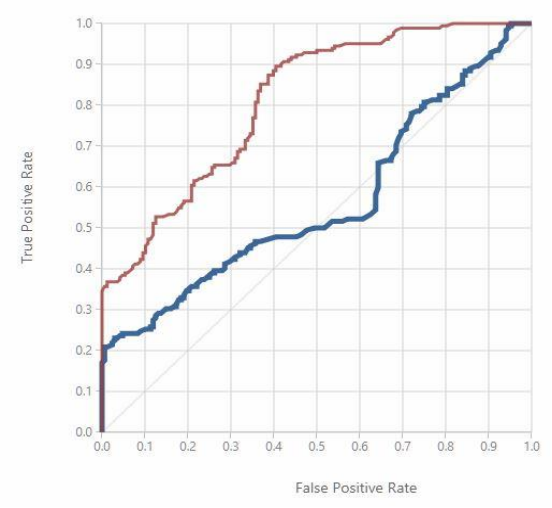
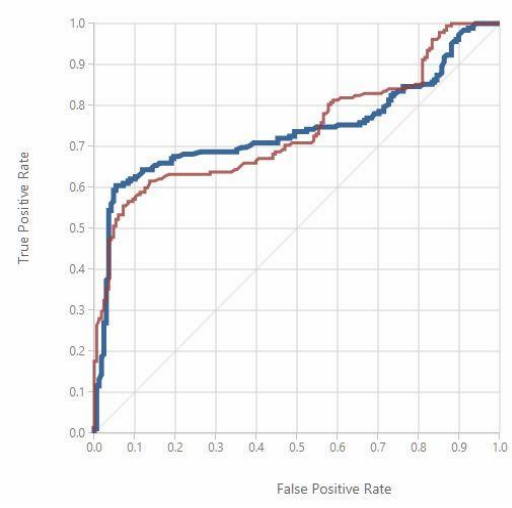


