

POLITECNICO DI TORINO

**MASTER OF SCIENCE COURSE
IN BIOMEDICAL ENGINEERING**

MASTER OF SCIENCE THESIS

**EEG Signal Processing for Semantic
Feature Extraction: A Study on Perception,
Imagination, and Vividness.**



**Politecnico
di Torino**

SUPERVISOR

PROF. LUCA MESIN

CO-SUPERVISOR

PHD. HOSSEIN AHMADI

CANDIDATE

MARTINA IMPAGNATIELLO

POLITECNICO DI TORINO
MASTER OF SCIENCE COURSE
IN BIOMEDICAL ENGINEERING

MASTER OF SCIENCE THESIS

**EEG Signal Processing for Semantic
Feature Extraction: A Study on Perception,
Imagination, and Vividness.**



**Politecnico
di Torino**

SUPERVISOR

PROF. LUCA MESIN

CO-SUPERVISOR

PHD. HOSSEIN AHMADI

CANDIDATE

MARTINA IMPAGNATIELLO

Index

1 ELECTROENCEPHALOGRAM	1
1.1 BRAIN ANATOMY	1
1.2 NEUROPHYSIOLOGICAL FOUNDATIONS	5
1.3 METHODS FOR RECORDING BRAIN SIGNALS	6
1.4 HISTORICAL MILESTONES IN ELECTROENCEPHALOGRAPHY	8
1.5 MOST COMMON APPLICATIONS OF EEG	8
1.6 ERP	9
1.7 EEG RECORDING TECHNIQUES	12
1.8 ARTIFACTS	14
1.9 FEATURES EXTRACTION	17
2 SEMANTIC FEATURES	19
2.1 NEURONAL MECHANISMS OF SEMANTIC MEMORY	20
2.1.1 CONTEXT DEPENDENCE	20
2.2 BRAIN AREAS INVOLVED IN SEMANTIC FEATURE EXTRACTION	21
2.3 EVENT RELATED POTENTIALS (ERPs) IN SEMANTIC PROCESSING	23
2.3.1 WORD2VEC	25
2.3.2 GLOVE	25
2.4 OSCILLATORY ANALYSIS	25
2.4.1 GAMMA BAND OSCILLATION IN SEMANTIC FEATURE EXTRACTION	26
3 METHODOLOGY	27
3.1 DATASET DESCRIPTION	27
3.1.1 VVIQ	27
3.1.2 BAIS	29
3.2 DATA PROCESSING AND FEATURE EXTRACTION	29
3.3 CLASSIFICATION	30
3.3.1 SUPPORT VECTOR MACHINE	30
3.3.2 RANDOM FOREST	31
3.3.3 LOGISTIC REGRESSION	31
3.3.4 MULTILAYER PERCEPTRON	31
3.4 STATISTICAL ANALYSIS	32
4 RESULTS	34

4.1 AUDIO MODALITY	34
4.2 ORTHOGRAPHIC MODALITY	41
4.2 PICTORIAL MODALITY	48
5 DISCUSSION.....	56
5.1 AUDIO MODALITY	56
5.2 ORTHOGRAPHIC MODALITY	57
5.3 PICTORIAL MODALITY	57
6 CONCLUSIONS.....	59
REFERENCES	60

*To my brother.
With you,
I will always find the way back home,
the light still on.*

Abstract

The extraction of semantic features from EEG signals presents a novel approach to understanding the neural mechanisms underlying perception, imagination, and vividness. This study examines the correlation between EEG-based brain activity and subjective imagery ratings derived from the Vividness of Visual Imagery Questionnaire (VVIQ) and the Bucknell Auditory Imagery Scale (BAIS-V). We used a dataset containing recordings of 15 subjects subjected to audio, pictorial and orthographic stimuli. Before proceeding with feature extraction, the signal was appropriately preprocessed. As regards feature extraction, Power Spectral Density (PSD) and frequency bands were used. In this way we ensure that the neural representations that we consider are significant. To classify perception and imagination states, we implemented machine learning models—Support Vector Machines (SVM), Random Forest (RF), Logistic Regression (LR), and Multilayer Perceptron (MLP)—and evaluated their performance using accuracy, area under the curve (AUC), and cross-validation techniques. In the end, to reveal the influence of self-reported vividness scores on classification accuracy, statistical analysis were performed. The findings indicate that EEG-based features have a great potential for modeling semantic cognition, offering valuable insights into the neural basis of mental imagery. This research contributes to advancing EEG-based brain-computer interface (BCI) applications and cognitive neuroscience studies on imagination and perception.

1 ELECTROENCEPHALOGRAM

1.1 BRAIN ANATOMY

The cerebrum is divided into two hemispheres—the right and the left. Each hemisphere is responsible for processing sensory inputs, controlling movement, language, emotions, learning, and reasoning. Intriguingly, the two hemispheres govern opposite parts of the body and have specialized functions: the left hemisphere, often considered the rational side, regulates language, writing, and logical reasoning; conversely, the right hemisphere is deemed the artistic side, overseeing creativity, intuition, and musical abilities [1].

Beneath the cerebrum lies the cerebellum, which maintains balance and attention. The brainstem, comprising the midbrain, medulla oblongata, and pons, regulates autonomic functions such as breathing, temperature control, and heartbeat. It also serves as a connection between the cerebrum, cerebellum, and spinal cord.

The brain's surface is enveloped by the cerebral cortex, commonly called the gray matter, which lies above the axons that make up the white matter. Table 1-1 summarizes the key attributes of these two types of brain tissue [2]:

GREY MATTER	WHITE MATTER
Comprises neuron bodies, dendrites, and unmyelinated nerve fibers.	Consists of myelin-covered fibers.
Located in the brain's cortex.	Situated beneath the cortex's gray matter.
Plays a crucial role in processing and initiating information and serves as a point for motor input.	Facilitates the connection and interaction of motor stimuli.

Table 1.1: Distinctions between white matter and gray matter in terms of, content, location and function.

The human brain is anatomically divided into four lobes, each associated with distinct functions as illustrated in Figure 1.1:

- 1) **Frontal lobe:** This lobe is pivotal for controlling voluntary movements, such as walking and running. It is critical in formulating and storing long-term memories and translating thoughts into words. The frontal lobe is also essential for managing attention, empathizing with others, and classifying objects. Additionally, it governs the so-called reward system and the capacity to foresee the consequences of specific decisions. It significantly influences personality traits and individual behavioral characteristics [3].
- 2) **Parietal lobe:** This lobe processes information relating to tactile sensations, including pain, temperature (such as cold and heat), and more. The upper region aids in spatial orientation,

while the lower part works with the frontal lobe to support memory, interpret language, and enhance calculation abilities [3].

- 3) **Occipital lobe:** Primarily involved in visual perception, this lobe facilitates reading and the recognition of colors, shapes, and movements. It is also crucial for depth perception and identifying moving objects [3].
- 4) **Temporal lobe:** It is involved in the perception of sounds, the interpretation of visual stimuli, and object recognition. This lobe plays a crucial role in understanding spoken and written language and maintaining long-term memory [6].

Each of these lobes contributes uniquely to the brain's complex functionalities, enabling a diverse range of human cognitive and physical abilities.

