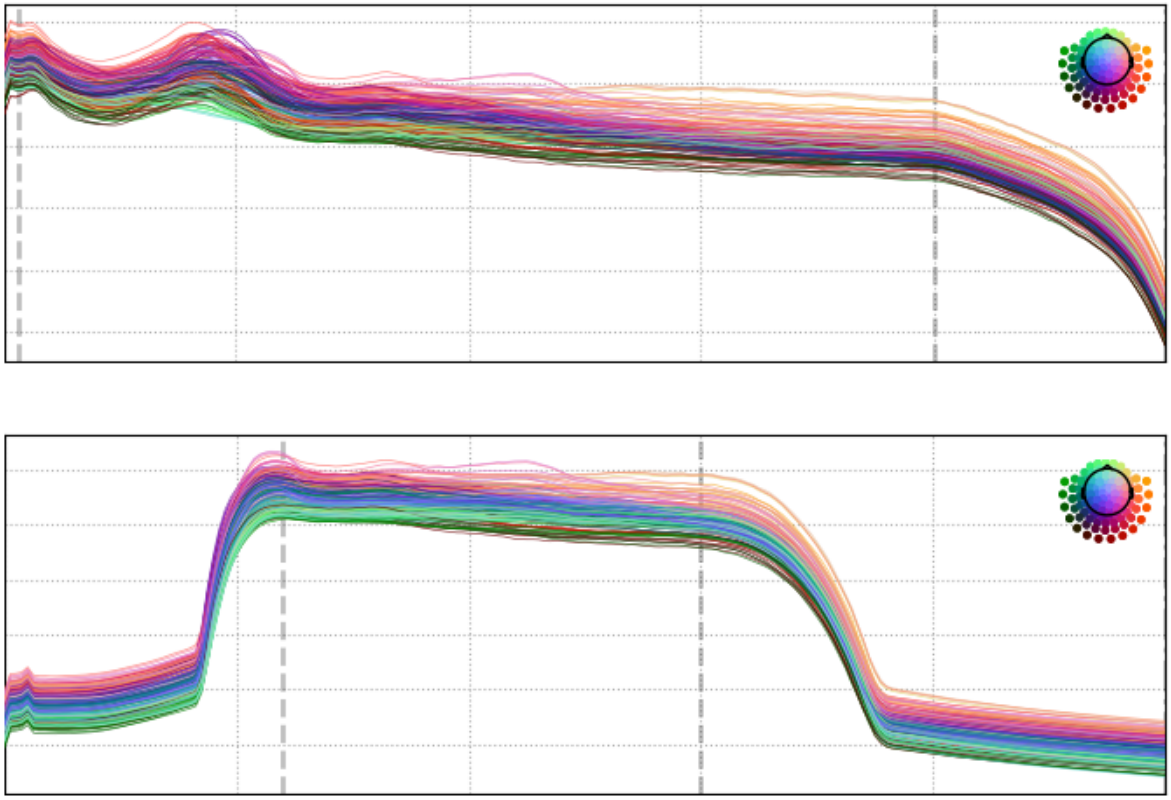
CHALMERS
UNIVERSITY OF TECHNOLOGY



Guiding AI-based classification: Can conventional functional neuroimaging analysis improve deep learning methods for identifying risk for essential hypertension?

Master's thesis in Biomedical Engineering

Christina Chau and Emma Nordanger

DEPARTMENT OF ELECTRICAL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2021
www.chalmers.se

Master's thesis 2021

Guiding AI-based classification: Can conventional functional neuroimaging analysis improve deep learning methods for identifying risk for essential hypertension?

CHRISTINA CHAU
EMMA NORDANGER

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Chalmers University of Technology
Gothenburg, Sweden 2021

Guiding AI-based classification: Can conventional functional neuroimaging analysis improve deep learning methods for identifying risk for essential hypertension?
Christina Chau
Emma Nordanger

Supervisor: Justin Schneiderman, MedTech West and the Department of Clinical Neuroscience
Supervisor: Alice Deimante Neimantaite, Syntronic Research and Development AB
Examiner: Paolo Monti, Department of Electrical Engineering

Master's Thesis 2021
Department of Electrical Engineering
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Power spectral density of MEG data recorded by the gradiometers in the MEG helmet. Upper figure displays the MEG data filtered around 0.5 Hz and 40 Hz. Lower figure displays the MEG data filtered around 12 Hz and 30 Hz.

Typeset in LaTeX
Gothenburg, Sweden 2021

Abstract

Hypertension is a highly prevalent disease and a major risk factor for cardiovascular diseases. Therefore, it is important to find methods to identify the risk of developing hypertension at an early stage in order to work in a preventive manner. It has been shown that the invasively-measured muscle sympathetic nerve activity (MSNA) response to stressful stimuli is likely to predict this risk. Furthermore, a strong correlation between MSNA and beta oscillations in the brain has recently been identified. With these known correlations, the neural activity in the beta band could potentially be used as a non-invasive biomarker recorded with magnetoencephalography (MEG) for evaluating the risk of developing hypertension.

The aim of this thesis is to investigate how the learning and classification of two previously developed neural networks can be improved with conventional MEG analysis methods. We attempted to give emphasis to the beta oscillations by filtering out frequencies outside this beta band. Multiclass classification was also tested on different groupings of subjects in the MEG study based on the level of risk, i.e., high risk, medium risk etc. The resulting accuracies were generally lower for the filtered data, indicating that the filtering had removed some aspect in the data that the network had previously learned form, perhaps some characteristics of the MSNA-related brain response profile. Accuracies improved for multiclass classification but were still deemed to be essentially equivalent to previous results. Thus, pure frequency-based filtering does not appear to achieve the desired effect, which could be because the AI methods used are not designed to identify frequency-based changes in power. Dividing the subjects into multiple classes also failed to improve classification, which is likely due to the limitation of having a low number of individuals in the MEG study. We conclude with a framework for potentially overcoming such challenges by training the networks on the data after transforming it into the time-frequency domain.

Keywords: Deep learning, Magnetoencephalography, Muscle sympathetic nerve activity, Beta oscillations, Preprocessing, Brain activity, Frequency filtering, Binary classification, Multiclass classification.

Acknowledgements

Throughout writing this thesis we have learned many new things and have improved in many ways. This would not have been possible without the help of many people.

We would like to start by showing our gratitude to our main supervisor at Medtech West, namely Justin Schneiderman. Who expertly guided us on this journey and brought this thesis to a higher level. His encouragement and insights where especially valuable and gave us confidence in our own abilities.

The assistance and expertise provided by Alice Deimante Neimantaite and Lisa Sjöblom at Syntronic is also greatly appreciated. We thank them for their knowledge and readiness to help, which was invaluable to this thesis.

We also wish to express our gratitude to our advisor Mohammad Hossein Moghaddam and examiner Paolo Monti for taking on this task and helping us finalize our work.

A special thank you to the previous works that laid the foundation for this project. To Anna Bakidou and Jonas Nord Odhner for a clear and structured thesis that made it fun and easy to follow. To Bushra Riaz Syeda, who generously took the time to explain the data set and share ideas and code for time-frequency analysis with us.

Additionally, we would like to thank our close friends, families and partner who supported and helped us find our way throughout our years at Chalmers.

Christina Chau and Emma Nordanger, Gothenburg, June 2021

Declaration of Contributions

The current master's thesis is a collaboration between Medtech West and Syntronic and is a direct continuation of the work by Bakidou and Nord Odhner. They developed the source code and data structure which was used when designing, optimizing and training two neural networks for the purpose of classifying the risk of developing essential hypertension. For further understanding of these networks we refer to [1]. Data acquisition and preprocessing for the data set that was used for both theses was performed by Bushra Riaz Syeda in her PhD dissertation [2]. Further development of the data structure, training of the networks and preprocessing was performed by Emma Nordanger and Christina Chau.

Contents

List of Figures

List of Tables

List of Acronyms

AI Artificial intelligence. 1–3, 10

ANNs artificial neural networks. III, 2, 9, 10, 30, 33

ANS autonomic nervous system. 5

CNN convolutional neural network. 30, 31

ENS enteric nervous system. 5

LSTM long short-term memory. 2, 9, 17–21, 23, 25, 29

MEG magnetoencephalography. I, II, 1–4, 7, 8, 13, 17, 27, 29–31, 33

MLP multilayer perceptron. 2, 9, 17, 20, 21, 23, 24, 26, 29

MSNA muscle sympathetic nerve activity. II, 1, 2, 6, 7, 12–16, 19, 21, 27–31, 33

PNS parasympathetic nervous system. 5, 6

SNS sympathetic nervous system. 5, 6

SQUID superconducting quantum interference device. 8

1

Introduction

Hypertension is a common disease with a high prevalence in the world, causing both morbidity and mortality as it is a major risk factor for cardiovascular disease [3]. There is no established clinical predictor for development of hypertension and it would be meaningful if such could be identified for prevention or treatment purposes. Indeed, many people experience repeated stressful input in everyday situations, such as inputs from traffic, construction or other demanding situations of modern society, which is one of the contributing factors for the development of hypertension [4], [5].

During a stressful or arousal input, the muscle sympathetic nerve activity (MSNA), which is a vasomotor activity controlling constriction of blood vessels, is involved and this response to arousal input differs for each person. It can either be an inhibition or a non-inhibition of the MSNA response, leading to a less elevated blood pressure for the former and an elevated blood pressure for the latter [6]. Having a non-inhibited MSNA response would be a disadvantage in today's modern environment, that include many stressful or arousal-inducing factors, as the frequently elevated blood pressure could lead to hypertension and subsequent cardiovascular disease. Currently, MSNA is measured with the delicate and invasive neurophysiological method, microneurography. This method is both invasive and time consuming which makes it not suitable for large scale clinical studies [7]. For this reason, it has given rise to the interest of finding alternative methods. In a previous study, it was found that beta oscillations in response to stressful or arousal stimuli are correlated to MSNA [2]. The beta oscillations represent brain activity in the frequency range of 13 - 30 Hz and can be measured with magnetoencephalography (MEG) which is a non-invasive method with high temporal and spatial resolution. This means that the beta oscillations can potentially be used as non-invasive biomarker for the MSNA response and subsequently be used to estimate the risk of developing hypertension.

Artificial intelligence (AI) has potential of being utilized more in clinical settings since these methods are able to manage and classify large amounts of clinical data and information. Therefore, combining AI with conventional MEG analysis methods could potentially be a new method to classify people as inhibitors or non-inhibitors. This has led to a collaboration between MedTech West and Syntronic to develop a classification method using deep learning. As a result of this collaboration, a previous master's thesis was conducted by Bakidou and Nord Odhner [1] with the aim to explore if deep learning could be used on MEG data to predict the risk

of developing essential hypertension. Their results showed that artificial neural networks (ANNs) had potential to be used for classification of the MSNA response. However, other aspects, mainly in how the MEG data is preprocessed, have to be further examined to see if the ANNs can learn the general MSNA-characteristics and not something else from the MEG signal.

1.1 Previous Work

The work herein was based upon the aforementioned master's thesis that investigated AI-based methods for classification of MEG signals gathered during a study on stress and heart disease. The MEG study was aimed at identifying functional neuroimaging-based correlates to a previously developed measure of the physiological response to stress. A correlation was discovered between an individual's brain response (i.e., in the MEG signals) and the physiological response (i.e., MSNA) to sudden stressors, in this case, unexpected electrical stimuli to the finger. An attempt was therefore made to use deep learning to classify subjects with their MEG data according to their MSNA response. To that end, the relevant signal segments were selected around the stimulus-responses in time because it was presumed that it is those segments contained the most relevant information for classification (i.e., as opposed to including the "resting state" MEG data that was recorded in between the stimuli). The data was furthermore broken down in the spatial domain by dividing the MEG sensors into three sets. First one that included sensors that were most likely to detect the activations that were already known to be coupled to the risk profile, one with sensors in a larger area, and lastly one with all sensors. Two ANNs, one long short-term memory (LSTM), and one multilayer perceptron (MLP), were developed to examine classification of MEG signals. Binary classification (i.e., to discriminate "low" from "high" risk individuals) was tested in different ways that included or excluded individuals that were determined to be on extreme and/or borderline values of the risk spectrum. Finally, multiclass classification was also attempted to see whether AI could identify individuals based on their brain responses.

The results of the previous work were mixed. Binary classification performed better than chance, but not nearly as well as that which was achieved via conventional MEG data analysis methods. Multiclass classification performed far better than chance and showed that the MEG signal could be mapped to an individual which indicated that the networks have learned individual features of the subject data that were most probably not related to the MSNA response.