Table 39: Subject 11 - Preliminary Experiment #2

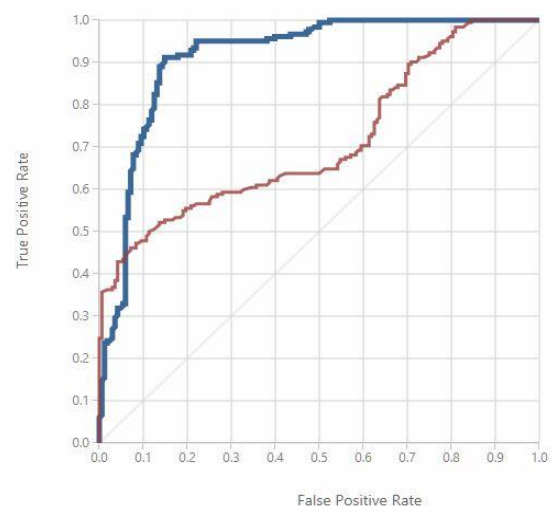
B. ML Prediction Condition



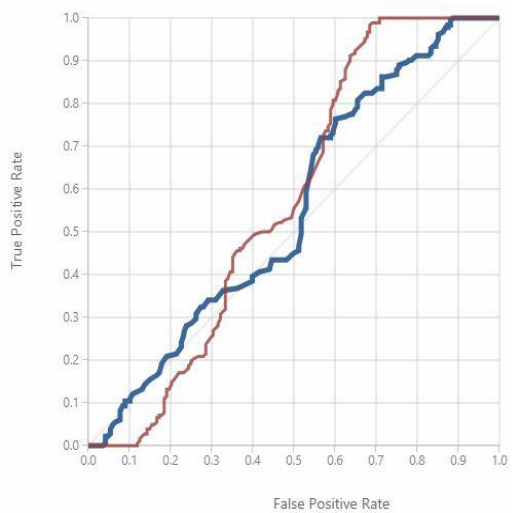
Subject #1 - 1



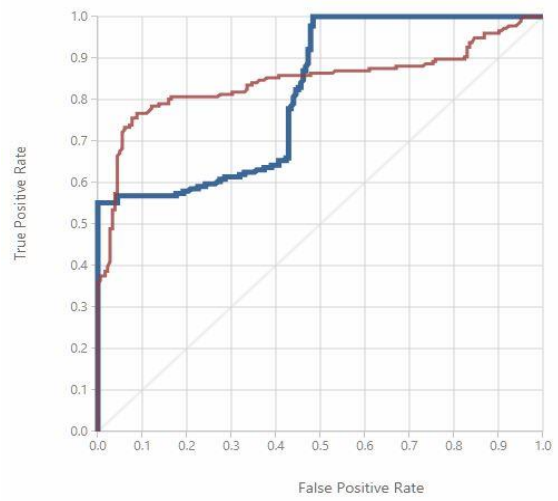
Subject #1 - 2



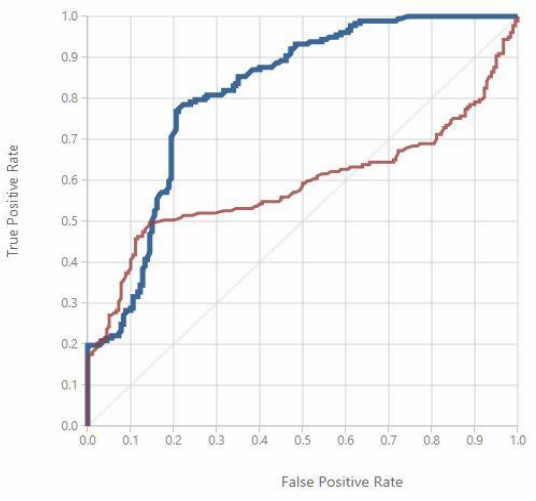
Subject #2 - 1



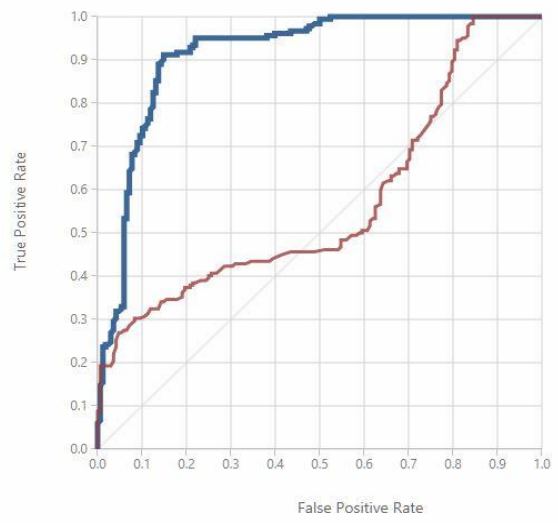
Subject #2 - 2