Human Brain Anatomy

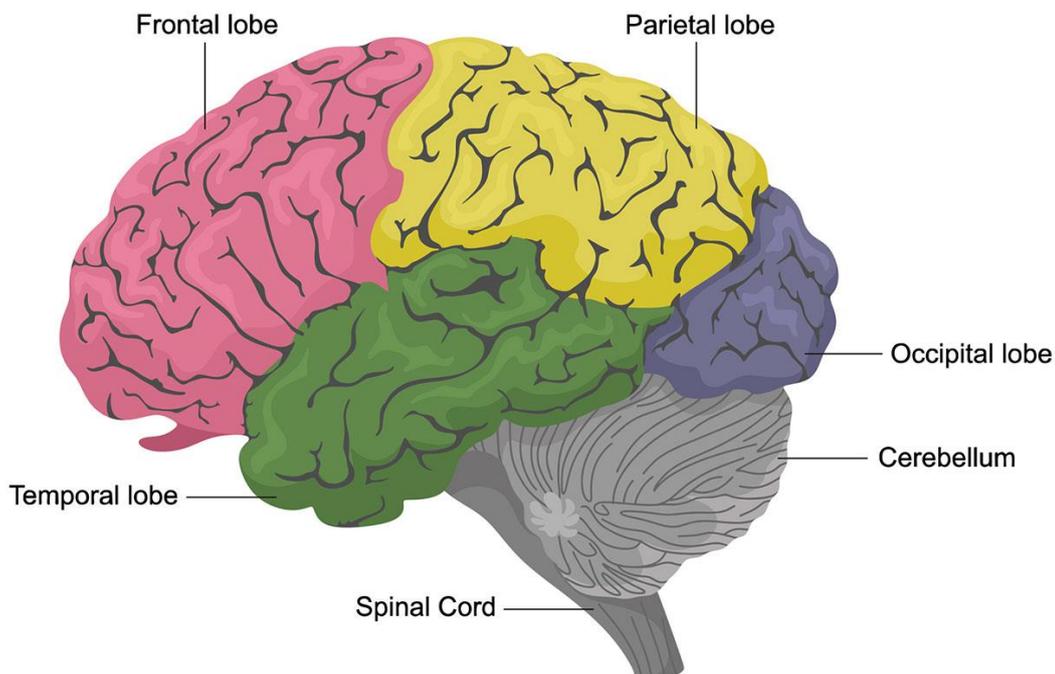


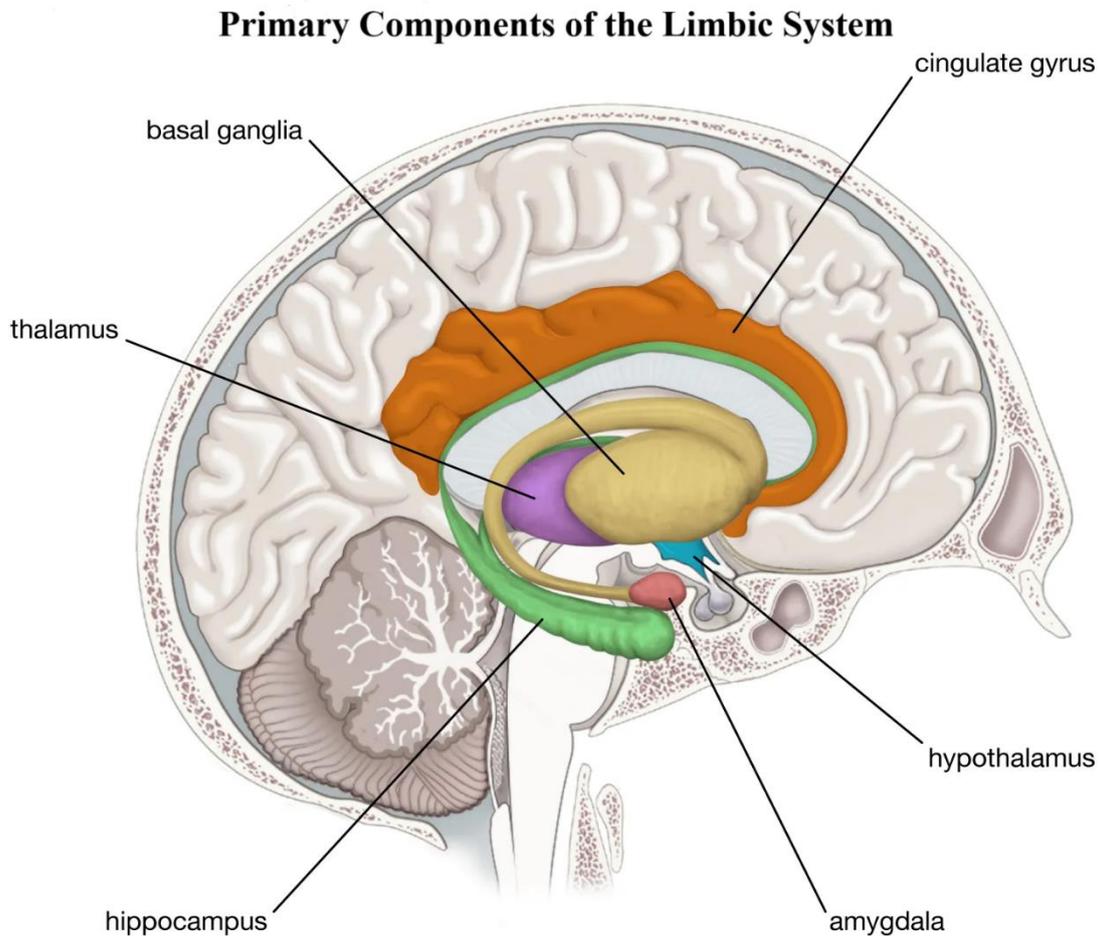
Figure 1.1: Human Brain Anatomy. This illustration delineates the four principal lobes of the brain and highlights their connectivity to the cerebellum and spinal cord.

As illustrated in Figure 2.1, certain brain lobes achieve specific functions through communication facilitated by the limbic system. This system encompasses several vital structures:

- **Hypothalamus:** This region governs the autonomic nervous system (ANS) and regulates fundamental needs such as hunger, thirst, and sleep cycles.
- **Thalamus:** Acts as a communication hub between the cerebral hemispheres, enhancing inter-hemispheric coordination.
- **Hippocampus:** Essential for memory formation and recall, pivotal in storing long-term memories.

- **Amygdala:** Functions as the emotional center of the brain, predominantly managing responses related to fear and other emotional processes.

All these components of the limbic system, work together to highlight psychological processes. They are characterized by specific functions and their efforts give information about complex behavioral and emotional responses [1].



© Encyclopædia Britannica, Inc.

Figure 2.1: Limbic System. This figure illustrates the primary components of the limbic system: the thalamus, hypothalamus, hippocampus, and amygdala.

In addition to the limbic system, the brain contains two pivotal endocrine glands: the pituitary gland, which orchestrates the body's endocrine system, and the pineal gland (epiphysis), which regulates sleep-wake cycles and melatonin production.

Neurons, the principal cells of the brain, feature a central body with extensions called dendrites and axons. Dendrites receive signals, while axons transmit them. The brain functions analogously to a vast electrical circuit, facilitating communication by electrical impulses initiated by stimulus perception at the neuron. These impulses travel to the axon, triggering the release of neurotransmitters that act as messengers. These neurotransmitters cross synapses to connect to receptors on the receiving cell. Additionally, the brain contains glial cells, which nourish and support neurons.

Graded potentials and action potentials are two types of electrical signals that mediate the neuronal communication by influencing ion channels activity. The former electric signal operates over short distances and vary proportionally to the stimulus that induced them. While, action potentials are electrical signal that can operates over long distance. They are characterized by a rapid reversal of membrane polarity and they have the capability to not diminish their strength despite the distance [4].

To understand the utility of these electrical signals further, it is essential to define hyperpolarization, depolarization, and repolarization. A neuron's membrane is polarized at a resting potential of -70mV . Hyperpolarization occurs when the potential moves toward more negative values, while depolarization occurs as it shifts toward less negative values. Repolarization is the process by which the membrane potential returns to its resting state. Depending on the affected neuron, the type of stimulus, and the specific ion channel involved, graded potentials can either be depolarized or hyperpolarized (Figure 3.1).