1.2 Aim

The aim of this thesis is to take another step towards classification improvement and understanding of the brain functional activation recorded with MEG and its correlation with the physiological MSNA response. To achieve this aim frequency filtering will be added around the beta oscillations, to examine whether filtering

of the MEG data to remove signals outside this frequency band can improve the classification accuracy. This will hopefully give us the ability to recognize what features the network is learning from and bring the research one step closer to being able to correctly classify the risk of developing essential hypertension. The investigation of what the network has learned also brings the method one step closer to explainable AI. To furthermore achieve the aim, this thesis will explore whether the multiclass classification approach can be improved by data management in the form of defining groups of individuals based on their physiological stress response. To that end groups will be formed that are labeled after increasing risk.

1.3 Demarcations

One demarcation in this master's thesis is that the preprocessed data will only be fed into the neural networks developed by the previous master's thesis. This means that other neural networks will not be examined. In order to accurately compare the results between the master's theses and to evaluate the results from the preprocessing steps, the architecture of the networks will not be changed or will be changed minimally. Furthermore, due to the limitation of time, all available conventional analysis will not be implemented for the preprocessing step. The primary method will be to filter the data to get a concentrated area of interest, i.e., the area around the beta oscillations. The next step in preprocessing was decided to be a transformation of the data with time-frequency analysis but due to time restriction, only a framework for this methodology will be presented in Appendix A.

1.4 Ethical and Societal Aspects

The use of machine learning in clinical applications can greatly reduce the burden on healthcare professionals. However, when human lives are involved, the interpretability of the machine learning predictions are of great importance. Clinicians have to trust and understand, to some degree, the predictions. Therefore, to use any kind of implementation of machine learning in healthcare brings to light the problem with black box approaches. That being, the decision making process is hidden in the models and not always intuitive [8], which could bring into question the reliability of the results in certain settings. Therefore, due to the increase of AI applications in many fields, the concept of explainable AI has recently emerged and gained more importance [9]. If people are going to trust these applications then it is essential that there are explanations for how the system works. To understand and explain the networks, different methods are used to either see what the network is learning or to open up the black box to examine the parameters in the network. This master's thesis is a first step in the direction of explainable AI as conventional neuroimaging analysis will be used to analyze what the network is learning from the data. The personal privacy must be considered as well since the data will be acquired from real subjects. Different regulations such as General Data Protection Regulation (GDPR) exist to ensure that the personal privacy of the used data is protected [10]. However, in the case of this master's thesis, the concern of violating

the personal privacy will not be an issue since the used MEG data is anonymized. Thus, it cannot, to our knowledge, be used to identify an individual, which means that the personal privacy will be preserved.

In today's world, data has an important role in everyday life and within healthcare. There are large amounts of data generated every day and it will only be more in the future. It is therefore important that methods managing the generated data are continuously developed. Especially for healthcare, as it is important that everyone in our growing population can get the medical help they need. Deep learning can be a powerful tool to handle all this data and help people, while also producing results that have great impact on the future. The result of this thesis could therefore increase the social sustainability by working in a preventative fashion, possibly making it easier for people to see if they are at risk of getting hypertension, and thus can take precautionary actions.

2

Theoretical Background

Hypertension, or high blood pressure, is a public health issue and it is prevalent in a large part of the adult population around the world. An estimated number of one billion people are affected by hypertension globally and according to projections, the number will continue to increase [11]. The danger of having prolonged high blood pressure is that it can cause damage to the arterial walls and can lead to aneurysms or even ruptures if left untreated. Thus, hypertension increases the risk for many cardiovascular diseases [12]. Hypertension can be divided into two types, primary hypertension and secondary hypertension. The condition is separated into these two types depending on the cause of the high pressure. If the cause is due to another disease, such as a kidney disease, then it is defined as secondary hypertension. Around 5 % of hypertension cases have this type [13]. The majority of people that suffer from hypertension are afflicted by primary or, as it is also called, essential hypertension which is not connected to any single known cause. It is however linked to genetic factors and environmental factors that are on the rise in the modern society, e.g., high salt intake, sedentary life style, and obesity [13], [12]. One prominent factor is stress or exposure to stressors in everyday life [11]. These stressors can be sharp lights or sounds but also emotional stressors such as worry, to which people can react in different ways.

2.1 The Sympathetic Nervous System and Stress

The sympathetic nervous system (SNS) is one of three parts of the autonomic nervous system (ANS), which is the system generally involved in regulation of bodily functions that are not consciously controlled, e.g., digestion, blood pressure, and heart rate [14]. The other two parts of the ANS are the parasympathetic nervous system (PNS) and the enteric nervous system (ENS). The ENS is somewhat independent and is mainly involved in managing gastrointestinal behavior. While the SNS and PNS together cooperate to maintain the balance between action and rest in the body. These systems are constantly modulating bodily functions in response to both external stimulus (e.g., a bright flash or a sharp sound) and internal stimulus (e.g., a stimulus originating from inside the body) [15], [14]. It is the triggering of SNS that enables responses to stressors. This can be an adjustment to heat as the body triggers sweat glands to produce sweat, or a reflex response to a sudden stimulus, increasing the heart rate and blood flow to skeletal muscles [14]. For this

reason, an activation of the SNS is usually known as the "fight-or-flight" response. The opposite is said for PNS, which is related to the "rest-and-digest" response that regulates the body in relaxed situations. These two systems are often described as opposites and it could be interpreted as a "push-and-pull" association, i.e., when the activation of SNS lessens the PNS increases. This is true for some organs in the body however many are just activated by one system. The constriction of most blood vessels, for example, are only activated by SNS [15]. As mentioned, the SNS is activated by stressors and this is necessary to regulate the body's functions. However, if the SNS is engaged too much it can be dangerous as overactivity of the SNS has been connected, among other things, to cardiac failure and hypertension [16].

To study sympathetic activations, it has become a standard to record nerve traffic in order to measure responses to stimulus [15]. The muscle sympathetic nerve activity (MSNA) is this type of response and can be measured with microneurography, which is an invasive method for recording nerve traffic. The MSNA activates the smooth muscles that act upon the blood vessels and triggers vasoconstriction. This constriction modulates the blood flow and subsequently the blood pressure [17]. Individuals' MSNA and blood-pressure response to stressful or arousing stimuli has, in a study of healthy male subjects, been shown to be strongly correlated. The study showed that there is also a significant variance between individuals with respect to both responses, where 50 % of the subjects inhibited the MSNA activation (inhibitors) resulting in lower blood pressure spikes while the other 50 % presented less MSNA inhibition (non-inhibitors) which resulted in a prolonged blood pressure spike after stimulus [6].

2.2 Neural Signals

The brain is the control center of the nervous system and hence the control center of the human body as it is the organ that processes incoming signals, coordinates, and sends out signals to the rest of the body. This complex organ, which consists of signal transferring neurons and neuron supporting glial cells, has various divisions with different functions and responsibilities for processing different kinds of information [18]. The stimuli are sent to the brain for processing and signals are subsequently sent from the brain to stimulate response actions, for example muscle movements.

The activity of neurons can synchronize, and this synchronization of the neural activity leads to rhythms, called neural oscillations. These neural oscillations varies depending on the brain state and function [19]. The neural oscillations are usually divided into different approximate frequency bands: delta (1 - 4 Hz), theta (4 - 8 Hz), alpha (8 - 12 Hz), beta (13 - 30 Hz) and gamma (> 30 Hz) [19], [20], [21]. There is a hypothesis that neural oscillations, within the same frequency band, are associated with different processing domains of the brain depending on the function [19]. For example, theta oscillations are associated with working memory functions, whereas, are beta oscillations are associated with, for instance, motor functions [21].

In a previous study, it was found that there is a correlation between the beta oscillations, more specifically the beta rebound, in the Rolandic area of the brain and the MSNA response [2]. In this case, the beta rebound refers to the signal increase in beta power that occurs at the same time of stressful or arousal stimuli, which are used for establishing MSNA profiles. The Rolandic area of the brain refers to the area that includes the central sulcus (Rolandic fissure), motor cortex, and somatosensory cortex [1], seen in Figure 2.1. Beta oscillations of the brain can be recorded with the non-invasive neuroimaging technique magnetoencephalography (MEG), and have been seen as a potential non-invasive biomarker for the MSNA response [2].

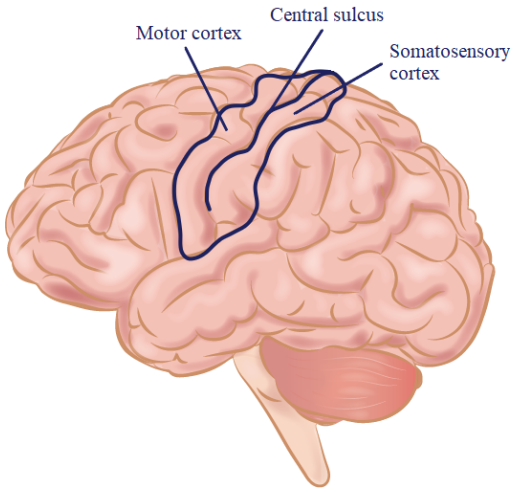


Figure 2.1: *Rolandic area of the brain. The Rolandic area, marked with dark blue, comprises the central sulcus, motor cortex, and somatosensory cortex. Source: Image of the brain is acquired from [22]. Modifications were made to highlight the Rolandic area of the brain.*

2.3 Magnetoencephalography

MEG is a non-invasive neuroimaging technique with both high temporal and spatial resolution used to record and study brain activity. During neural activities, signals are transmitted between neurons through action potentials, which are electrochemical impulses giving rise to induced magnetic fields that MEG measures and records with sensors outside the scalp [23]. Rather than the action potential, the main contributor to the MEG recordings are the postsynaptic currents, which are the ion flow induced within the neuron after the action potentials have passed a synapse [24]. The neuromagnetic signals are picked up by flux transformers and are coupled to a measuring device, which is the sensing element. The changing magnetic field from the brain, induce electric currents in the coil of the flux transformer enabling coupling and measurements in the measuring device. These flux transformers have different coil configurations to achieve different measurement properties. For example magnometers, consisting of one single coil, measures orthogonal magnetic fields and are more sensitive to external magnetic fields, whereas gradiometers, consisting of two coils, are more insensitive to magnetic fields that are distant [23].

The neuromagnetic signals from the brain, that are recorded by MEG, are on the scale of femtotesla (10^{-15} T), which are 7 - 8 orders of magnitude smaller compared to the magnetic field of the earth [23], [24]. Nowadays, the standard technology with sufficient sensitivity used to measure the weak magnetic fields of the brain is a superconducting quantum interference device (SQUID) [24], which is the measuring device coupled with the flux transformers. SQUID functions at very low temperatures, below 10 K [2]. In order to keep the low temperature, the SQUIDs are stored in a thermal insulated dewar filled with liquid helium. For MEG measurement, around 300 SQUID channels in liquid helium are arranged in a helmet to cover the entire head [24]. The set-up that was used to acquire the MEG data used in this master's thesis consisted of 306 SQUID channels, 102 being magnetometers and 204 being planar gradiometers [2]. Each sampling location in the helmet consisted of 1 magnetometer and 2 gradiometers, resulting in 102 locations with sensors in the helmet. Compared to the surrounding magnetic fields such as the magnetic fields due to cardiac activity or external electric equipment, the induced magnetic fields of the brain are much weaker, which means that surrounding magnetic fields can cause significant interference to the measurements. In addition to the coil configuration of the flux transformers, MEG measurements are usually carried out in a magnetically shielded room to further reduce external magnetic disturbances [2], [23], [24].

2.4 Frequency Filtering

Frequency filtering is usually implemented to isolate signals of interest or improve the quality of the data by reducing noise, but often for both reasons. It is an operation that removes or reduces features from a signal that correspond to a specified frequency or frequency band of the signal, which is determined by cut-off frequencies [25], [26]. Depending on how the filter is designed the filter can have different properties. The most common filter types are: high-pass, low-pass, band-pass, and band-stop filter. They are all defined by which frequencies they pass though and which they attenuate, e.g., low-pass filter attenuates high frequencies while allowing low frequencies to pass, band-pass filter only passes a specific band of frequencies [26]. The filter type is also described by the filter's impulse response, which can be either a finite impulse response (FIR) or a infinite impulse response (IIR). FIR is usually more stable and can have a linear phase, which means that the relationship in time can be preserved since this filter does not produce a phase distortion. The FIR filter is more computationally heavy though. IIR on the other hand is more memory efficient and utilized when linear phase is not as important, this filter type is however less stable [25]. A filter is usually complimented by a windowing function as this limits the passband and stopband ripple, which are amplitude fluctuations of the signal in the corresponding band [25], [26].