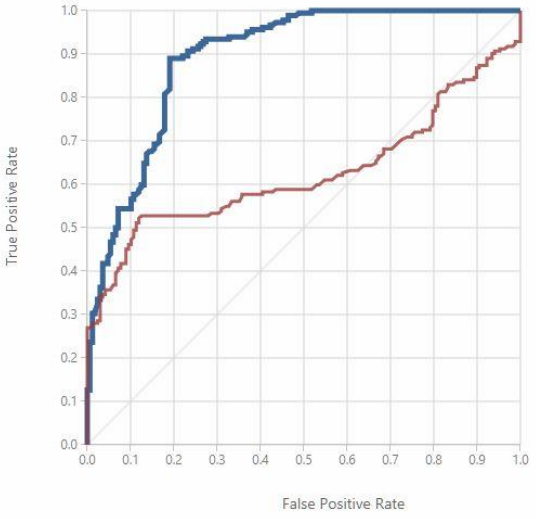
Subject #3 - 1



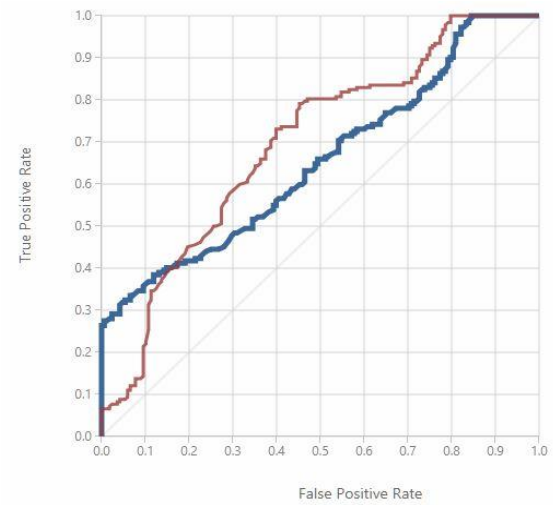
Subject #3 - 2



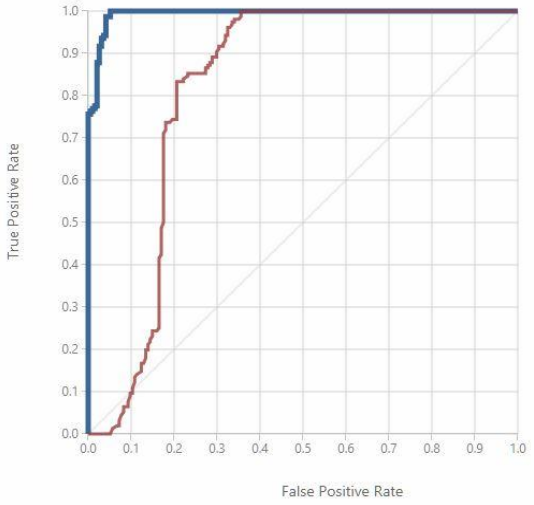
Subject #4 - 1



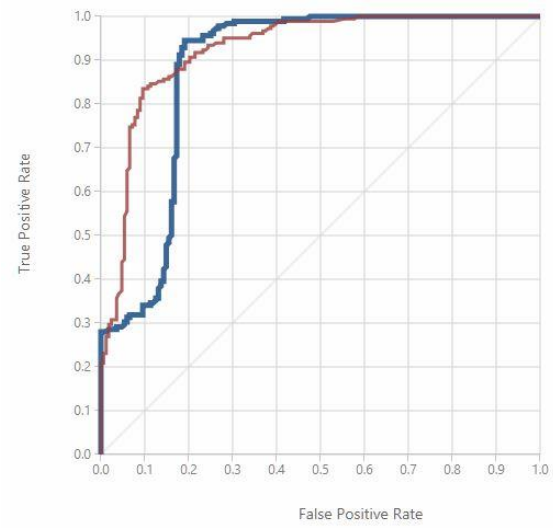
Subject #4 - 2



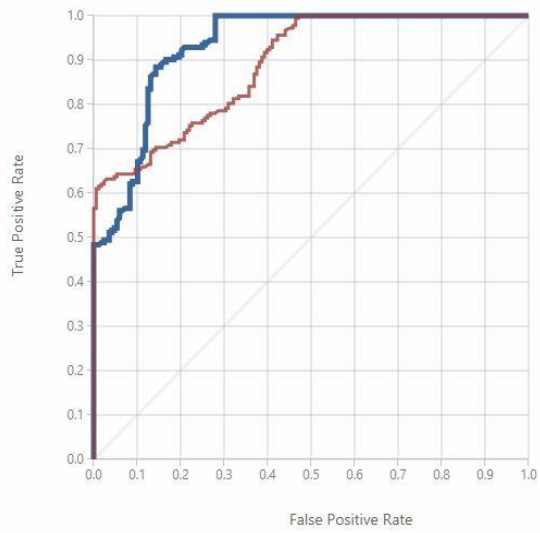
Subject #5 - 1



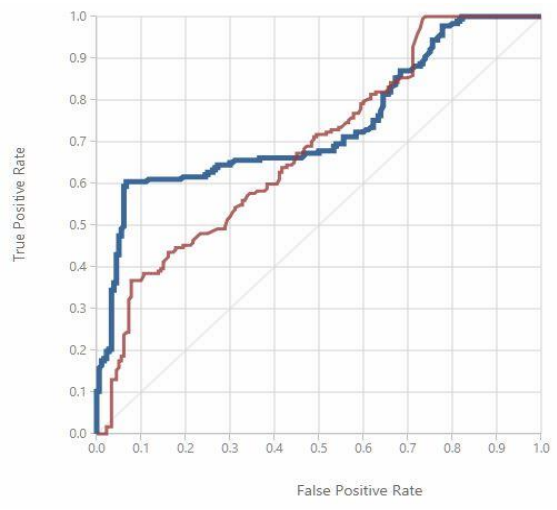
Subject #5 - 2



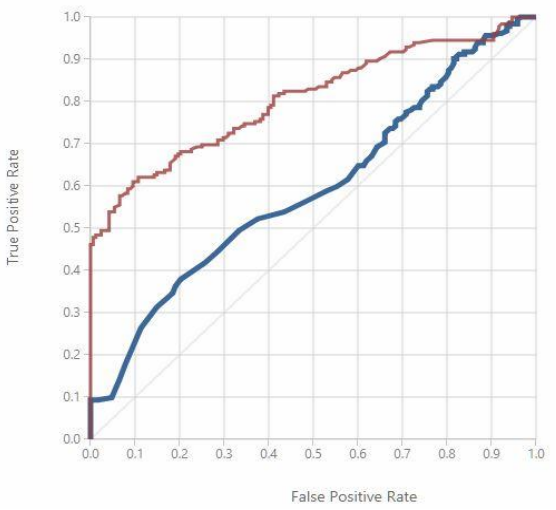