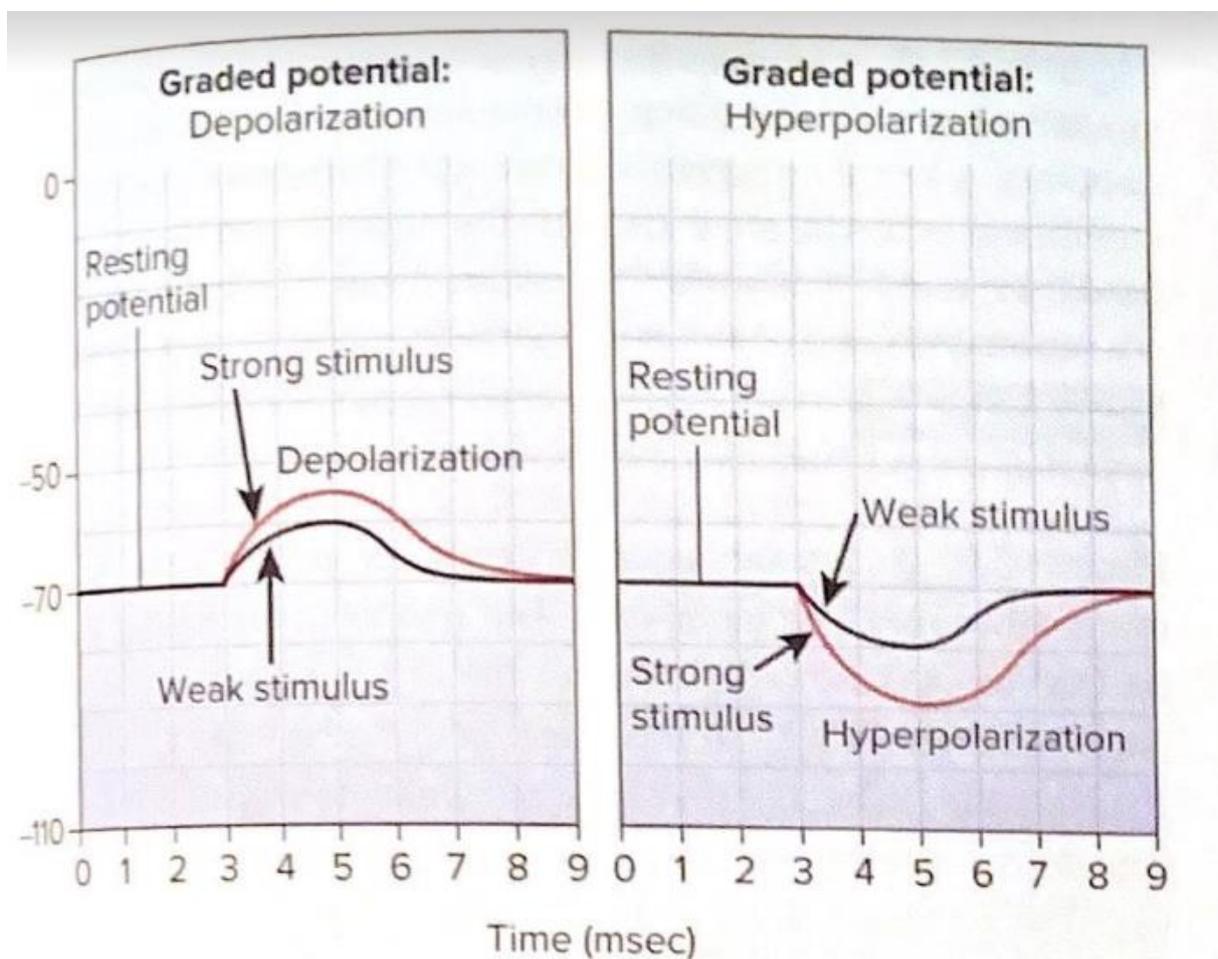


Figure 3.1: Depolarization and Hyperpolarization Phases. These diagrams illustrate how neuronal responses vary with the intensity of the stimulus, depicting the phases of depolarization and hyperpolarization.

The action potential operates through a slightly more complex mechanism. It is initiated at the axon hillock when a graded potential depolarizes the membrane to -55mV . The action potential progresses through three phases (Figure 4.1):

1) Rapid Depolarization: As sodium ions (Na^+) enter the cell, the membrane potential shifts from -55mV to $+30\text{mV}$. Due to high sodium conductance, the membrane potential approaches the sodium equilibrium potential of $+58\text{mV}$. A polarity reversal occurs at the peak of this phase.

2) Repolarization: The potential returns from $+30\text{mV}$ back to -70mV . During this phase, sodium conductance ceases entirely, and potassium (K^+) conductance increases. The efflux of potassium ions resets the potential to its resting value.

3) Posthumous Hypopolarization: Before stabilizing at resting value, the membrane potential briefly becomes more negative, approximating the potassium equilibrium potential. The closure of voltage-gated potassium channels ultimately restores the resting potential [4].

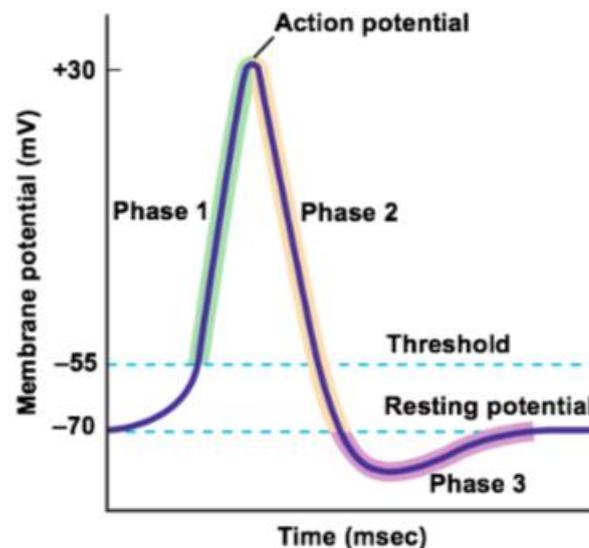


Figure 4.1. Development of an Action Potential. This illustration shows the progression of an action potential. Specifically in phase 1 of rapid depolarization it reaches a peak, it is followed by phase 2 of repolarization, and in the final phase it is characterized by hyperpolarization, which includes a negative peak before returning to the resting potential.

1.2 NEUROPHYSIOLOGICAL FOUNDATIONS

The brain comprises two principal types of cells: neurons and glial cells. Both cell types maintain a resting potential of -80mV . The electrical activity in neurons arises from generating action potentials and postsynaptic potentials. Action potentials occur when the potential difference across the cell membrane surpasses a specific threshold, whereas postsynaptic potentials are subthreshold events. More specifically, evoked potentials result from increased membrane permeability to sodium ions, causing the internal potential of the cell to shift from negative to positive (approximately $+30\text{mV}$). Following this, the membrane enhances potassium ion permeability while reducing sodium permeability, re-negativizing the potential and generating a spike-shaped action potential that lasts about one millisecond.

Postsynaptic potentials (PSPs), with typical magnitudes of $5\text{-}10\text{mV}$, occur at the postsynaptic membrane. When an action potential reaches a synapse, it triggers the release of a neurotransmitter, facilitating ion passage across the membrane and creating a potential difference. If the neuron's internal negativity decreases, an excitatory postsynaptic potential (EPSP) is generated; conversely, an increase in negativity results in an inhibitory postsynaptic potential (IPSP).

1.3 METHODS FOR RECORDING BRAIN SIGNALS

Among the known methods for recording brain signals are Magnetoencephalography (MEG) and Electroencephalography (EEG).

Magnetoencephalography allows measuring the magnetic field generated by neurons. Like EEG, it is a non-invasive technique, often combined with magnetic resonance imaging (MRI) to obtain a structural perspective of the magnetic source. This technique has very high temporal resolution and good spatial resolution. From the anatomy of the brain, it is known that neurons are traversed by electrical current. This electrical current is always associated with a magnetic field perpendicular to its direction, as taught by the right-hand rule.

MEG detects tangential dipoles since the radial dipole, remaining within the cranial cavity, has a magnetic field that cannot be detected. MEG fields are not distorted because the magnetic permeability of biological tissues is very similar to that of space in a vacuum, so neither the scalp nor the skull will be disturbing elements, which is extremely advantageous from the point of view of spatial and temporal resolution. This technique is used especially for epilepsy surgery and preoperative brain mapping, being more sensitive to superficial cortical activity.

Like all techniques for acquiring biological signals, MEG also presents advantages and disadvantages, as shown in Table 2.1.

ADVANTAGES	DISADVANTAGES
Ability to track temporal dynamics of events	Requires sensors in superconducting regime
No signal distortion	Complex management
Utilizes specialized signal processing algorithms	High cost
High spatial and temporal resolution	

Table 2.1: Advantages and disadvantages of MEG.