2.5 Neural Networks

Artificial neural networks (ANNs) are inspired by how the biological neural networks work. Depending on the connections or weights between the, so called, neurons in

the network, ANNs can have different architectures. Each architecture has different properties and are therefore suited for different tasks of learning. A feedforward network is one type of ANN architecture. In this type of network, the information is only moving in one direction, which is forward from the input layer, through the hidden layer(s), to the output layer [27], [28]. These kinds of ANNs are nonlinear function approximators and they have the ability of mapping input and output nonlinearly. One example of a feedforward ANN is a multilayer perceptron (MLP), which is a high dimensional function approximator [29], [28]. As the name indicates, a MLP network contains one or several hidden layers, and the information is processed in a forward manner. A general architecture of a MLP network can be seen in Figure 2.2. MLP networks are used for nonlinear function approximations and nonlinear classification tasks. In image processing, MLP has also been applied for problems including optical character recognition and medical diagnosis [29].

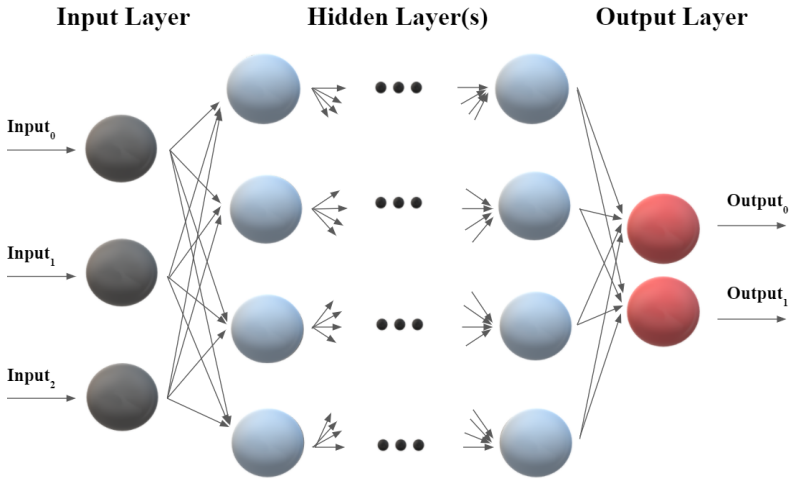


Figure 2.2: *General architecture of a MLP network.*

Another type of ANN is a feedback network, which is a feedforward network but with the additional capability of feeding the output information of the layers back into the model, creating a loop that enables a memory function in the network [27], [28]. One type of such a network is the long short-term memory (LSTM) network, which is a gated recurrent neural network. It consist of gated internal recurrence (self-loops), which is recurrence to the "LSTM cell", and outer recurrence to the LSTM network. They are suited for processing sequential data, that is data appearing in a sequential order, such as speech or time series [28]. A general architecture of a LSTM network can be seen in Figure 2.3, where I is the input and O is the output for the LSTM cell. The subscript t represents the time step where the input and output are vectors for the corresponding time step. The cell incorporates the memory function of the layer. LSTM networks are used in many applications including, speech and unconstrained handwriting recognition, machine translation, parsing, and image captioning. It has been shown that LSTM networks can learn long-term dependencies more easily compared to simpler recurrent neural networks [28].

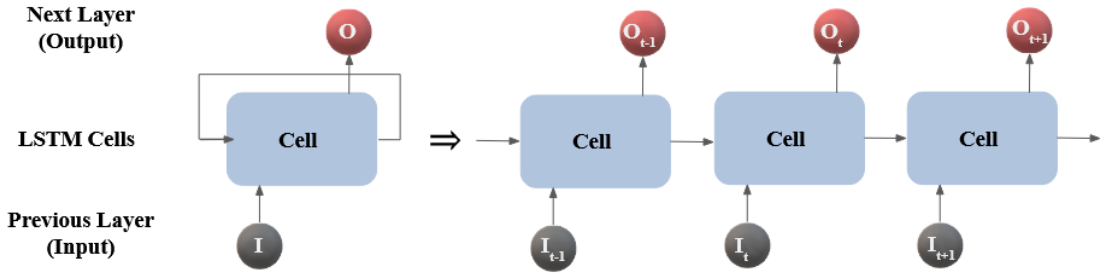


Figure 2.3: *General architecture of a LSTM network.*

Having a neural network with more than one hidden layer results in a deep neural network and deep learning. With this, the information is processed through the hidden layers which enables self-learning of the network [28]. In deep learning the goal is often to classify data, such as an image or a signal, by predicting the correct label or name to a given input. Two types of these classification problems are binary classification and multiclass classification. In binary classification, the predicted label is either true or false, i.e., there are only two cases that the input can be classified as. In multiclass classification the predicted label can be classified as one of multiple classes [30]. Classification accuracy is often the metric used to measure the performance of the neural networks' predictions. The accuracy is simply the amount of correctly predicted labels divided by the total number of predictions made [28], [30].

2.5.1 Interpretability and Explainable AI

ANNs, and deep learning methods in general, are attractive for problem solving and data analysis since they have a unique way of learning relationships and patterns in the data that the networks are presented with [28]. They are not bound to any pre-existing rules of what the data should look like, but the networks can still find complex connections in data. However, this way of learning and the increased complexity of neural networks that are broadly used today can make the results somewhat difficult to interpret or explain due to the inherent black box nature of the networks. The need for more explainable models have hence emerged and the term explainable AI is often used to encompass many efforts to make networks more transparent or explainable [9]. One way to do this is to "open" pre-trained black box models and evaluate how they learn, by estimating the parameters that the network is learning or by computing feature importance for the network [9], [31]. Another is to remove or mask off parts of the data set, or otherwise change the data set, to force the network to learn from another feature and analyze how this have affected the performance, which also gives information of what the network is learning [31]. However, when creating more interpretable models there is usually a trade-off between interpretability and accuracy [9].

3

Methods

To achieve the objectives and ultimately the aim of this master's thesis, different methods were executed. Firstly, the acquired MEG data was further processed by filtering the data and then fed to the neural networks to examine how the classification accuracy of the developed neural networks would be affected. Furthermore, multiclass classification with different group divisions was also explored to see if improvements could be made to this kind of classification. To analyze the effect of the filtering, the filtered MEG data was used for both binary classification and multiclass classification. All code for filtering and classification was implemented in Python 3.7. The package MNE-python [32] was used for the preprocessing, analysis and visualization of the MEG data.

3.1 Data Background

MEG data was acquired from MedTech West and had followed MEG processing methods to reduce noise and remove artifacts such as head position indicator (HPI) coils, signal space separation (SSS) and independent component analysis (ICA). Additional artifact removal methods such as active shielding and Maxfilter software were also utilized [2]. As a part of this process, the data was filtered between 0.5 Hz and 40 Hz. The data was recorded from 19 subjects, 10 non-inhibitors and 9 inhibitors. As mentioned in Section 1.1, the provided MEG data had already been split into time segments around the stimulus, these time segment are known as epochs. The epochs spanned from -2000 ms to 2000 ms with the stimulus trigger at t = 0. The stimulus was an electrical pulse administrated to the index finger and was delivered three consecutive times, forming a pulse train. The pulses of the train are denoted as pulse 1 (p1), pulse 2 (p2), and pulse 3 (p3). Each subject received 72 pulse trains with 30 - 60 s intervals between each train.

Furthermore, the data was also spatially divided based on the correlation between the beta rebound in the Rolandic area and the MSNA response [2]. Thus, the sensors, and by this the sensor data, of the MEG helmet were divided into three groups called All, Decent and Peak. This was to make it possible to focus on the sensors that are likely to pick up the activations that have been shown to be correlated with MSNA. Peak is the group of sensors closest to the Rolandic area of the brain and therefore most likely to pick up the activations. The Peak set contains

the fewest number of sensors. Decent is the groups of sensors that covers a larger area around the Rolandic area and incorporates the Peak sensors. All denotes the group where all 306 sensors of the MEG helmet are included.

3.2 Filtering over Frequency Domain

To get a more concentrated window over the signal of interest, i.e., the beta oscillations, the epoched data was filtered over the frequency domain with a FIR band-pass filter with filter boundaries at 12 - 30 Hz. Beta oscillations are not strictly between 13 - 30 Hz in every source, hence a lower boundary at 12 Hz was chosen. The band-pass filter was generated and applied with the function `mne.filter` from the MNE package. The default filter parameters were considered appropriate for the filtration, thus the filter used was a non-causal, one-pass, zero-phase band-pass filter and the window function was a Hamming window with 0.0194 passband ripple and 53 dB stopband attenuation. The filter length was 1101 samples. To confirm that the filtering of the epoched data had been successful, the MNE function `mne.plot_psd` was used to plot the power spectral density of the signal to visualize and inspect the frequency content of the signal after filtration. The data filtered in this way will henceforth be called filtered data, whereas the data only having the previous pre-processing, including filtration between 0.5 Hz and 40 Hz, will be called unfiltered data.

3.3 Data Structure

The data structure is constructed in the same manner as Bakidou and Nord Odhner and will thus be explained briefly. After filtration, the MEG data was cropped to only contain data from -100 ms to 1500 ms. To examine different aspects of the neural networks, categories were constructed with different data included. The previous master's thesis created five categories, A, B, $\tilde{A}$, $\tilde{B}$, and X. However, in this thesis multiclass classification was examined further. Therefore, we added two more multiclass categories, *X3* and *X4*. Moreover, category X was renamed *X19* in this thesis to reflect the 19 classes. Category A and B were used for binary classification (i.e., inhibitor vs. non-inhibitor) and both categories had the same test set consisting of data from one inhibitor, being subject 0283 with a MSNA inhibition value of 58.64 %, and data from one non-inhibitor, being subject 0179 with a MSNA inhibition value of 7.9 %. These subjects were randomly selected to be in the test set. The difference between category A and B is the inclusion of a borderline case in category B. As described in the previous master's thesis, subject 0293, with MSNA inhibition value of 30.01 %, is considered a borderline case because the threshold between inhibitor and non-inhibitor is at 30.0 % [1]. From the results of the previous master's thesis the borderline case did not seem to impact the training to any significant extent but it was decided that we would keep all five categories to investigate if the borderline case had any effect on the training after filtering. Furthermore, category $\tilde{A}$ and $\tilde{B}$ were also used for binary classification but had the extreme cases as test set instead, which is data from the

subjects having the lowest and highest MSNA inhibition values. Therefore, the test set for category $\tilde{A}$ and $\tilde{B}$ consisted of subject 0301 (non-inhibitor) with -132.47 % MSNA inhibition value and subject 0158 (inhibitor) with 83.74 % MSNA inhibition value. Like category A and B, the borderline case was included in $\tilde{B}$ and not in $\tilde{A}$. The last three categories created were category *X19*, *X3* and *X4*. These categories were created for the purpose of multiclass classification. Category *X19* was constructed for classification on individual level, i.e., with 19 classes, whereas category *X3* consisted of 3 classes and category *X4* of 4 classes. Each set for the multiclass categories, training, validation and test, contained a percentage of data from all subjects including the borderline case. The test set for these categories consisted of 10 % of available epochs from all subjects. A summary of the content and purpose of the categories are presented in Table 3.1.

Table 3.1: *Summary of all categories. The table presents the test set for each category and what classification the categories will be trained for as well as if the borderline case is included. Source: Adapted with permission from [1].*

	Category	Borderline case inclusion	Test case
Binary classification	A	No	Random (IDs 0179, 0283)
	B	Yes	Random (IDs 0179, 0283)
	$\tilde{A}$	No	Extreme (IDs 0301, 0158)
	$\tilde{B}$	Yes	Extreme (IDs 0301, 0158)
Multiclass classification	*X19*	Yes	10 % of available epochs
	X3	Yes	10 % of available epochs
	X4	Yes	10 % of available epochs

For each category, three data packages were created for the different sensor groups, one data package for All sensors, one data package for Decent sensors, and one data package for Peak sensors. Furthermore, each package of the MEG data was sorted into the 3 different pulses, p1, p2, and p3, representing the arousal stimuli in the pulse train given at the experiment. This results in 3 divisions of each package with the three pulses, each containing their own training, validation and test sets.