Subject #6 - 1



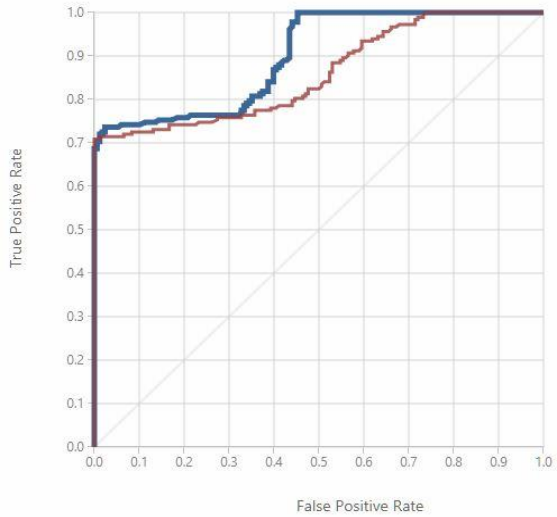
Subject #6 - 2



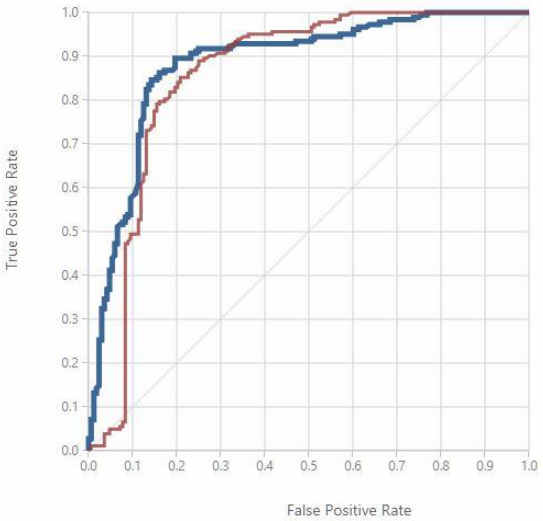
Subject #7 - 1



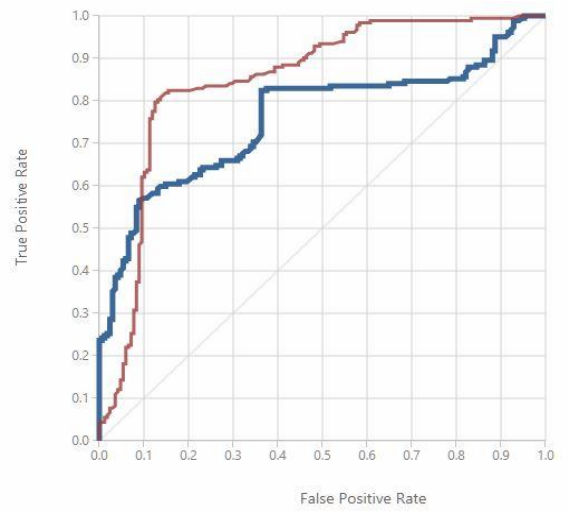
Subject #7 - 2



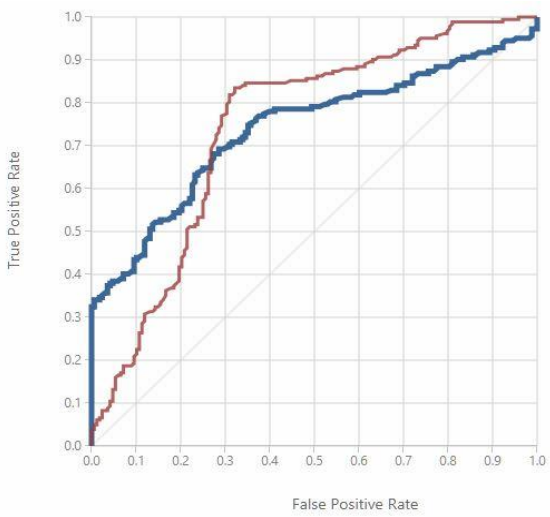
Subject #8 - 1



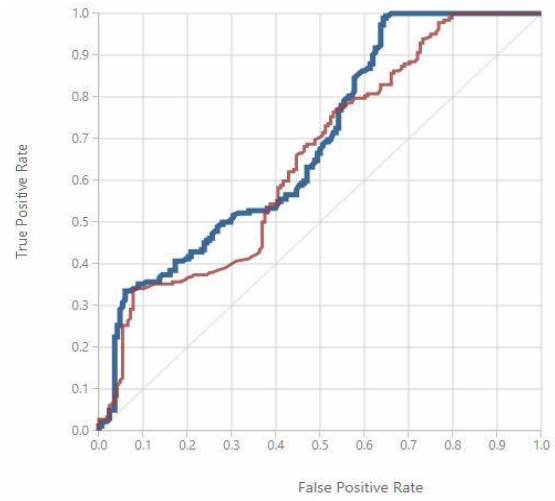
Subject8 - 2



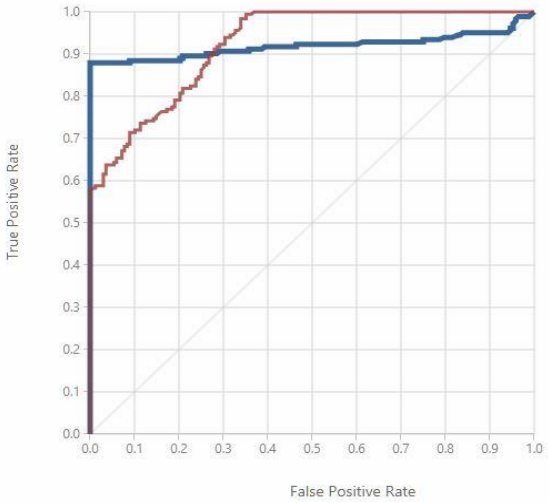
Subject #9 - 1