Electroencephalography involves the recording and analysis of the electroencephalogram, a term derived from the Greek words 'enkephalo' (meaning brain) and 'graphein' (meaning to write). An electroencephalogram represents the recording of electrical signals produced by the collective activity of brain cells. These signals can be measured using electrodes placed on the scalp, known as an electroencephalogram, or directly on the cortex, which is referred to as an electrocorticogram (Figure 5.1).



Figure 5.1: EEG setting example

In EEG studies, five distinct frequency bands have been identified, each associated with different brain states (Figure 6.1):

1. Delta (0.5-3Hz): Delta waves are characteristic of deep sleep and exhibit amplitudes ranging from 75mV to 200mV.
2. Theta (3-7Hz): Although rare in adult humans, theta waves are predominantly observed in rodents and are crucial for transferring information between brain structures.
3. Alpha (7-14Hz): Commonly observed during waking, alpha waves are prominent when the individual is relaxed and typically with eyes closed.
4. Beta (14-40Hz): indicates heightened alertness and cognitive engagement.
5. Gamma (40-60Hz): This range is like the brain's superpower. It's associated with processing sensory stimuli and initiating voluntary movements, like reacting quickly to a sudden sound or making a conscious decision to move the hand [5].

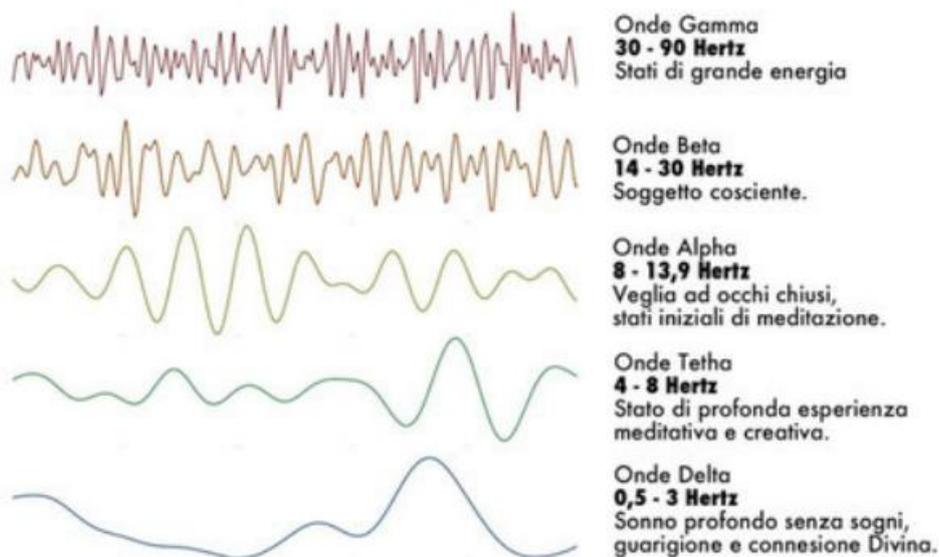


Figure 6.1: waveforms of distinct frequency bands.

Electroencephalography (EEG) has several advantages and, like any diagnostic technique, it has its pitfalls [6]. The pros and cons are summarized in Table 3.1.

ADVANTAGES	DISADVANTAGES
Ease of integration with other techniques	Requires expertise in signal cleaning
Economical	
High Temporal Resolution	Low temporal resolution
Portable and Wireless	
Non-invasive	

Table 3.1: Advantages and disadvantages of EEG examination.

Starting from basic knowledge and integrating with the characteristics contained in Tables 2.1 and 3.1, it is possible to make a comparison between these two methods of acquiring brain signals. MEG is more sensitive to tangential currents on the scalp, while EEG is sensitive to both tangential and radial activities of neurons. For detecting an epileptic spike, MEG requires only 3-4 cm², whereas EEG needs 6 to 20 cm of synchronized cortical area. MEG has an excellent spatial resolution of the source, approximately 2-3 mm, compared to EEG's 7-10 mm, but it is costly and also faces a significant engineering challenge: it requires shielding from the Earth's magnetic fields, which are stronger than those derived from the brain [7].

1.4 HISTORICAL MILESTONES IN ELECTROENCEPHALOGRAPHY

Pioneer work in electroencephalography dates back to the late 19th century Richard Caton was among the early investigators of the first potentials of the brain in the cat and the rabbit. At the same time, a researcher named Adolf Beck was investigating spontaneous electrical activity in the brains of annexed animals, finding rhythmic oscillations in the process. Beck was pioneering in his identification of oscillations and their relation to visual stimuli, leading to detection of the alpha band. In 1929 Hans Berger performed the first electroencephalography (EEG) using surface electrodes placed on human scalp. Berger noticed that the activity recorded was influenced by the state of the brain—whether the brain was relaxed, asleep, hypnotized, or hypoxic. His insights laid the foundation for modern electroencephalography by showing that brain activity is deeply affected by the subject's awareness, ease and alertness [5].

1.5 MOST COMMON APPLICATIONS OF EEG

The applications of electroencephalography extend beyond basic brainwave analysis. Among the most common applications is Neurological Polysomnography, which is useful for identifying sleep-related disorders. It is a non-invasive examination conducted during sleep, utilizing EEG traces and eye movement (EOG) to assess the correlation between brain activity and eventual body spasms. It is particularly useful for diagnosing conditions such as chronic insomnia, parasomnias, narcolepsy, excessive snoring, as well as sleep apnea and behavioral disorders during the REM phase [8].

EEG is widely used for the diagnosis of epilepsy. This is a neurological disorder caused by abnormal electrical discharge in the cerebral cortex. EEG, in a non-invasive manner, records electrical activity

through electrodes placed on the patient's skin. In addition to standard EEG, there are techniques where it is possible to film the subject being monitored. This is the Video-EEG technique, useful when both the trace and the patient need to be analyzed simultaneously.

With the advent of new technology, it is also possible to perform Dynamic EEG using portable recorders to record the trace for extended periods (24-48 hours). For a complete diagnosis, EEG is associated with Neuroimaging, particularly brain magnetic resonance imaging (MRI), which highlights areas of the brain that have been damaged. Using one diagnostic technique over another is not functional for diagnosing the disease [9].

The Brain-Computer Interface (BCI) is one of the rapidly developing fields. BCI systems are EEG systems that allow users to control an external actuator. The use of this growing technology could enable a subject to control a computer without following conventional methods, but it can also be used to monitor the subject's mental states such as emotions, concentration, or drowsiness. With this type of technology, paralyzed, amputated, or individuals with reduced neurological functionality can be supported, as well as for art and entertainment. It spans from a medical to a non-medical field [10].

One application that has attracted the attention of scientists in recent years is Motor Imagery (MI) EEG. It is an application of BCI in which the patient imagines performing a specific task without actually performing it [11]. It is an innovative technique because, besides being cost-effective, it can be used as a rehabilitative therapy after a stroke. However, it is challenging from a classification and processing standpoint due to the instability of EEG signals. After data acquisition, the data are processed by removing noise and artifacts. Subsequently, EEG features are extracted, and the most significant ones are selected to classify which motor movement was imagined by the patient [12]. Among the most well-known and widely used techniques are the Event-Related Potentials (ERP) signals, which are discussed below.

1.6 ERP

Evoked potentials are voltage changes that occur in response to internal or external stimuli. They are important for examining cognitive processes like perception, memory, emotions, and language in both adults and children, and they are also useful to examine people with neurological and psychiatric conditions. In contrast to the larger spontaneous components of electroencephalogram (EEG), the amplitudes of evoked potentials are not readily discernible in raw EEG data. Consequently, signal averaging is necessary to enhance the visibility of these potentials. As stimulations increase, the signal's morphology becomes increasingly apparent, forming an ERP signal. Averaging represents the brain's mean response to a stimulus, derived from the summation of multiple epochs synchronized with the stimulus or event. This process isolates the evoked activity by averaging out the background activity, which is random relative to the event and thus tends to diminish or cancel out, accentuating the evoked activity. Graphically, evoked potentials manifest as a series of positive and negative inflections and peaks, varying according to the wave's polarity and the latency involved; unlike an electroencephalogram that records the brain's ongoing electrical activity, an evoked potential captures specific changes in the biological signal following a sensory or motor stimulus. Each inflection indicates the presence of postsynaptic potentials, and the localization of these inflections helps identify which cortical areas are active in response to specific stimuli [13].