3.4 Multiclass Classification

Multiclass classification can give different insights into the data compared to a binary classification. In fact, the developed neural networks have shown promise in

the previous master's thesis for classification at individual level, however the result implicated that it was not the MSNA response that the networks were learning. Considering this, together with the idea that the MSNA inhibition value is believed to be approximately normally distributed, it invited the idea to examine if the learning and accuracies can be improved with different divisions of data than category *X19*, while also investigating the effects of filtering. The divisions, that will be described in detail below, were distributed as evenly as possible but due to the number of subjects being prime, there did not exist a division that would have an equal number of subjects in each class. This created a small bias for the classes that contained extra subjects but was deemed to not affect the outcome greatly. The multiclass categories were trained with both the filtered and unfiltered data. By having one unfiltered set and one filtered set of data, the approach with multiclass classification could be evaluated as well as the effect of filtration in the frequency domain. To aid the interpretation of the results within and between the different group divisions, chance level for each case, indicating the level in which the network is classifying at random, was also considered. The chance level was calculated with this following equation:

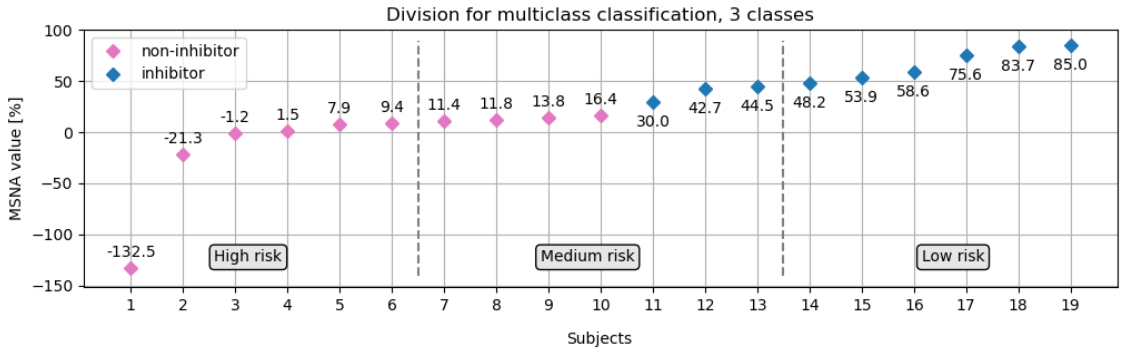
$$chance\ level = \frac{1}{number\ of\ classes} \tag{3.1}$$

The chance level was thus, 50 % for two classes, 33 % for three classes, 25 % for four classes and 5.3 % for 19 classes.

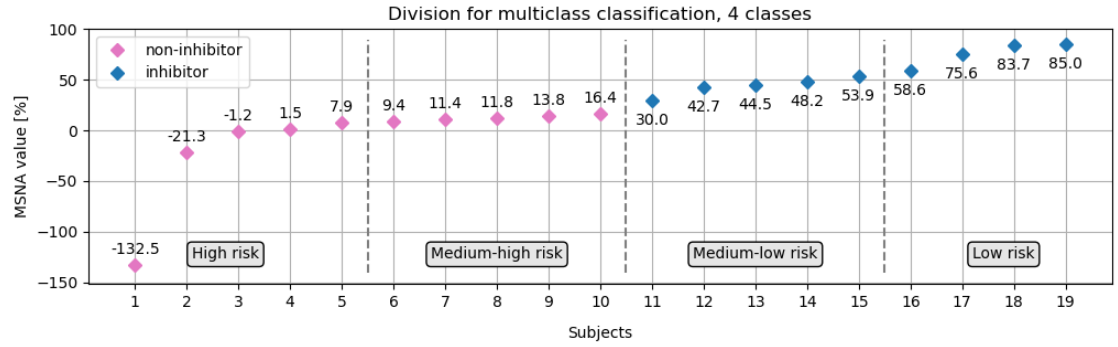
3.4.1 Group Division

Two main elements were considered when grouping of the subjects was performed. Firstly, the classes were not allowed to be too small, i.e., contain too few subjects. If the classes where too small the network might just learn the individual activations instead of the general signal that could be used to predict the inhibition type. The division possibilities were limited due to the number of subjects in the study. Secondly, the MSNA value is believed to be approximately normally distributed across the population which means that it is believed that most of the population have a MSNA value in the range around 30 %, which is the threshold of being an inhibitor and non-inhibitor. The subjects were arranged in an ascending order according to their MSNA inhibition value. After arranging the subjects, they were divided into different classes depending on their MSNA value, meaning that neighboring subjects belonged to the same class. Divisions of the 19 subjects into groups with 3 and 4 classes were explored and a summary of the groupings are presented in Figure 3.1. Groupings with 5 or 6 classes were also considered but deemed unnecessary once early result from the first divisions were evaluated.

The first grouping, category *X3*, was done over all subjects, not considering the inhibitor and non-inhibitor threshold at 30 %. With the aforementioned criteria in mind, a division with 3 classes, being high, medium and low risk, was considered suitable in this case. This grouping is visualized in Figure 3.1(a). The high risk class



(a)



(b)

Figure 3.1: *Group division of the 19 subjects for multiclass classification, arranged in ascending MSNA inhibition value. Subjects being non-inhibitors are represented by the pink items, whereas subjects being inhibitors are represented by blue items. The dashed lines distinguish the classes. (a) Shows the group division where the subjects are divided into 3 classes, high risk, medium risk, and low risk, containing 6, 7, and 6 subjects, respectively. The inhibition threshold at 30 % was not considered in this grouping. (b) Shows the group division where the subjects were divided into inhibitors and non-inhibitors and into 4 classes, high risk, medium-high risk, medium-low risk, and low risk, containing 5, 5, 5, and 4 subjects respectively.*

and the low risk class consisted of 6 subjects each with subjects having a MSNA value of 48.2 % and higher, and a MSNA value of 9.4 % and lower respectively, whereas the medium risk class consisted of 7 subjects with MSNA values between 11.4 - 44.5 %. The medium risk class consisted of 7 subjects to reflect the idea that the MSNA value is normally distributed. With this division, the medium risk class would have both non-inhibitors and inhibitors which might seem counter-intuitive when one of the goals is to classify these two groups. The reason for why this grouping approach was considered acceptable was because the classification results could still be usable in clinical applications. Because it is more important to be able to classify the subjects that have high risk of developing hypertension (strong non-inhibitors) than to strictly classifying individuals as inhibitors and non-inhibitors. This means that this division with 3 classes would focus more on the two side groups, high risk and low risk, and therefore classify people who are extreme cases, which could be the results that matter in the clinical setting as mentioned above.

The second grouping, category *X4*, contained 4 classes, high, medium-high, medium-low and low risk. This grouping is displayed in Figure 3.1(b). In this case the inhibitor and non-inhibitor threshold at 30 % was considered, meaning that no class in this division would contain both inhibitors and non-inhibitors. The idea with this grouping was to investigate the effect of dividing the subjects into more groups, having values closer to each other, and to have inhibitors and non-inhibitors in separate groups, which explains the consideration of the inhibitor and non-inhibitor threshold at 30 %. Furthermore, after having two classes (binary classification) and three classes (Category *X3*), four classes were naturally the next class division to examine without potentially having too few subjects in each class. To minimize bias in the data, the classes were divided as evenly as possible, with the number of subjects in each group nearly the same. To summarize, the subjects were divided into 4 classes; two classes, high and medium-high risk, being non-inhibitors and two classes, medium-low and low risk, being inhibitors. One inhibitor class, high risk, contained 4 subjects with MSNA inhibition value at 58.6 % and above while the other three classes contained 5 subjects each with a MSNA value of 53.9 % and below.

4

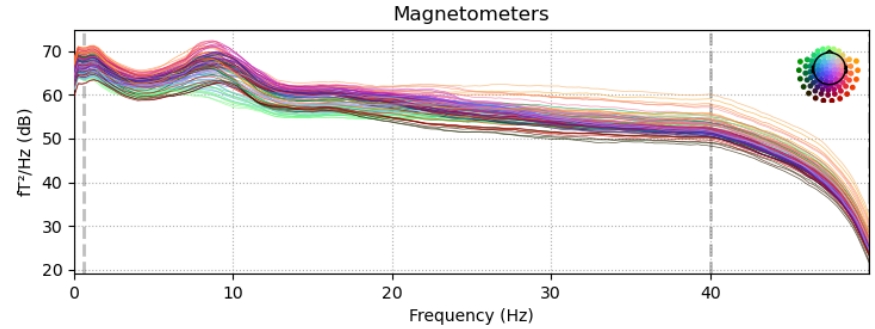
Results

Filtering the MEG data in the frequency domain had a major role in this thesis since subsequent analysis was dependent on the MEG data being filtered. For that reason, it was important to ensure that the MEG data had been filtered properly by visualizing and studying the data after filtration. Binary and multiclass classification were then performed on the filtered data and the unfiltered data with the developed LSTM and MLP networks. The results from the LSTM and MLP networks are presented in figures in this section. The accuracies for training, validation, and test, are presented for each pulse, p1, p2, and p3, in the pulse train, for All, Decent, and Peak sensors. Furthermore, the results for the filtered data were plotted together with the results from the original[1] unfiltered data in a bar chart to make the comparison more clear. For all bar charts in this section, the results from the unfiltered data have darker bars whereas the results from the filtered data have lighter bars with black frames and have displayed accuracies on the bars. When viewing the bar charts, if the lighter bar with the black frame is visible, it means that the filtered data got higher values. When interpreting the results of the networks, greater emphasis was put on the test accuracy as this accuracy is more important compared to the training and validation accuracy. Furthermore, in binary classification, a comparison of the resulting accuracies was made between the filtered data and the unfiltered data. In multiclass classification, an additional comparison of the results from the unfiltered data was made between the different group divisions.

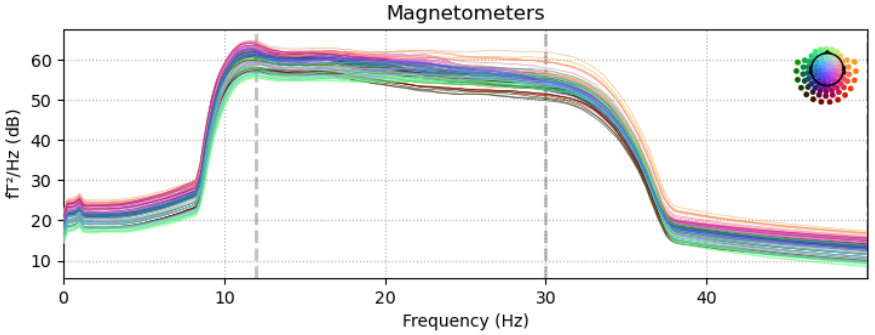
4.1 Filtering

The MEG data was visualized in the frequency domain to confirm the filtration. To inspect the frequency content, the power spectral density of the MEG sensors was analyzed. The power spectral density from the magnetometers in the sensors for subject 0284 are presented in Figure 4.1. Comparing the power spectral density of the unfiltered data, shown in Figure 4.1(a), and the filtered data, shown in Figure 4.1(b), it can be seen that the MEG signal has been successfully filtered as the signal now mainly contains information in the frequency range of 12 - 30 Hz which represent the beta oscillations of interest.

[1] Original here meaning the recreated results from the previous master's thesis [1].



(a)

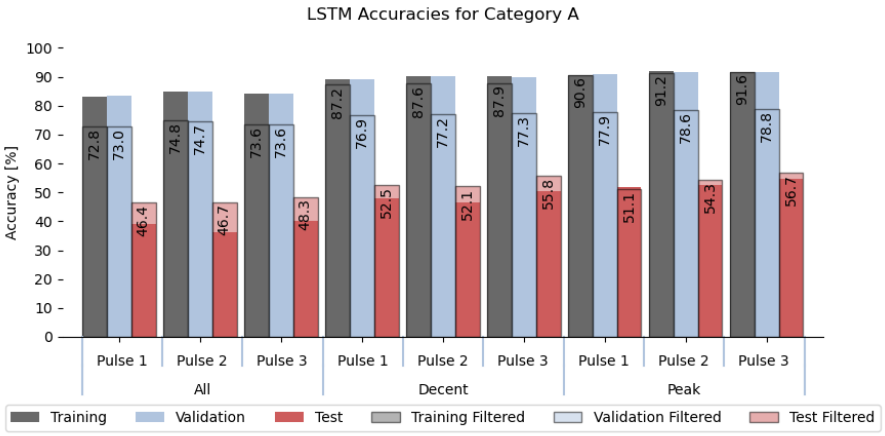


(b)

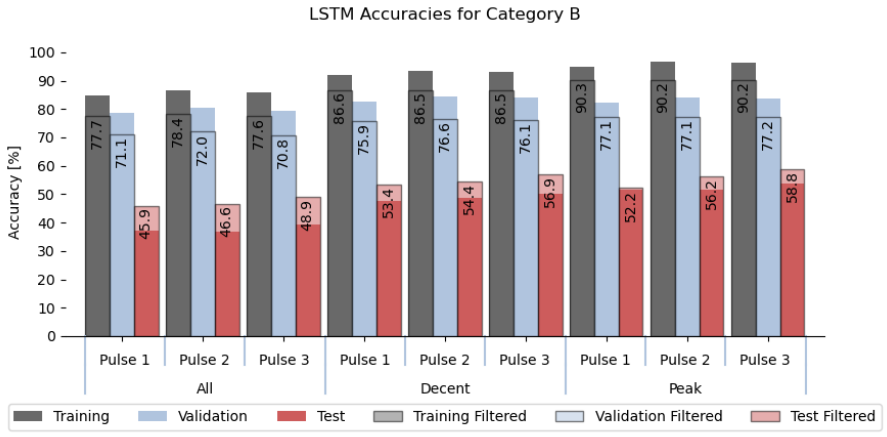
Figure 4.1: *Power spectral density of the recorded MEG data for the magnetometers for subject 0284. The small image on the right corner displays the locations of the MEG sensors in relation to the head. The dashed lines indicate the boundaries of the applied filters. (a) Frequency content for the unfiltered data. (b) Frequency content for the filtered data.*

4.2 Binary Classification

Starting with the results for the LSTM network, the combined bar charts of the results for the LSTM network with category A and B are presented in Figure 4.2. The filtered result got better classification results, i.e., higher test accuracy, for both category A and B. An important note is that the validation accuracy went down while the test accuracy went up for the filtered data. This means that the filtration has definitively made an impact and has made a change in what the network is learning form these categories. Even though the test accuracies have clearly improved with the filtered data, the results are still and not applicable in clinical settings. There is no set criterion for approval in the clinical setting but the test accuracy should be above 90 % [33]. The difference between the categories is the inclusion of the borderline case in category B. In accordance with the previous master's thesis, the inclusion of the borderline case did not seem to affect the accuracies differently in category A and B for the filtered data as the results from both categories had similar trends of increase and decrease of accuracy rate.