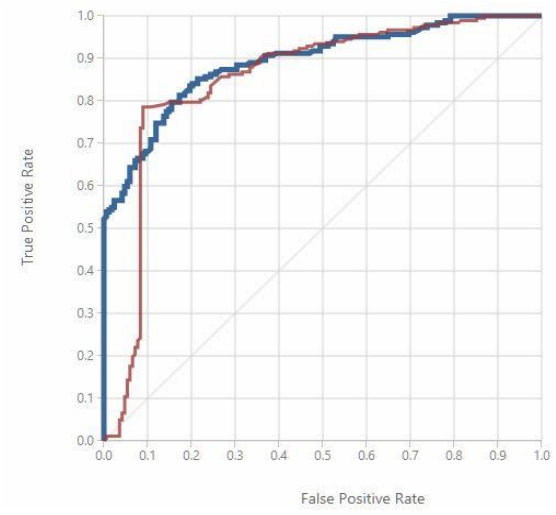
Subject#9 - 2



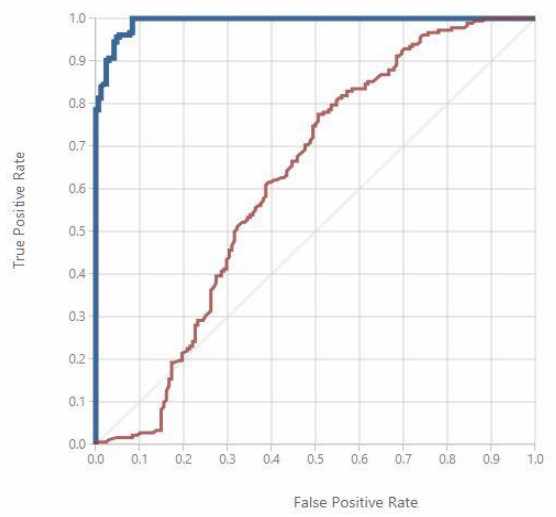
Subject#10 - 1



Subject#10 - 2

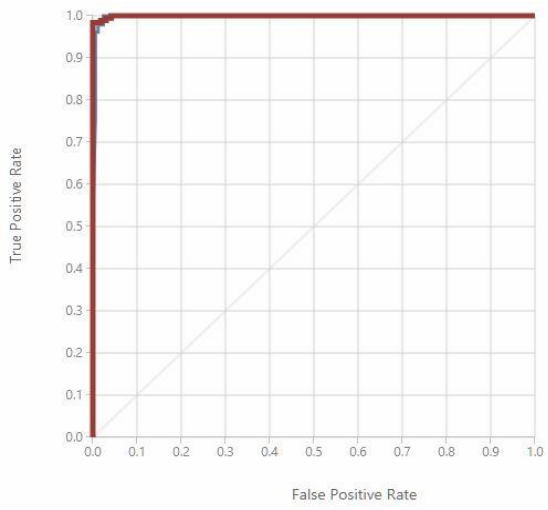


Subject #11 - 1

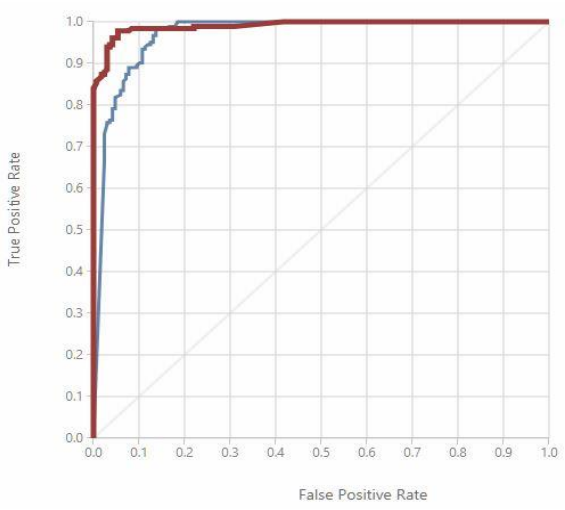


Subject #11 - 2

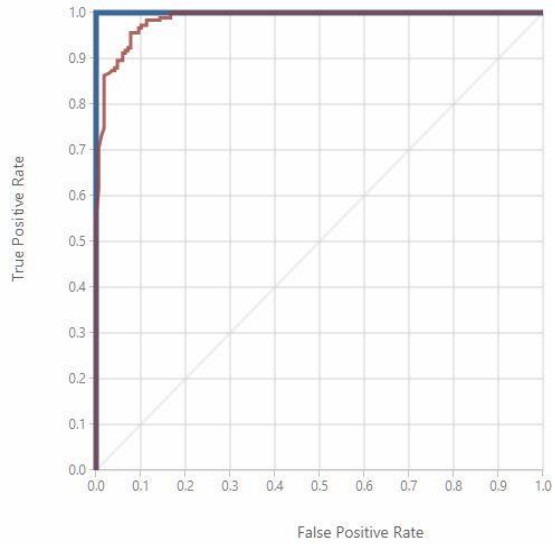
C. Improved ML Prediction Condition



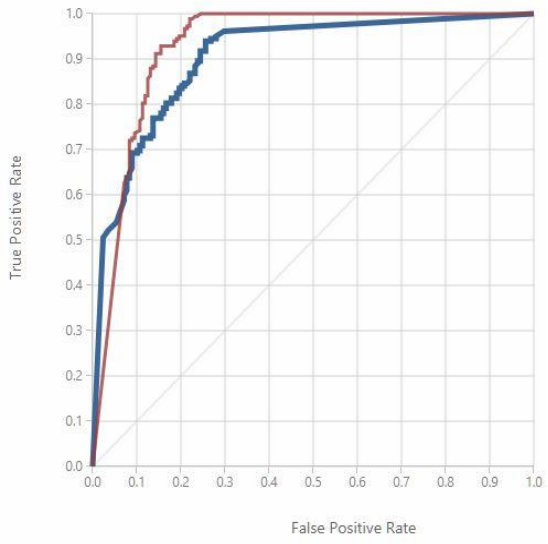