Evoked potentials are characterized by two primary parameters: latency and topography. Latency is the time, expressed in milliseconds, between a stimulus's application and an inflection's appearance. Topography denotes the position on the skull where the inflection reaches its maximum amplitude.

Evoked potentials can be categorized into stimulus-related potentials, contingent upon the stimulus's physical characteristics, and ERPs, which hinge on the stimulus's information content and manifest only when a subject actively attends to the stimulus. Discussion of evoked potentials typically pertains to the cognitive functions derived from recorded brain activity. Notably significant in this study are the P300 and N400 potentials.

The N400 is a negative deflection in the EEG signal employed to analyze the understanding of word semantics and sentence meanings. Conversely, the P300 is a positive deflection that occurs exclusively in response to target stimuli, indicating a non-cognitive function as it arises whenever the subject updates their spatial perception.

The most frequently utilized evoked potentials in diagnostic contexts are somatosensory, visual, and auditory. By examining the shape and latency of these potentials, abnormalities in the afferent pathways can be diagnosed, thereby quantifying the extent of alterations that could indicate chronic degenerative diseases.

One of the most prevalent types of evoked potentials is the Visual Evoked Potential (VEP), depicted in Figure 7.1. This potential is typically elicited using a checkerboard pattern of alternating dark and bright squares that reverse in brightness. The standard electrode montage for recording VEPs includes Oz, O1, and O2, referenced to either a scalp electrode or a more anterior ear electrode. Several factors can influence the VEP, including the luminance of the stimulus, contrast, field size, square size, frequency of inversion, the subject's age, visual acuity, and the ability to maintain visual fixation at the center of the checkerboard.

Among the visual evoked potentials, it is necessary to mention the SSVEP, which stands for Steady-State Visual Evoked Potentials. These involve responses to visual stimuli at specific frequencies. In particular, the retina excited between 3.5Hz and 75Hz generates brain electrical activity at the same frequency as the visual stimulus. This type of evoked signal can be useful in cases where the subject is exposed to stimuli with different flashing frequencies to determine which stimulus is of interest [14].

The response to this stimulus is characterized by a sequence of negative (N75), positive (P100), and again negative (N145) waves. The negative responses may vary among individuals, whereas the positive response, specifically the P100 wave, is consistently observed in all subjects without pathological conditions. Instances of prolonged P100 response typically indicate a defect in the optic nerve of the stimulated eye. Analyzing the characteristics of the P100 response, such as prolonged latency and interocular differences, as well as the absence or presence of this response, can be instrumental in conducting abnormality analyses.

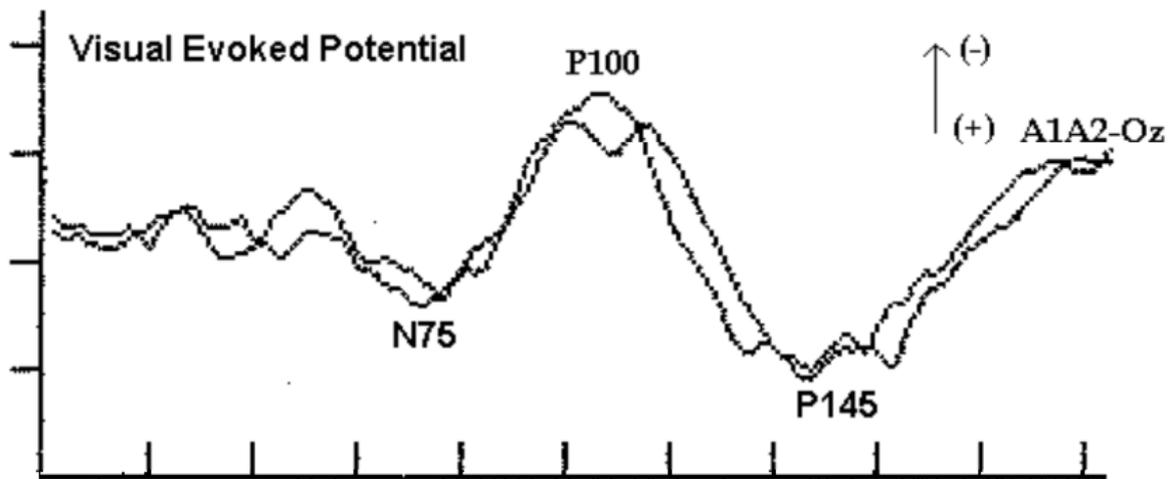


Figure 7.1: Example of a VEP. This illustration demonstrates the response to a visual sensory stimulus featuring negative N75 and N145 waves and a positive P100 wave.

Another notable evoked potential is the Brainstem Auditory Evoked Response (BAER), illustrated in Figure 8.1. These responses are elicited by monaural auditory stimuli, where white noise is delivered to the contralateral ear. The typical response includes five primary positive waves recorded from the Cz electrode, referenced to an ear electrode on the ipsilateral stimulated ear and the contralateral ears (Ac). Wave I originates from the distal portion of the VIII nerve near the cochlea, and wave II is generated by the proximal segment of the nerve. The superior olivary nucleus produces wave III. Waves IV and V arise from multiple generators within the rostral pons and the caudal midbrain, collaborating for their initiation.

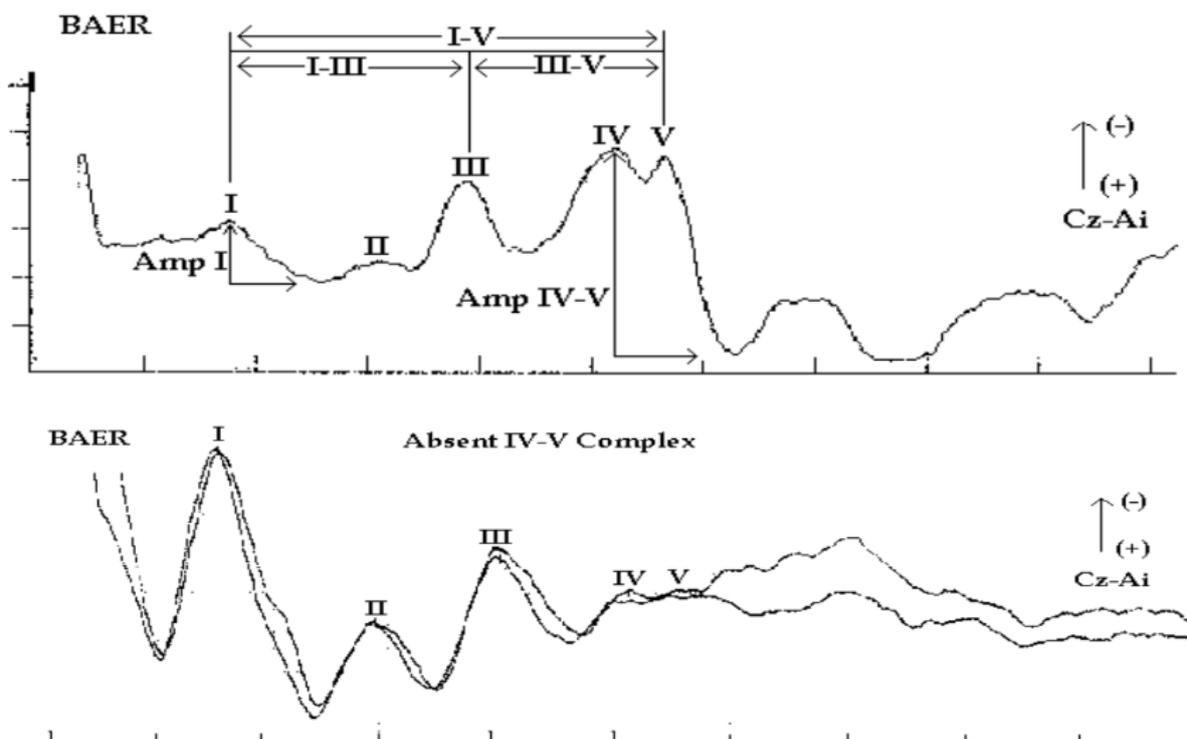


Figure 8.1: BAER shows how waves develop in response to an auditory stimulus.