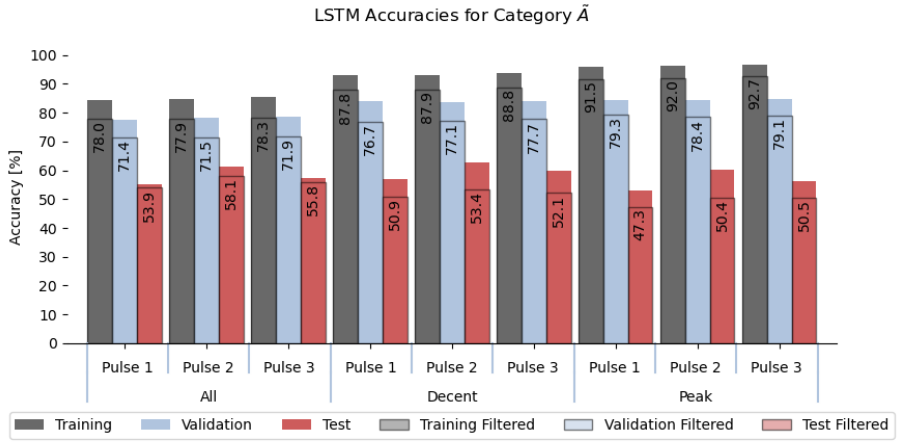
(a)



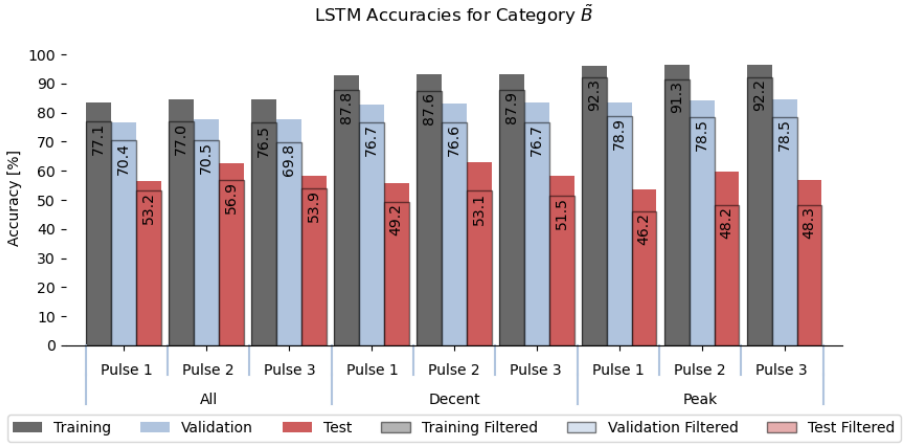
(b)

Figure 4.2: *Training, validation and test accuracies for the LSTM network, both unfiltered (dark bars) and filtered (light bars) accuracies. The labeled percentages always represent the filtered data. Category A and B have the same test set (random) but borderline case is excluded for A. (a) Shows the results for category A and (b) shows the results for category B.*

The results for the filtered data for category $\tilde{A}$ and $\tilde{B}$ for the LSTM network showed that all accuracies, training, validation, and test, were lowered compared to the unfiltered data. The resulting accuracies for category $\tilde{A}$ and $\tilde{B}$ for the LSTM network are presented in Figure 4.3. Remember that these categories have the extreme cases as the test set, i.e., subjects that have the largest negative and positive MSNA inhibition values. One interpretation of the results is therefore that it is more challenging for the network to predict these extreme cases after filtration, having values deviating from the other MSNA inhibition values, when the network was trained on filtered values around the inhibitor threshold at 30 %. This mostly applies to subject 0301 with MSNA inhibition value -132.47 % since this subject is an outlier with a much lower value compared to their closest neighbor, subject 0295 with -21.3 % MSNA inhibition value.



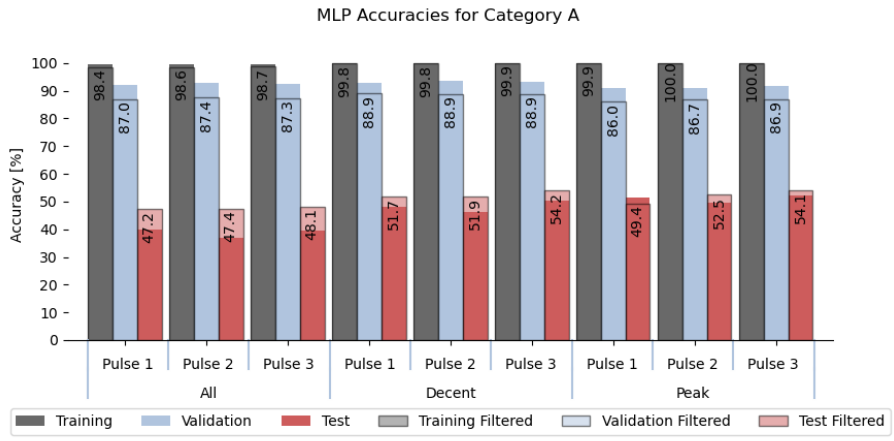
(a)



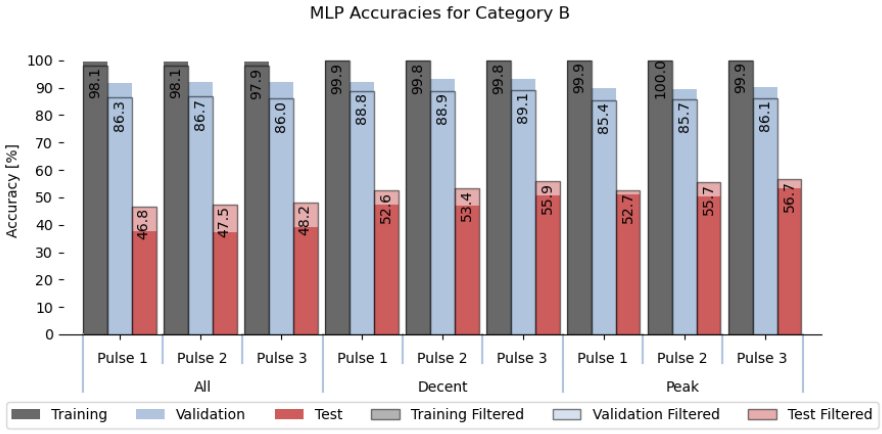
(b)

Figure 4.3: *Training, validation and test accuracies for the LSTM network, both unfiltered (dark bars) and filtered (light bars) accuracies. The labeled percentages always represent the filtered data. Category $\tilde{A}$ and $\tilde{B}$ have the same test set (extreme) but borderline case is excluded for $\tilde{A}$. (a) Shows the results for category $\tilde{A}$ and (b) shows the results for category $\tilde{B}$.*

In the case of the MLP network, the results for category A and B showed a similar trend as the LSTM network, a decrease in validation accuracies and an increase in test accuracies. These results are shown in Figure 4.4. Like the results from the LSTM network, the filtering of the data made the gap between validation and test accuracies smaller. However, the gap is still larger for the MLP network compared to the LSTM network. The smaller gap between validation and test accuracies after filtration still indicates an improvement of classification because this could mean that the network has learned a more general characteristic of the data.



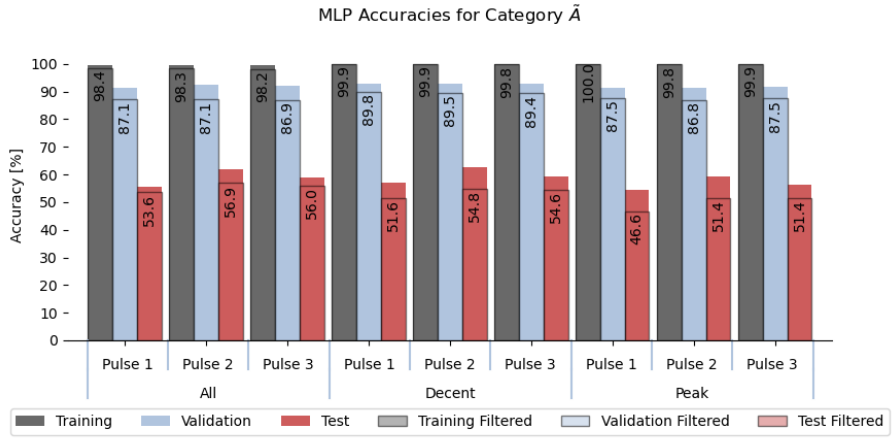
(a)



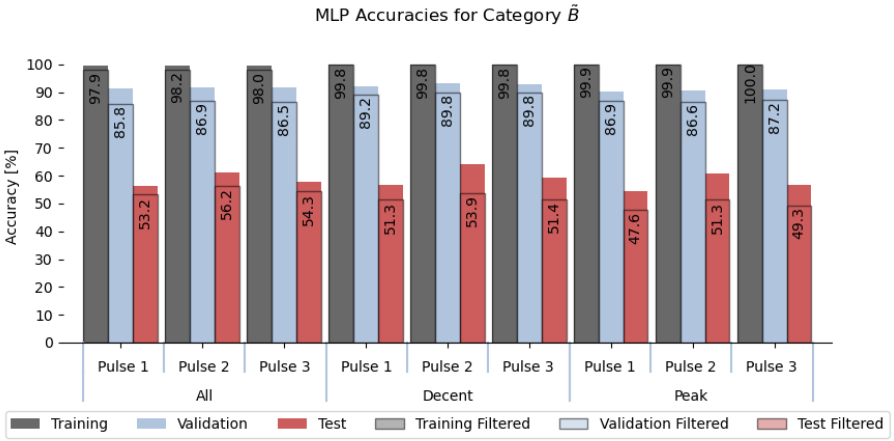
(b)

Figure 4.4: *Training, validation and test accuracies for the MLP network, both unfiltered (dark bars) and filtered (light bars) accuracies. The labeled percentages always represent the filtered data. Category A and B have the same test set (random) but borderline case is excluded for A. (a) Shows the results for category A and (b) shows the results for category B.*

Category $\tilde{A}$ and $\tilde{B}$ for the MLP network showed no real difference in results compared to the same categories in the LSTM network, where all accuracies were lower after filtration of the data. Figure 4.5 shows the results for $\tilde{A}$ and $\tilde{B}$ for the MLP network. The same interpretation as in the case of the LSTM network can be used for this case as well, that after filtration it may be more difficult for the network to classify a subject with an extreme value of MSNA when the network was trained on individuals that had MSNA values closer to each other.