Subject #1 - 1



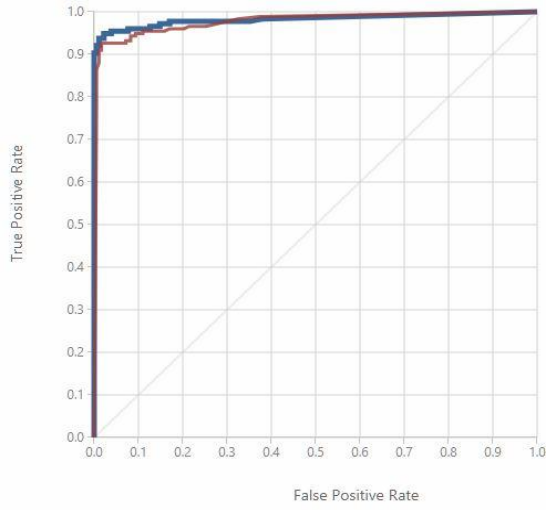
Subject #1 - 2



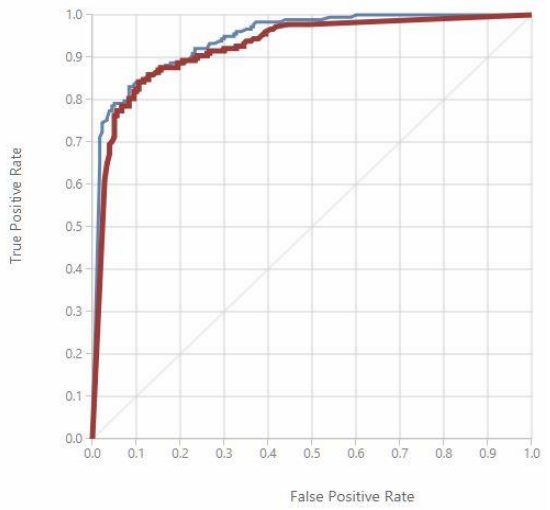
Subject #2 - 1



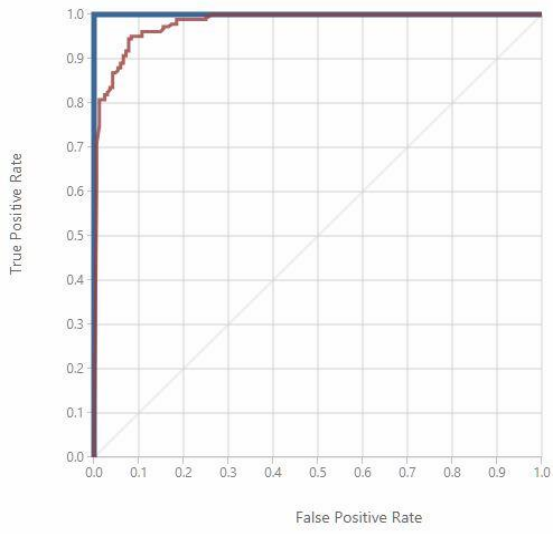
Subject #2 - 2



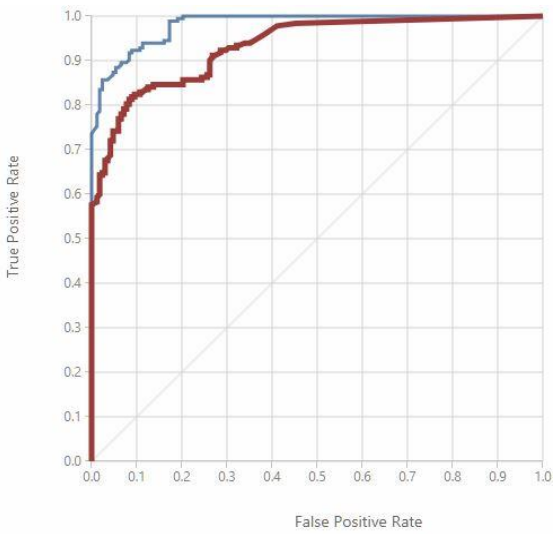
Subject #3 - 1



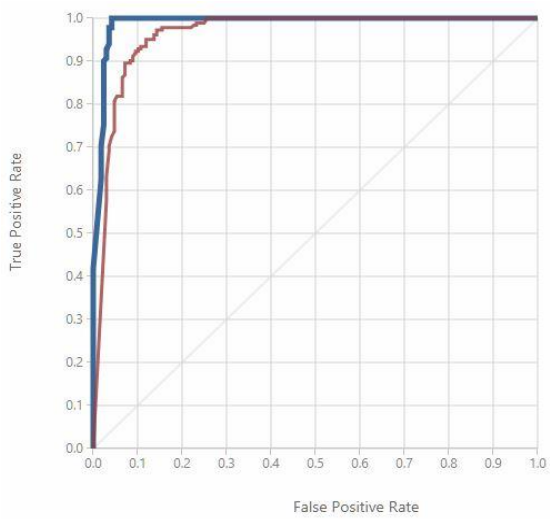
Subject #3 - 2



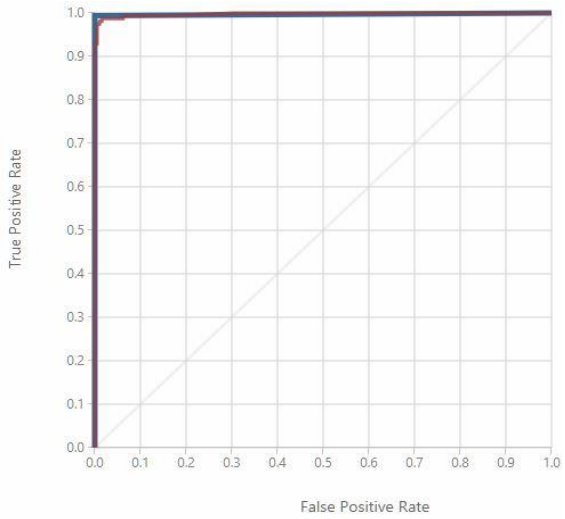
Subject #4 - 1