1.7 EEG RECORDING TECHNIQUES

Electroencephalogram (EEG) recordings utilize electrodes with conductive media, amplifiers with filters, an A/D converter, and a recording device (Figure 9.1).



Figure 9.1: EEG signal measurement set.

Electrodes are commonly placed on the scalp with glue or in a head helmet. Since the recording captures temporal variations of an electrical system between the active (signal) and reference electrodes. A ground electrode which is needed to get a differential voltage must be used as a third electrode. They can have a single-channel configuration (one active, one reference, and one ground electrode) or a multi-channel configuration for up to 128 or 256 active electrodes. Standardization is required, as correct placement of electrodes is critical for the accurate interpretation of signals and for the comparison of results. The most frequently used system is the 10-20 system [15], which is based on 21 EEG electrodes positioned at certain anatomical reference point. The acronym '10-20' denotes proportional distances (percentage) between the ears and the nose, informing the site selection of the electrodes.

The 10-20 System

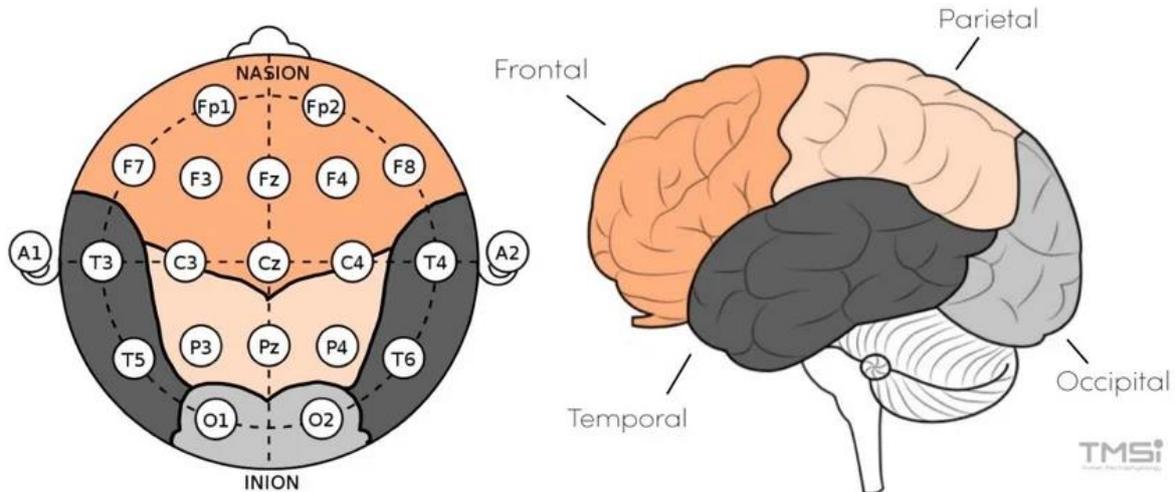


Figure 10.1: International Standard 10/20 System. This illustration displays two key anatomical reference points: the nasion, the depression between the nose and the forehead, and the inion, the prominence at the base of the skull along the midline. These landmarks measure the length of the skull along the transverse and median planes. These measurements are then segmented into 10% and 20% intervals to accurately determine the electrodes' positions.

The placement of EEG electrodes is designated according to brain regions: F (frontal), C (central), T (temporal), P (posterior), and O (occipital). Each label is recognizable by odd numbers on the left side of the head and numbers on the right, as viewed from the subject's perspective. In order to mitigate the signal distortion, the impedances applied on the scalp should be maintained below 5 KOhm and ideally within 1 KOhm.

The choice of the reference electrode is critical as it can introduce topographical distortions. Various reference points can be used, including the vertex (Cz), ears, mastoids, ipsilateral ear, contralateral ear, C7 vertebra, bipolar configurations, and the tip of the nose. Using earlobes or mastoids as references minimizes the risk of activity inflation in one hemisphere but may also result in deviations from the proper reference point. The C7 reference has been shown to be beneficial in the context of active electrodes. The reference choice is a controversial aspect, so there are various reference-free techniques, each with its benefits and drawbacks.

There are various types of electrodes available, each with distinct characteristics:

- **Single-use:** Useful to ensure sterility and minimize cross-contamination.
- **Reusable disc:** This disc is made from sterilizable and reusable materials such as gold, silver, stainless steel, or tin.
- **Headbands and electrode caps:** Facilitate multi-channel EEG recording by housing multiple electrodes in fixed positions.
- **Saline-based:** Utilize a saline solution to enhance electrode and scalp conductivity.
- **Needle electrodes:** Typically used for more precise recordings, though they may cause discomfort.

The choice of electrode depends on the specific EEG setup required. Cap systems are preferred for multi-channel recordings as they allow electrodes to be precisely and stably positioned on the scalp.

Scalp electrodes are generally made from Ag/AgCl and range in diameter from 1 to 3 mm. They are capable of recording prolonged potential variations.

Skin preparation is crucial to ensure accurate signal acquisition. It is typically recommended to clean the skin with diluted alcohol and lightly rub any dry patches to reduce impedance. Care must be taken to avoid skin abrasions from the needle when using cap systems. Repeated EEG tests require careful repositioning to prevent skin irritation or bleeding. If silver-silver chloride (Ag/AgCl) electrodes are used, a conductive paste should be applied between the electrode and the skin to minimize contact impedance and enhance signal quality.

Amplifiers and filters are integral components of the EEG measurement system, designed to ensure signals are compatible with displays, recorders, and converters. The input signal to an amplifier typically consists of five components: desired biopotentials, unwanted biopotentials, a mains interference signal, and its harmonics, interference signals generated at the tissue/electrode interface, and noise. A vital characteristic of an amplifier is its gain, defined as the ratio of the output signal to the input signal. To minimize noise, an optimal gain value ranges between 100 and 100,000. Additionally, amplifiers are designed to reduce the impact of electrically noisy environments. They are characterized by a common mode rejection ratio (CMRR) of at least 100 dB and a high input impedance of at least 100 MOhm. The CMRR measures the ratio of differential mode gain (for the desired signal) to common mode gain.

Another vital component is the A/D converter, which transforms analog signal channels into digital format by sampling at fixed intervals. In order to circumvent aliasing, it is imperative that the sampling rate exceed the maximum frequency of interest by a factor of at least two. Moreover, the resolution of the converter assumes a pivotal role, as it delineates the minimum amplitude change that can be measured with precision. That is to say, an enhanced resolution facilitates the discernment of even the most minuscule variations in the signal.

To mitigate low-frequency disturbances from sources such as respiration, a high-pass filter is appropriate. In contrast, a low-pass filter, with a cutoff frequency set to the highest frequency of interest, is useful to prevent aliasing issues.

The final step involves calibrating the entire measurement system. Calibration should ensure no discrepancies and that any output noise is primarily attributable to the analog amplifier and A/D converter circuitries.

1.8 ARTIFACTS

During a recording signal, artifacts produce unwanted noise. There are several types of artifact each categorized into physiological artifacts, generated by the human body, and external artifacts, which may arise from the surrounding environment. In order to have a clear and interpretable signal, these artifacts must be minimized [16].

Physiological artifacts originate from the eyes and eyelids, as shown in Figure 11.1. These artifacts affect the electrodes Fp1 and Fp2 (frontoparietal). Specifically, eye movements introduce artifacts in the EEG signal due to the potential difference between the retina and the cornea. Similarly, eye blinking generates high-amplitude signals that substantially exceed the amplitudes typically observed in EEG signals.