(a)



(b)

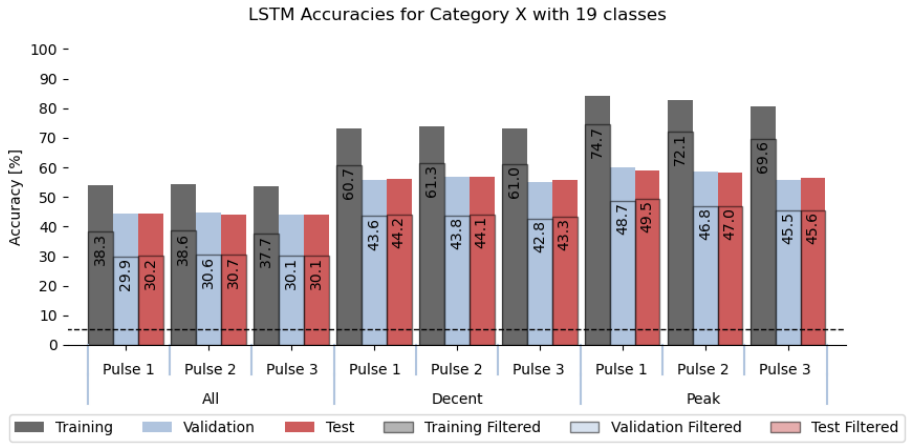
Figure 4.5: *Training, validation and test accuracies for the MLP network, both unfiltered (dark bars) and filtered (light bars) accuracies. The labeled percentages always represent the filtered data. Category $\tilde{A}$ and $\tilde{B}$ have the same test set (extreme) but borderline case is excluded for $\tilde{A}$. (a) Shows the results for category $\tilde{A}$ and (b) shows the results for category $\tilde{B}$.*

4.3 Multiclass Classification

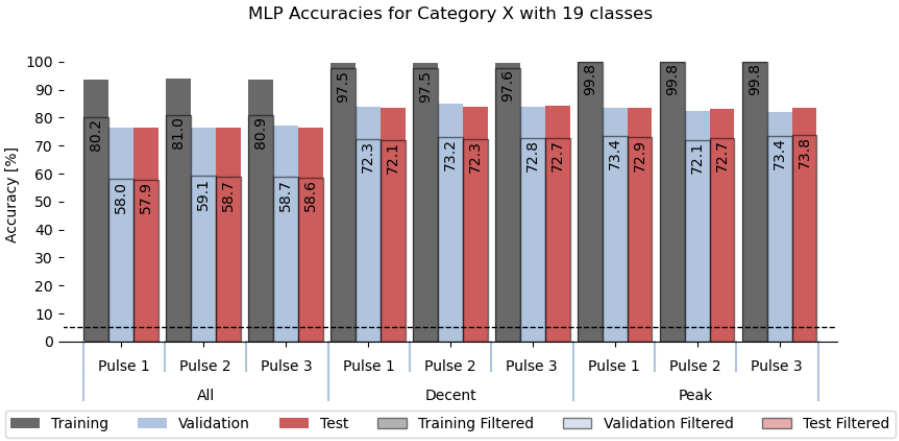
The results for the multiclass classification are presented in the same way as the binary classification with the addition of a dashed line indicating the threshold for the calculated chance level. If the accuracy lands below this level that means the network has worse performance than random classification for that number of classes. The chance level is important to consider in the case of multiclass classification since the results from the unfiltered data are compared between categories. The chance level was not highlighted for the binary classification since the focus was on comparing the unfiltered and filtered results. Whereas in the multiclass classification, it is of more interest as a reference point when comparing the different group divisions in order to evaluate the performance of the networks for the different groupings

of the data.

The first step with multiclass classification was to apply the filtered data to category *X* that had 19 classes, i.e., one class for each subject. The results from *X19* showed that all accuracies of the filtered data decreased compared to the unfiltered data, both for MLP and LSTM. The largest decrease was for MLP and it decreased with as much as 18.5 percentage points for All sensors. Accuracies for category *X19* for both networks are presented in Figure 4.6. The accuracies were low for the LSTM network, but they were still significantly higher than the chance level at 5.3 %, which indicate that the network is not randomly classifying the filtered data.



(a)



(b)

Figure 4.6: *Classification results for category X with 19 classes (individual level classification). The chance level at 5.3 % is marked by the dashed line and the labeled percentages always represent the filtered data. (a) LSTM accuracies for the unfiltered (dark bars) and filtered (light bars) data sets. (b) MLP accuracies for the unfiltered and filtered data sets.*

Next, was category *X3*. The results for the filtered data and the unfiltered data for both networks are displayed in Figure 4.7. The networks showed different classification results for this division of category *X*. For LSTM, on the filtered data, the accuracies were noticeably lower for all pulses when looking at All sensors. How-

ever, when looking at the Decent and Peak sensors, the accuracies were only lowered by a small amount for the Decent sensors and the accuracies for the Peak sensors increased instead. The increase was not significant, however it is a positive sign that it is the Peak sensors that has shown an increase, since those sensors should have picked up the strongest activations related to MSNA. This indicates that the division into three classes coupled with the filtration could enable the network to better classify the more general data. In the case of the MLP network, all accuracies decreased when the data was filtered but not as much compared to *X19*.

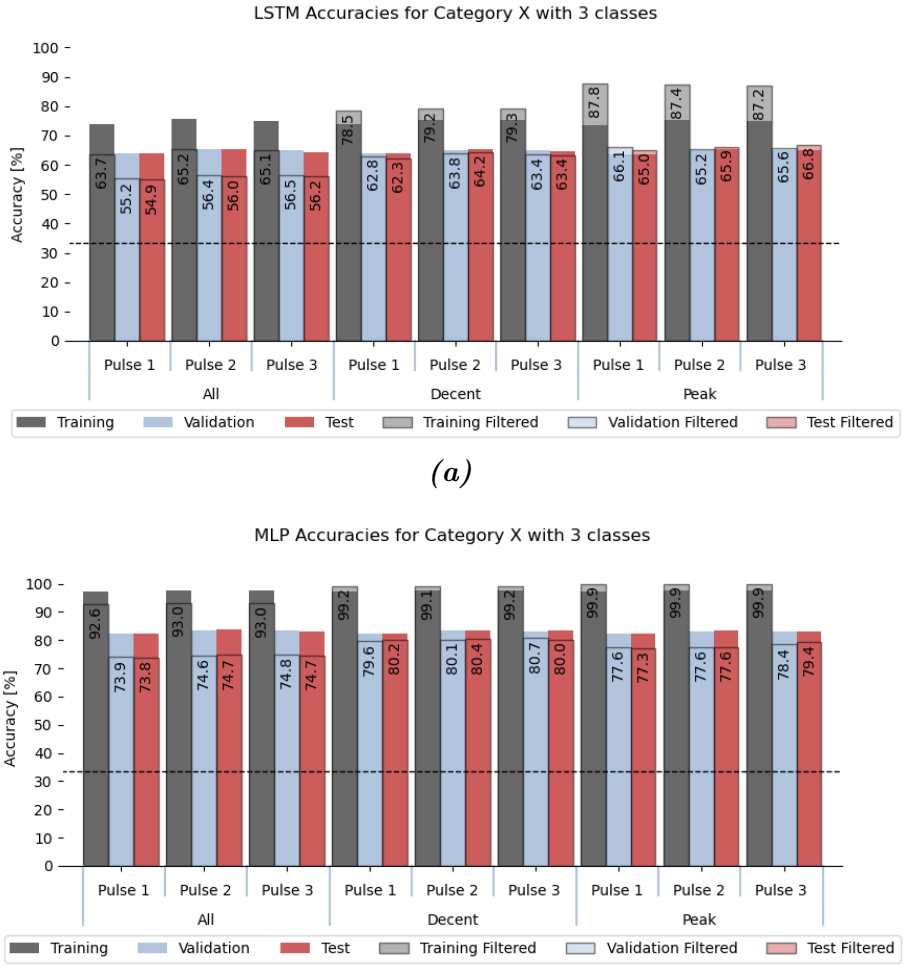
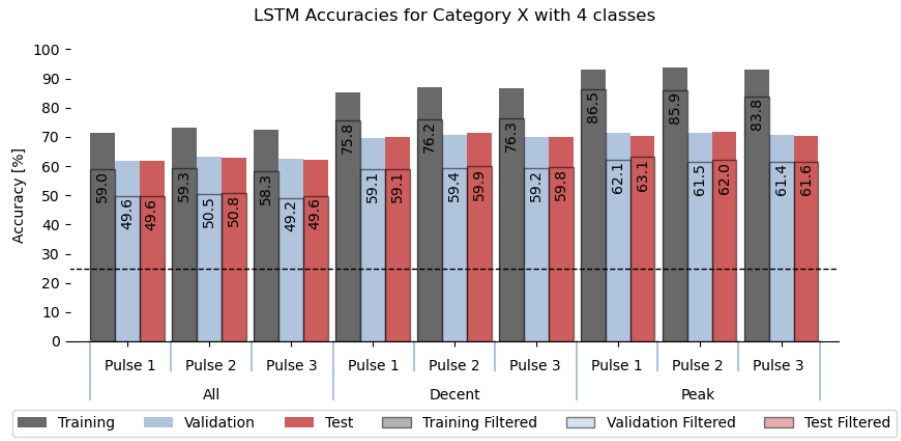


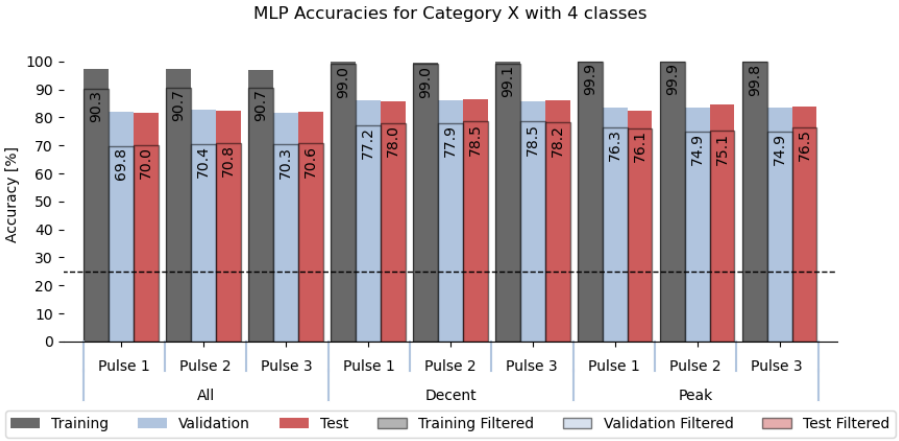
Figure 4.7: *Classification results for category X with 3 classes (3-class classification with inhibition-threshold disregarded). The chance level at 33.3 % is marked by the dashed line and the labeled percentages always represent the filtered data. (a) LSTM accuracies for the unfiltered (dark bars) and filtered (light bars) data sets. (b) MLP accuracies for the unfiltered and filtered data sets.*

Category *X4* was the last multiclass classification that was tested for this project, since splitting the data into more classes would most likely result in too few subjects in each class. Like the results from category *X19*, all the classification results for category *X4* had a decrease in validation and test accuracies when the data was filtered, see Figure 4.8. Also, note that the filtration accuracies for *X4* are at similar

levels as for *X3* but the chance level is slightly lower for this category, because *X4* contains more classes. This could influence what the accuracies are representing.



(a)



(b)

Figure 4.8: *Classification results for category X with 4 classes (4-class classification with inhibition-threshold taken into account). The chance level at 25 % is marked by the dashed line and the labeled percentages always represent the filtered data. (a) LSTM accuracies for the unfiltered (dark bars) and filtered (light bars) data sets. (b) MLP accuracies for the unfiltered and filtered data sets.*

At this point, when all *X* categories had been tested, the results from the unfiltered data can be compared between the categories to evaluate the effect on the training and classification of the network when dividing the unfiltered data into different classes. When the X categories are plotted, an increase between the categories can be seen in validation and test accuracies for the LSTM network, see Figures 4.6(a), 4.7(a) and 4.8(a). The increase of accuracies in the different divisions could imply that the multiclass classification has improved results. However, with fewer classes, the chance level increases and therefore the network has a higher probability of randomly getting a correct prediction. Indeed, when taking the chance level into consideration, it could be argued that category *X3* is the weakest division with the lowest accuracy since the chance level for category *X3* is the highest at 33 % and

category *X3* and *X4* acquired similar results. The MLP network remains around 80 % for every sensor group and group divisions with only slight deviations, see Figures 4.6(b), 4.7(b) and 4.8(b). These accuracies for MLP might make it seem that the network is unchanged when divided into different categories but keep in mind that the chance level changes here as well.

5

Discussion

In this thesis, the already preprocessed and epoched MEG data was further filtered to contain data in the range of 12 - 30 Hz. By doing so, ringing artifacts, which are ripples of artifacts at sharp transitions or discontinuity in a signal, could have been introduced to the MEG data at the edges, due to the sharp properties of the band-pass filter. These ringing artifacts can have had an impact on the training of the networks. Because it could be that the network sees the artifacts as a part of the signal. Therefore, it cannot be excluded that the networks could have been influenced by the artifacts, but likely not to the extent that it has overshadowed the other properties in the signal. Due to the sharp feature of the band-pass filter, it is not possible to get rid of ringing artifacts completely. However, it is possible to reduce the amount of ringing artifacts obtained by filtering the raw MEG data before extracting the epochs but handling the raw MEG data was outside the scope of this thesis.

Filtering over the epochs results in the MEG data having information over fewer frequencies, mainly in the range of the beta oscillations. Subsequently, this has had an impact on the training of the networks. The combination of the fact that it is not known which exact signals of the MEG data are associated with MSNA, and that it is not completely known what information is being removed or suppressed by filtration, has therefore led to various interpretations of the obtained results for both binary and multiclass classification. Indeed, when filtering, information is being removed or suppressed. This removed information could be the information the network learned from prior to the filtration such as the individual brain activation oscillations or measurement specific noise. There is also the possibility that there exists MSNA-related characteristics outside the beta band in the lower or higher frequencies. Furthermore, it could also be that what is left of the filtered signal might not have been enough for the network to learn a general pattern for classification, which could also implicate a limitation for actually learning the brain signals that are coupled to MSNA inhibition.