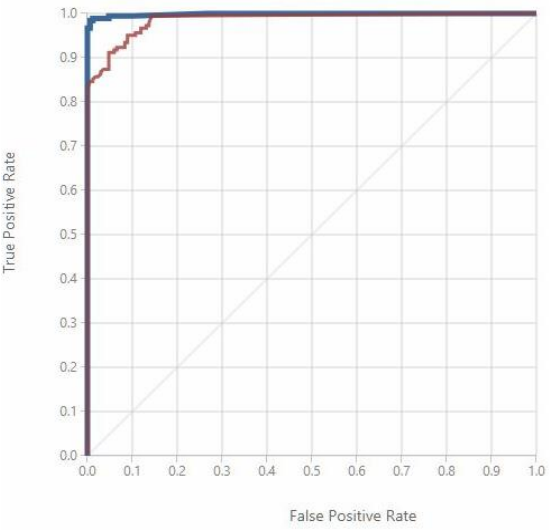
Subject #4 - 2



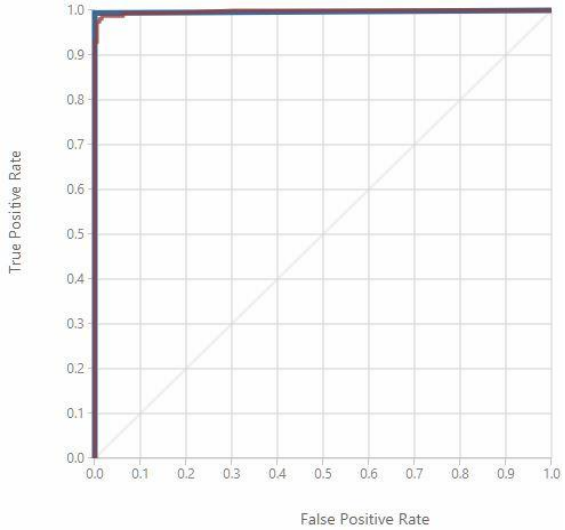
Subject #5 - 1



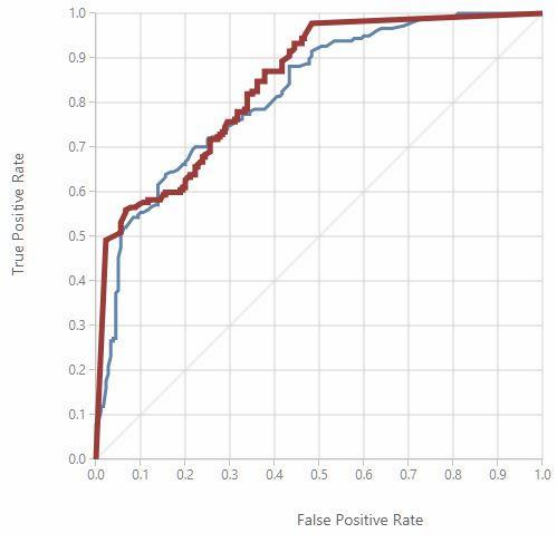
Subject #5 - 2



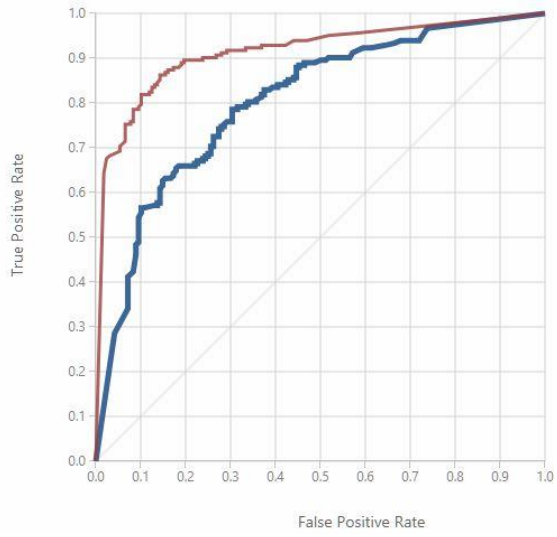
Subject #6 - 1



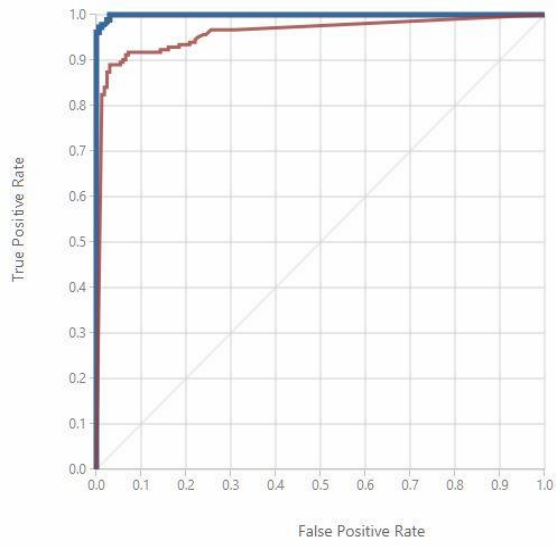
Subject #6 - 2



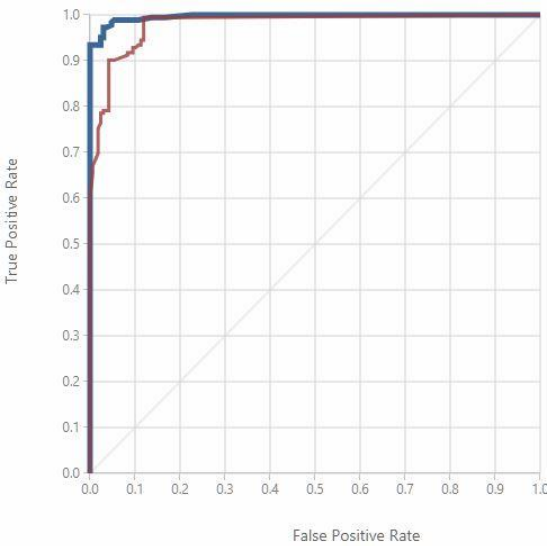
Subject #7 - 1



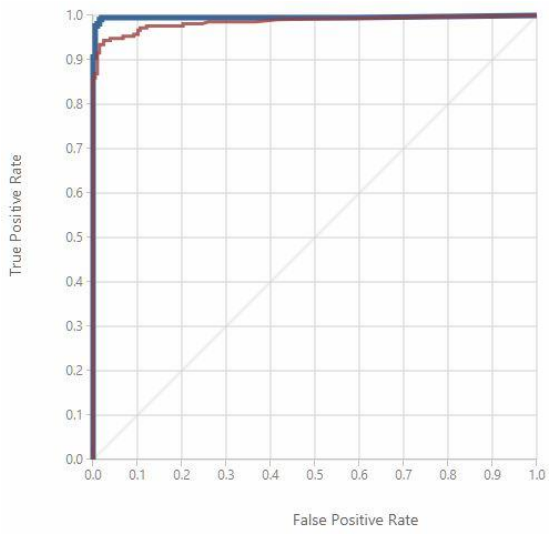