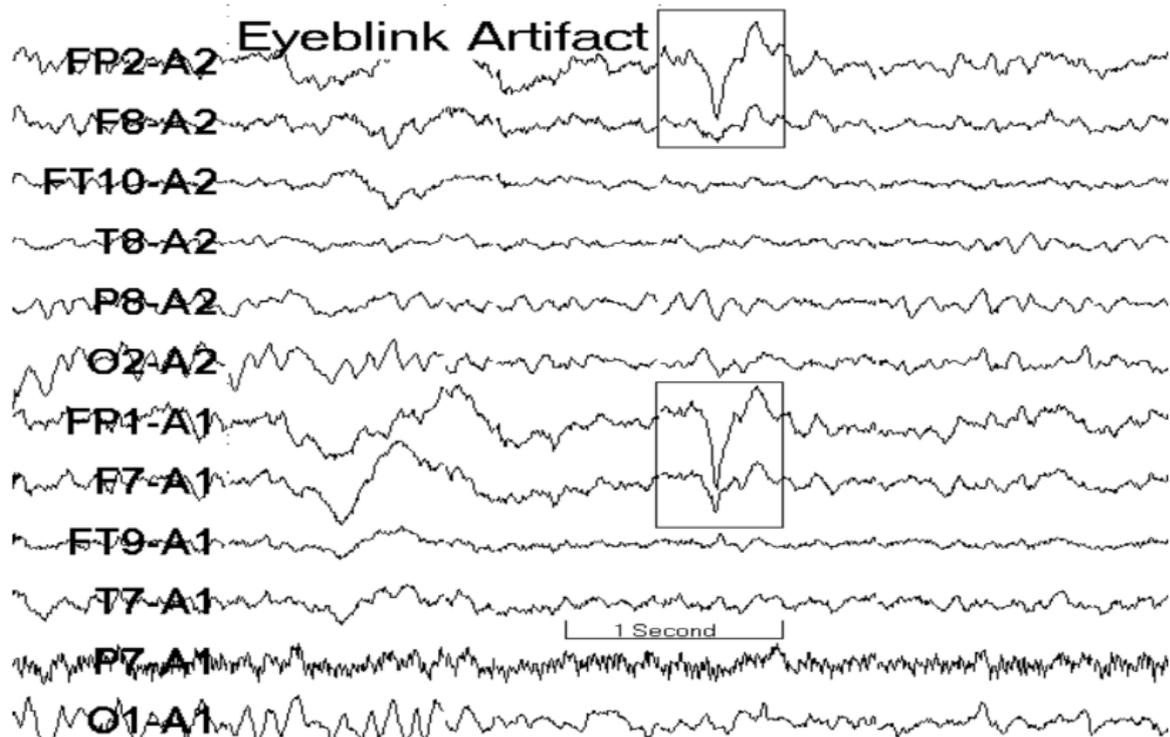


Figure 11.1: Eyeblink Artifact. This illustration highlights the points where the patient blinks, causing peaks of greater amplitude in the recorded signal.

Muscle artifacts can result from involuntary muscle movements, such as tongue movements during swallowing, grimacing, or chewing. The amplitude of these artifacts varies with the degree of muscle contraction. These artifacts are less pronounced during the beta band sleep phase when muscle activity is reduced and relaxation is more profound.

One of the most significant sources of artifacts is cardiac activity (Figure 12.1). The heartbeat generates electrical and mechanical artifacts, manifesting as an ECG signal detected by scalp electrodes. They are artifacts that appear as spike-shaped, similar to the QRS complex (ventricular depolarization) of the electrocardiographic trace. The artifact occurs at the same heart rate frequency as the subject. It can affect all electrodes but is mainly visible on the posterior channels [6].

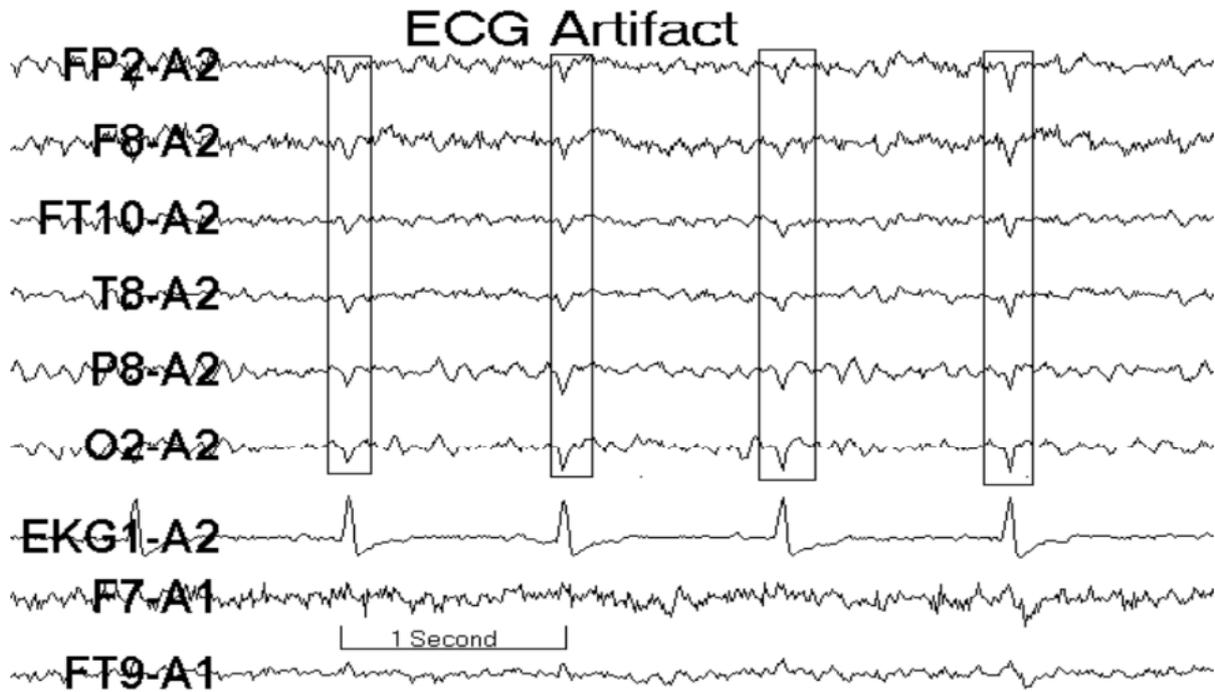


Figure 12.1: ECG Artifact. This illustration shows the appearance of a QRS. Complex, characteristic of the ECG signal, within the trace.

Electronic devices and transmission lines often induce external artifacts. Specifically, as shown in Figure 13.1, the transmission line artifact arises due to the line frequency of 50Hz, which can be easily mistaken for the frequency range within which EEG operates (0.5-60Hz). To effectively eliminate this artifact, a 50Hz notch filter can be employed.

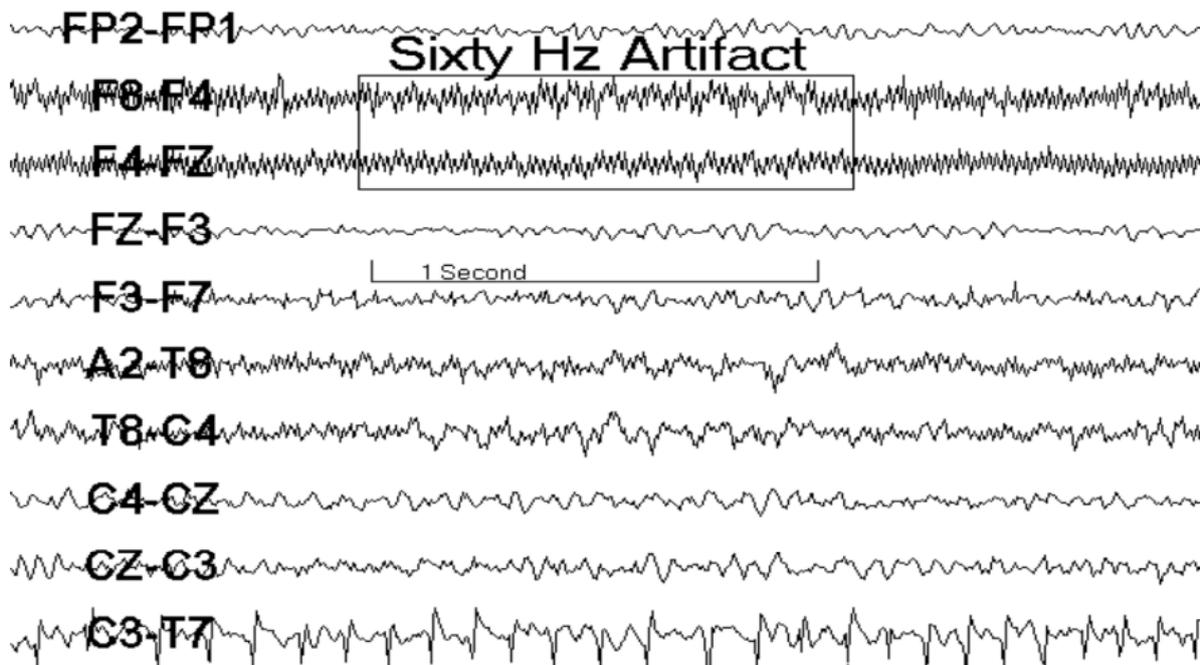


Figure 13.1: Network interference.