5.1 Binary Classification

For all results, including results from the multiclass classification, it was clear that the filtration generally produced accuracies that were lower than the unfiltered coun-

terpart results. We have mentioned reasons as to why this could be the case and one explanation for the results is that some aspect of the signal that was suppressed or removed had been contributing to the learning before filtration. Since most accuracies, except test accuracies for category A and B, show a drop it is likely that information has been lost. This, however, does not implicate that this information was the correct one to learn, only that it made it easier for the network to make a distinction between classes. What that information was is hard to assess, it could possibly be measurement specific noise or some individuality in the brain oscillations. We also speculated, based on the results, that there could be some characteristic of the MSNA signal that is also present in the lower or higher frequencies. Through filtration we wanted to focus on the beta oscillations since these had been correlated with MSNA. Therefore, the results could indicate that MSNA is not just active in the beta band and that this is the information that was removed after filtration.

The idea with the filtration was that it would remove a notable part of the individuality of the brain activation recordings and leave behind a more general MSNA characteristics, restricting the network to learn from this instead. Recall then, that the networks trained on data sets with the random test set, category A and B, showed a resulting smaller gap between validation and test accuracies, indicating an improved classification. However, the opposite result was observed for the test accuracies for category $\tilde{A}$ and $\tilde{B}$ where the test set was the extreme cases. The filtered test accuracies decreased for these categories which indicate that it was more difficult to classify the extreme cases when the data had been trained on filtered data. These results could indicate that when the data is filtered around the beta oscillations, the MSNA inhibition response is more pronounced. Thus the networks had an easier time of classifying the individuals that have MSNA inhibition values closer to each other, like for category A and B where the test set was random. When the test set contained extreme cases, the network had a more difficult time classifying them, since the network was trained on individuals further from their values.

Also another interpretation of the results is that the filtration succeeded in concentrating the area for the MSNA response but the filtered data might not contain enough information or contained too few general similarities between all individuals for the network to learn on the binary level with the data available. If there would have been more variations in the data, i.e., more subjects included, it might have helped the network to learn the subtle pattern in the beta band after filtration.

5.2 Multiclass Classification

During the approach of multiclass classification, the first result that was noted was the filtration results of *X19*. There, all accuracies decrease and quite substantially as well. Expectations for *X19* in this approach, was that a large decrease would be observed. This was due to the view that the filtration would isolate the brain signals coupled with MSNA, which were believed to be less similar between individuals. Therefore, if there was a more general MSNA characteristic left in the data, then this decrease would make sense since category *X19* focuses on classifying the individual.

Another note is that the Peak sensors have the smallest decrease, which indicates that it is the MSNA response that the network is picking up. This is because the Peak sensor set should pick up the strongest activations for the MSNA response since they are the sensors closest to the Rolandic area.

With that said, the filtered results for category *X3* and *X4* were anticipated to improve compared to the unfiltered accuracies since those divisions should have better recognized the more general patterns left after filtering, as the divisions were meant to represent high- and low risk groups. The results did not entirely reflect this anticipation however and the only increase observed in test accuracy was for *X3* Peak sensors, see Figure 4.7(a). The fact that it was the Peak sensors that increased is nevertheless a positive sign. Moreover, filtration did seem to affect *X3* more positively overall compared to *X4*, which could be due to the notion that MSNA is normally distributed and the division of *X3* took that into consideration. Remember that *X3* had the class, medium risk, containing both inhibitors and non-inhibitors to reflect the normal distribution, whereas *X4* have classes where inhibitors and non-inhibitors are distinctly divided. This difference between the classes could explain why filtration affected the categories differently. When filtering the MEG data, it could be that a more general MSNA characteristics had been captured, as discussed above. Therefore, in addition to the normal distribution of MSNA, *X3* with the property of normal distribution, was less negatively affected by filtration, and even an increase for the Peak sensors, which have captured the stronger signal, could be seen. Regardless, it appears as if the approach of pure-frequency filtering does not produce the desired effect.

In addition to investigating the effect of filtration, *X3* and *X4* were constructed for the purpose of exploring the multiclass classification. Improvements were made but not considerably and the question of what the networks have learned still remains. Since the test set consisted of 10 % of data of each subject, it could be that the networks have learned the individual MEG data instead of the wanted MSNA-features and thus is able to classify on an individual level even if the subjects now are divided into different classes. This idea can potentially be reflected in the MLP results for the unfiltered *X* categories. As the subjects were divided into 3 and 4 classes the results were still quite similar or had only slight improvements. On the other hand, the results of the LSTM network showed changing accuracies for the different categories, suggesting that the network might have learned something common within the classes. This is because, when having the network classifying different classes instead of individuals the classification accuracy increased with increasing chance level. Once again, we come back to the question of whether there are a sufficient number of subjects for this type of classification. It could be that the number of subjects in each class was insufficient or that the variety of the data was not enough for the network to learn the common pattern for the concerning class.

5.3 Future Work

The multiclass classification executed in this thesis had 10 % of data from all available epochs as the test set. This was to match and build upon the execution process of the previous master's thesis. Another way to test the multiclass classification is to have one subject from each class as test data instead of having a fraction of data from all subjects. This approach would give additional information about the multiclass classification and what the networks are learning. Even so, the concern of having too few subjects in each class is even greater for this approach. Nevertheless, to continue in this direction with another approach of multiclass classification could provide more information about the networks. Other approaches such as constructing a confusion matrix or implementing leave-one-out cross-validation (LOOCV) could also be used to get more information about the networks and how they are classifying the test set. To improve the classification however, it is more attractive to investigate approaches using another conventional neuroimaging analysis, namely, time-frequency analysis.

Pure frequency-based filtering and division into multiclasses did not yield the desired results of an improved classification. The obtained classification results suggest that the networks find it difficult to classify the MSNA response profile in the available data. An idea emerged that it could be because the networks had difficulty in detecting the frequency-based changes in power present in the MEG data before and after frequency filtration. It was therefore decided to explore whether time-frequency analysis could be implemented to transform the frequency-based changes in the MEG data and overcome this challenge. Time-frequency analysis adds one more dimension to the data which enables one to analyze the time and frequency content of the data at the same time. Furthermore, time-frequency analysis also allows one to analyze the data in one dimension, e.g., only taking one single frequency band or by averaging the frequencies in the frequency domain. This investigation was started in this thesis but did not reach the end-point of classification results, which is why it is not presented in the main report. For details of the implemented approach see Appendix A. The investigation resulted in a framework, based on ongoing research at MedTech West, for how the MEG data could be presented as time-frequency representations and by averaging over the frequencies in these representations the data can be shaped into a format that can be fed into the developed ANNs. By shifting the data into the time-frequency domain the oscillatory changes in amplitude simply become changes in amplitude. Which is presumably easier to classify, both for humans and machines.

The MEG data contains information in both time and frequency domain, which explains the possibility of utilizing time-frequency analysis to present the data to the neural networks in a different manner that may be easier to learn from. Currently utilized networks in this thesis analyze data vectors. One disadvantage of the approach utilizing these ANNs on the MEG data is that only one domain can be analyzed at a time and the information relative to the other domain is lost. Thus, another approach worth exploring is the approach utilizing a convolutional

neural network (CNN) for analyzing the time-frequency content of the MEG data, in the form of images including both time and frequency domain of the data, see Figure A.1 in Appendix A. Indeed, image classification is an advanced field of deep learning where CNN is capable of analyzing data with more than one dimension, specifically images, and exceeds in detecting features in the data. By translating the time-frequency content of the MEG data into an image, such as a time-frequency representation, all information of the MEG data will be preserved. A CNN can then be used to analyze the data. Potentially a specific feature of the MEG data, correlated to MSNA, could be found and an improved classification could be achieved. Furthermore, this type of data would give a visual component to what the network is learning from. A heatmap can be obtained to visualize what parts in the image the network sees as important for prediction. How this can be obtained we refer to [34]. This approach using CNN would probably require more input data to achieve better classification but even without more data it could be utilized in the aspect of investigating what neural networks can learn from the MEG data.

6

Conclusion

Pure frequency filtration of the MEG data around the beta oscillations did not improve the classification accuracy considerably nor did changes in multiclass grouping. With that said, the results did provide us with some insights that could put the solution closer to classify the risk of hypertension. Firstly, the resulting accuracies after filtration indicate that the developed ANNs must have, at least partially, learned from information in frequencies outside the beta band. This means that there could be other frequencies in the data, apart from the beta rebound, that are correlated with MSNA inhibition. Therefore, it could be worthwhile examining a larger range of frequencies compared to the beta band. Secondly, the filtration approach, where only the beta band was selected to train the network, did not improve the accuracies when it is known from previous studies that oscillations in the beta band are correlated with MSNA inhibition. Another explanation for the filtered results is therefore that the developed ANNs do not lend themselves well to learning the amplitude changes of the oscillatory signals in data recorded with MEG. This could be tackled using time-frequency analysis, since it can transform the data into a format that is potentially easier for the network to classify. Lastly, multiclass classification has potential for classifying risk groups rather than individuals but results showed that it is probably limited by the number of subjects available for the study. This limitation also applies to the binary classification, thus indicating a need for an expanded data set for classification with the developed ANNs. In conclusion, further investigations are necessary in order to understand what the network is learning from the MEG data and functional neuroimaging methods can be explored to improve classification. The next promising step is to utilize time-frequency analysis.

Bibliography

[1] A. Bakidou and J. Nord Odhner, "Deep Learning for Brain Activity Analysis," Master's Thesis, Chalmers University of technology, 2020. [Online]. Available: https://www.syntronic.com/wp-content/uploads/2020/08/Deep-Learning-for-Brain-Activity-Analysis.pdf

[2] B. R. Syeda, "Bringing MEG towards clinical applications," Doctoral thesis, University of Gothenburg. Sahlgrenska Academy, 2018. [Online]. Available: http://hdl.handle.net/2077/56334

[3] C. M. M. Lawes, S. V. Hoorn, and A. Rodgers, "Global burden of blood-pressure-related disease, 2001," *The Lancet*, vol. 371, no. 9623, pp. 1513–1518, 2008, DOI: 10.1016/S0140-6736(08)60655-8.

[4] M. Timio *et al.*, "Blood pressure trend and cardiovascular events in nuns in a secluded order: A 30-year follow-up study," *Blood Pressure*, vol. 6, no. 2, pp. 81–87, 1997, DOI:10.3109/08037059709061804.

[5] N. K. Hollenberg *et al.*, "Aging, acculturation, salt intake, and hypertension in the kuna of panama," *Hypertension*, vol. 29, no. 1 Pt 2, pp. 171–6, 1997, DOI:10.1161/01.hyp.29.1.171.

[6] V. Donadio *et al.*, "Muscle sympathetic response to arousal predicts neurovascular reactivity during mental stress," *The Journal of physiology*, vol. 590, no. 12, pp. 2885–2896, 2012, DOI:10.1113/jphysiol.2012.228981.

[7] J. L. Greaney and W. L. Kenney, "Measuring and quantifying skin sympathetic nervous system activity in humans," *Journal of neurophysiology*, vol. 118, no. 4, pp. 2181–2193, 2017, DOI: 10.1152/jn.00283.2017.

[8] W. Nicholson Price, "Big data and black-box medical algorithms," *Science translational medicine*, vol. 10, no. 471, p. 5333, 2018, DOI: 10.1126/scitranslmed.aao5333.

[9] A. Adadi and M. Berrada, "Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," *IEEE Access*, vol. 6, pp. 52 138–52 160, 2018, DOI: 10.1109/ACCESS.2018.2870052.

[10] "General Data Protection Regulation (GDPR)," 2016, [Online]. Available: https://eur-lex.europa.eu/eli/reg/2016/679/oj (Accessed: Jan. 27, 2021).

[11] World Health Organization (WHO), *A global brief on hypertension: Silent killer, global public health crisis.* Geneva, Switzerland: WHO, 2013.

[12] A. W. Cowley, "The genetic dissection of essential hypertension," *Nature Reviews Genetics*, vol. 7, no. 11, pp. 829–840, 2006, DOI: 10.1038/nrg1967.

[13] O. Carretero and S. Oparil, "Essential Hypertension: Part I: Definition and Etiology," *Circulation*, vol. 101, pp. 329–35, 2000, DOI: 10.1161/01.CIR.101.3.329.

[14] J. A. Waxenbaum, V. Reddy, and M. Varacallo, *Anatomy, Autonomic Nervous System.* Treasure Island (FL): StatPearls Publishing, 2020, [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK539845/ (Accessed: Apr. 10, 2021).

[15] J. M. Karemaker, "An introduction into autonomic nervous function," *Physiological Measurement*, vol. 38, no. 5, pp. R89–R118, 2017, DOI: 10.1088/1361-6579/aa6782.