Subject #7 - 2



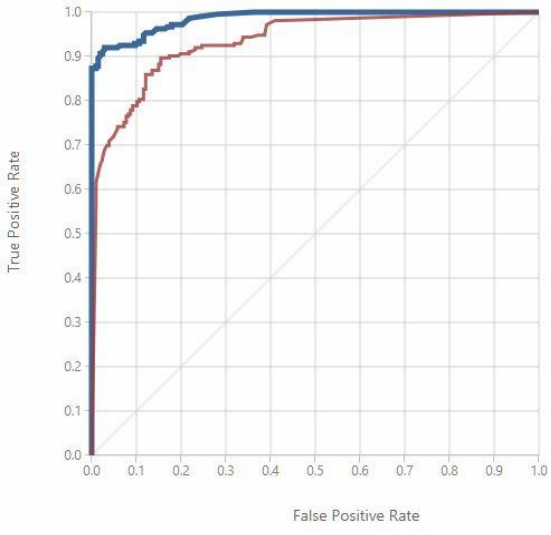
Subject #8 - 1



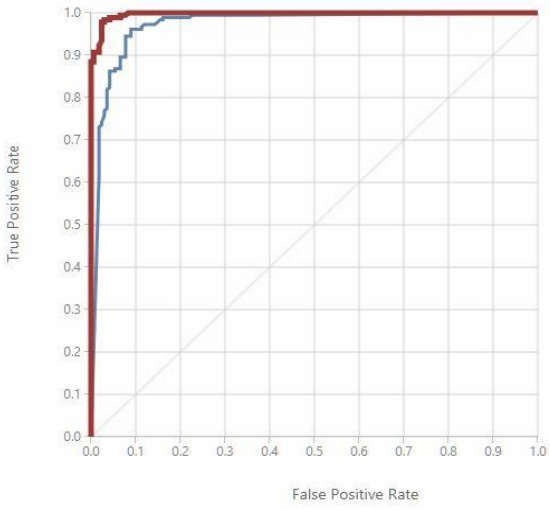
Subject #8 - 2



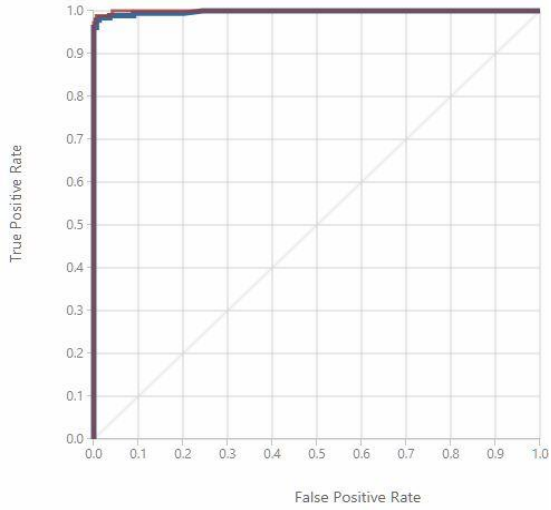
Subject #9 - 1



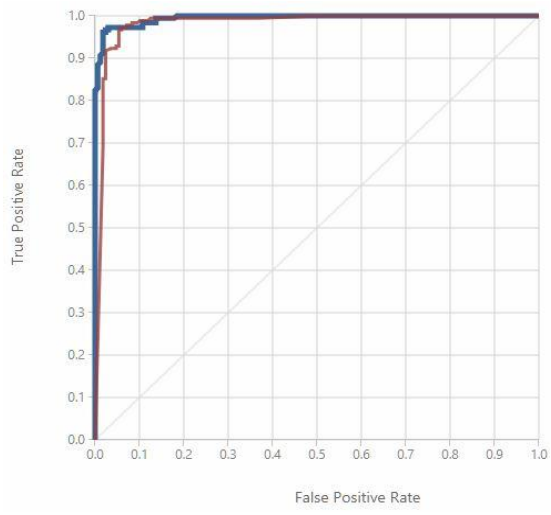
Subject #9 - 2



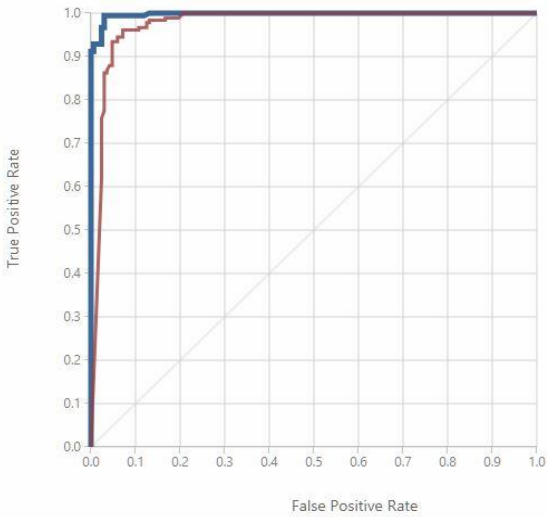
Subject #10 - 1



Subject #10 - 2



Subject #11 - 1



Subject #11 - 2