[16] A. J. Manolis, L. Poulimenos, M. Kallistratos, I. Gavras, and H. Gavras, "Sympathetic Overactivity in Hypertension and Cardiovascular Disease," *Current Vascular Pharmacology*, vol. 12, no. 1, pp. 4–15, 2014, DOI: 10.2174/15701611113119990140. PMID: 23905597.

[17] T. Mano, "Muscle sympathetic nerve activity in blood pressure control against gravitational stress," *Journal of Cardiovascular Pharmacology*, vol. 38, pp. S7–11, 2001, DOI: 10.1097/00005344-200110001-00003.

[18] K. A. Maldonado and K. Alsayouri, *Physiology, Brain.* Treasure Island (FL): StatPearls Publishing, 2021, [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK551718/?report=classic (Accessed: Apr. 19, 2021).

[19] W. Klimesch, "The frequency architecture of brain and brain body oscillations: an analysis," *European Journal of Neuroscience*, vol. 48, no. 7, pp. 2431–2453, 2018, DOI: 10.1111/ejn.14192.

[20] A. E. Symons, W. El-Deredy, M. Schwartze, and S. A. Kotz, "The functional role of neural oscillations in non-verbal emotional communication," *Frontiers in Human Neuroscience*, vol. 10, p. 239, 2016, DOI: 10.3389/fnhum.2016.00239.

[21] A. K. Engel and P. Fries, "Beta-band oscillations–signalling the status quo?" *Current Opinion in Neurobiology*, vol. 20, no. 2, pp. 156–65, 2010, DOI: 10.1016/j.conb.2010.02.015.

[22] Injurymap. *Brain* [Online image]. Injurymap. https://www.injurymap.com/free-human-anatomy-illustrations.

[23] M. Hämäläinen, R. Hari, R. J. Ilmoniemi, J. Knuutila, and O. V. Lounasmaa, "Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain," *Reviews of Modern Physics*, vol. 65, no. 2, pp. 413–497, 1993, DOI: 10.1103/RevModPhys.65.413.

[24] S. Baillet, "Magnetoencephalography for brain electrophysiology and imaging," *Nature Neuroscience*, vol. 20, no. 3, pp. 327–339, 2017, DOI: 10.1038/nn.4504.

[25] A. Widmann, E. Schröger, and B. Maess, "Digital filter design for electrophysiological data – a practical approach," *Journal of Neuroscience Methods*, vol. 250, pp. 34–46, 2015, DOI: 10.1016/j.jneumeth.2014.08.002.

[26] A. de Cheveigné and I. Nelken, "Filters: When, Why, and How (Not) to Use Them," *Neuron*, vol. 102, no. 2, pp. 280–293, 2019, DOI: 10.1016/j.neuron.2019.02.039.

[27] V. Renganathan, "Overview of artificial neural network models in the biomedical domain," *Bratislava Medical Journal*, vol. 120, no. 7, pp. 536–540, 2019, DOI: 10.4149/bll_2019_087.

[28] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, pp. 1-12, pp. 162-220, pp. 367-412, [Online]. Available: http://www.deeplearningbook.org.

[29] A. Meyer-Baese and V. Schmid, *Chapter 7 - Foundations of Neural Networks*. Oxford: Academic Press, 2014, pp. 197–243. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780124095458000078

[30] S. Sharma and S. Sharma, "Activation functions in neural networks," *Towards Data Science*, vol. 6, no. 12, pp. 310–316, 2017, ISSN: 2455-2143.

[31] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, "Explainable AI: A Review of Machine Learning Interpretability Methods," *Entropy*, vol. 23, no. 1, 2020, DOI: 0.3390/e23010018.

[32] A. Gramfort *et al.*, "MEG and EEG data analysis with MNE-Python," *Frontiers in Neuroscience*, vol. 7, no. 267, pp. 1–13, 2013, DOI: 10.3389/fnins.2013.00267.

[33] S. H. Park, J. Choi, and J. S. Byeon, "Key principles of clinical validation, device approval, and insurance coverage decisions of artificial intelligence," *Korean journal of radiology*, vol. 22, no. 3, pp. 442–453, 2021, DOI: 10.3348/kjr.2021.0048.

[34] V. Petsiuk, A. Das, and K. Saenko, "Rise: Randomized input sampling for explanation of black-box models," *arXiv preprint arXiv:1806.07421*, 2018.

[35] E. Sejdić, I. Djurović, and J. Jiang, "Time–frequency feature representation using energy concentration: An overview of recent advances," *Digital Signal Processing*, vol. 19, no. 1, pp. 153–183, 2009, DOI: 10.1016/j.dsp.2007.12.004.

[36] "Time-frequency analysis using hanning window, multitapers and wavelets," *FieldTrip*, 2020, [Online]. Available: https://www.fieldtriptoolbox.org/workshop/paris2019/handson_sensoranalysis/#time-frequency-analysis-using-hanning-window-multitapers-and-wavelets (Accessed: June 8, 2021).

[37] R. Oostenveld, P. Fries, E. Maris, and J. M. Schoffelen, "FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data," *Computational Intelligence and Neuroscience*, vol. 2011, 2011, Article ID 156869, DOI: 0.1155/2011/156869.

A

Time-frequency Analysis

Recorded data can be analyzed in time domain and/or frequency domain depending on the content of the data and what information is desired from the data. When analyzing data in one domain, the information in the data in the other domain is not described which means that the information acquired may be inadequate for subsequent analyzes or processes such as feature classification or recognition. Thus, for data containing information in both time and frequency domain, e.g., MEG data, it would be more advantageous to utilize time-frequency analysis to obtain a time-frequency representation of the data describing the time and frequency contents simultaneously [35]. With time-frequency analysis, one can, among other things, analyze the time-frequency content or single frequency bands. Different methods such as Fourier analysis, utilizing tapers (single or multiple), or wavelet analysis, utilizing wavelets, can be used to perform time-frequency analysis. The time-frequency representation of the data is obtained with the help of a sliding time window, with a window function of choice [36].

A.1 Framework

The approach of utilizing time-frequency analysis was based upon ongoing research at MedTech West. The code for implementing time-frequency analysis was performed in MATLAB and the open source package Fieldtrip [37] was used for the actual transformation and visualization of the time-frequency representations. For the time-frequency analysis the data set had a slightly different structure then in the main report to include a baseline epoch for each pulse train. It was divided into epochs and cropped from 0 ms (stimulus trigger) to 1500 ms and contained a baseline, of the same length, that was used to normalize the data with baseline correction. The baseline was taken pre-stimulus so that no arousal would be present in the baseline. The separate pulses and the baseline were concatenated into a combined data structure representing the pulse train with a baseline at the beginning. To create the time-frequency representations a 7-cycle Hanning-tapered sliding window was used, which means that the time window is dynamic and depends on the frequency. There were always 7 cycles (periods) in each sliding window, i.e., higher frequencies are going to have shorter time windows. The time-frequency representation is then plotted for visualization. Figure A.1 present the time-frequency representation of the MEG data from subject 0295, when an average is taken over all epochs. In

the time-frequency representation, the power intensity, in the unit dB, of the neural activity during the received stimuli is presented over both time (in the x-axis) and frequency (in the y-axis). The Figure A.1, from left to right, represents the baseline, pulse 1, pulse 2, and pulse 3. The pulses were described in the main report and represent the response of the received stimuli. An interesting thing to note is the expressed brain activity around the beta band in all pulses, which complies with previous findings that beta oscillations are correlated to the MSNA inhibition response.

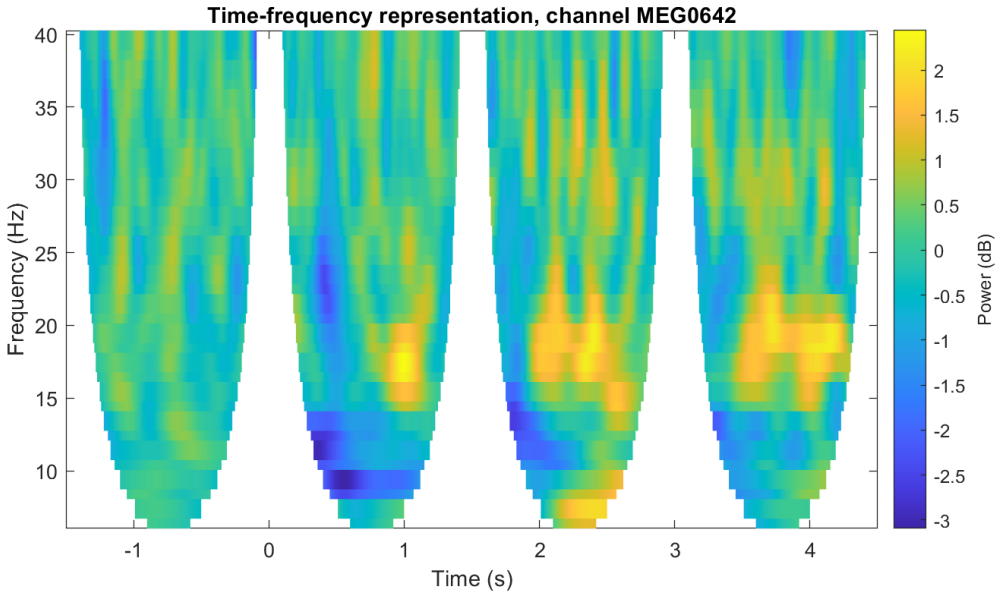


Figure A.1: *Time-frequency representation of the MEG data from subject 0295 for the channel MEG0642. The time-frequency representation is an average over all epochs. The activities from left to right describes the baseline, pulse 1, pulse 2, and pulse 3. The brain activity is expressed in power intensity over time and frequency.*

If the time-frequency representations would be used in the already developed neural networks, then they need to be shaped into a format that the networks can handle. Usually, time-frequency representation for the MEG data is produced by taking average over all epochs for each pulse and MEG channel. This is to reduce noise in the signal. However, for the purpose of having sufficient data to train the network, a time-frequency representation was created for each pulse train and MEG channel without taking the average over the epochs or MEG channels. Consequently, this would mean that the data contains more noise, but the networks might still find the pattern for MSNA inhibition. One of the conclusions drawn in this thesis was that the neural networks may have it difficult to analyze data with power changes in the frequency domain, thus, this complication was dealt with by averaging over the frequencies, resulting in an average power over the frequencies. An example of the averaged data, for subject 0295 and MEG channel MEG0642, is presented in Figure A.2.

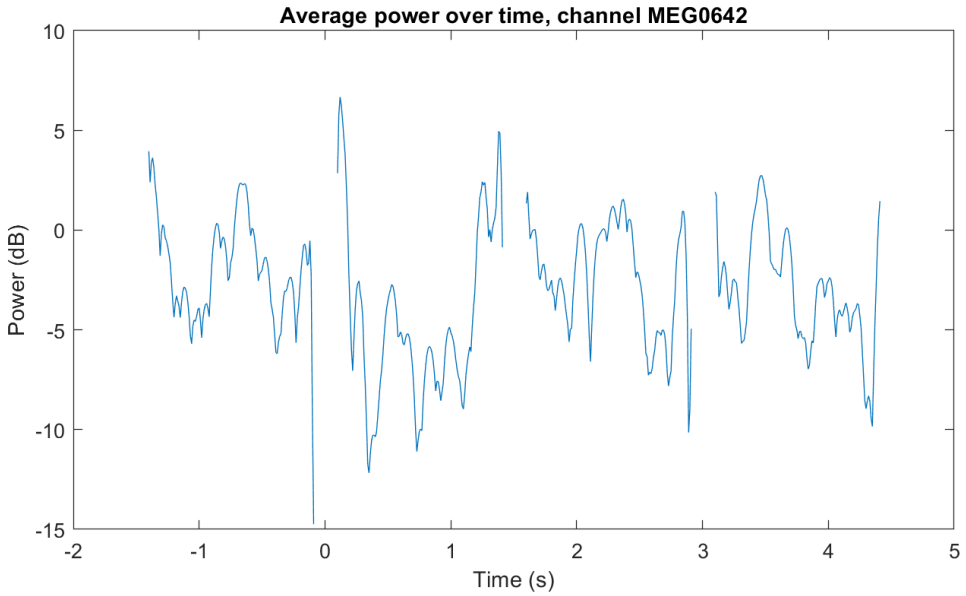


Figure A.2: *Average power over time. The frequency-based changes in power are averaged resulting in one average value of the power in the frequency domain.*

The next step, which this thesis did not implement due to lack of time, would be to convert the data into the correct format from MATLAB to Python and integrate the data with the already developed ANNs. Potentially the Python module `filedtrip2MNE` could be used to accomplish the format transformation.

DEPARTMENT OF ELECTRICAL ENGINEERING
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

CHALMERS
UNIVERSITY OF TECHNOLOGY