Telephone artifacts can occur due to high-frequency cell phone signals. To avoid such disturbances, it is advisable not to have a mobile phone present during recordings. Additionally, electrode artifacts may arise from inadequate contact between the electrode and the skin, leading to low-frequency artifacts. Ensuring proper electrode contact is crucial for minimizing these types of disturbances.

1.9 FEATURES EXTRACTION

Over the years, various methods have been implemented for extracting features from EEG signals, aimed at reducing the loss of important information contained within the signal. Let's briefly analyze these implementations below.

Fourier Transform Method (FFT): The extracted features are analyzed using Power Spectral Density (PSD), calculated via the Fourier transform on the autocorrelation sequence estimated with non-parametric methods. One method for calculating the PSD is the Welch method. In the Welch method, the signal is segmented, and a window is applied to the signal segments before the Fourier transform is performed. If the segments are non-overlapping, the variance of the mean of the periodogram is inversely proportional to the number of segments used, K .

$$S_w(\omega) = \frac{1}{k} \sum_{i=1}^k S_{w_i}(\omega)$$

In this formula, $S_w(\omega)$ represents the Welch PSD estimated by averaging the periodograms.

- **Wavelet Transform (WT) Method:** This method compresses the signal into a few parameters, which is crucial for diagnostics. It is a spectral estimation technique where any function can be expressed as an infinite series of wavelets. It is used in the time-frequency domain for extracting raw data and utilizes long-term WT windows to achieve more precise resolution at low frequencies [17].

Time-Frequency Distributions: A preprocessing phase is necessary to remove artifacts. The definition of TFD has been generalized by Choen.

$$P(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} A(\theta, \tau) \Phi(\theta, \tau) e^{-j\theta t - j\omega \tau} d\theta d\tau$$

Where $A(\theta, \tau)$ is the Ambiguity Function [17].

- **Autoregressive Method (AR):** The spectral power density (PSD) of EEG is estimated using a parametric approach. The estimation is achieved by computing the coefficients or parameters of the linear system under examination. Among autoregressive models, we find the Yule-Walker method, where the coefficients are computed by minimizing the least squares of the prediction error, and the Burg method, where error reduction is done both forward and backward without the need to compute the autocorrelation function. Parametric methods reduce spectral loss and yield increased frequency resolution [17].

In recent years, significant strides have been made in the field of brain-computer interface through the extraction of semantic features from EEG signals. This method has shown promise in

diagnosing neurodegenerative diseases such as Alzheimer's and Parkinson's, leveraging machine learning and artificial intelligence techniques. Identifying the relationship between words is a subfield of machine learning aimed at automatically extracting the meaning of words.

In the literature, there are studies where semantic features are used for diagnosing diseases that affect language capabilities. Subjects are typically exposed to external stimuli that may involve the use of various senses, followed by learning and classification methods to discern a discrepancy between perception and reality.

The goal of this thesis will be to leverage the normal characteristics of an EEG signal to search for a parameter that can express how clearly a subject is able to perceive the stimuli they are exposed to, including auditory and visual stimuli. In particular, the focus will be on a parameter known as vividness, and with the aid of machine learning models, we will search for a correlation between this parameter and the features extracted from the signal.

Before proceeding with the extraction of semantic features, there will be an initial phase of extracting the features of interest, followed by a second phase of signal preprocessing for artifact removal and noise cancellation. Healthy subjects will be compared with those affected by visual and auditory disabilities to see if there is a clear difference in terms of vividness between the two groups.

This could be key to highlighting the fact that even seemingly non-scientific parameters, when treated appropriately, can contribute to timely diagnostics.

2 SEMANTIC FEATURES

Semantic features are theoretical units useful for understanding the meaning of words, in order to understand the relationship between words in a language. Semantic concepts are the basis of the processes of understanding the world. The representation of semantics is an essential prerequisite for understanding human nature. Semantic representations perform two cognitive functions. First of all, they reflect the structure of conceptual similarity, that is, the ability to recognize that objects that are different in appearance may belong to the same category or be involved in similar actions. These similarities can be noticed from the first months of life and are used to allow the child to orient their behavior. Secondly, semantic representations allow us to recover or infer knowledge, attributing properties to objects or events that are not immediately visible or explicitly indicated. For example, seeing the image of a bird in a book, a reader may infer that the animal can fly, even if the image is still. So when we talk about semantic representations, we are referring to the cognitive and neural states that encode conceptual structure and facilitate information retrieval. [18]

Semantic memory, which represents the totality of accumulated knowledge about the world, is fundamental to almost all human activities. However, understanding its neurobiological mechanisms is only recently emerging. Recent neuroimaging studies have revealed two interesting findings: first, language comprehension involves not only general cognitive systems, but also modality-specific brain areas for sensory, motor, and emotional modalities. Second, there are large brain regions that, while active in comprehension tasks, are not limited to a single sensory modality. These regions, such as the inferior parietal lobe and large parts of the temporal lobe, lie at points of convergence between different perceptual processing pathways. Such convergences allow for the creation of more abstract and modality-independent representations that underpin conceptual functions such as object recognition, social cognition, language, and the unique human ability to recall memories and imagine future scenarios [19]. Today, the term **semantic memory** refers to all types of stored information about the defining features and attributes of concepts, as well as the processes that enable us to retrieve, use, and produce that information efficiently to support thought and language.

It is becoming increasingly clear that some significant forms of long-term memory do not fit easily into the traditional, most studied memory subtypes. A prominent example is **personal semantic memory**. Similar to episodic memory in that it is tied to an individual's identity and life, personal semantic memory differs in that, like general semantic memory, it is not tied to a specific encoding context. Retrieval of general semantic memory is known to generate an electrophysiological response called the **N400**, which has been well documented by over thirty years of research. [20]

Semantic feature extraction in the context of EEG signals is the identification process of semantic features through brain electrical activity. This field of research is interesting because it gives us information about the neural mechanisms involved in language comprehension, concept recognition, and abstract thinking. EEG studies, is crucial to analyze the encoding process of the brain thanks to its high temporal resolution.

Approaches to semantic feature extraction include the identification of specific brain response patterns, such as **event-related potentials** (ERPs), and the **analysis of brain oscillations** associated with the processing of semantic stimuli. In particular, the N400 ERP is widely studied and known to be sensitive to semantic processing, emerging as a robust response when the semantic stimulus is incongruent with the expected context.

The broad sensitivity of the N400 to semantic stimuli and meaning-related manipulations has made this component a useful tool for exploring how semantic information is stored in the brain, a topic related to semantic memory. Such research has included the study of various aspects, such as typicality, levels of representation, concreteness, and differences between word categories. In general, two significant findings emerge from such studies. First, the dissociation between reaction times (RTs) and the N400, as these two measures do not always behave in parallel. Such dissociations are common in cognitive ERP research, as ERP components reflect only a part of the processes that influence reaction times, and are a specific measure for this reason. Second, and perhaps more importantly, the N400 data do not fully support any of the existing theories in a clear-cut way, suggesting instead that different aspects of each are valid [21].

For example, N400 research aiming to understand why concrete words are processed more easily than abstract words offers an interesting insight. According to the "dual coding" theory, concrete words benefit from being represented in two semantic systems: a verbal one, predominantly in the left hemisphere (LH), which also represents abstract words, and another, based on imagery processing, located in the right hemisphere (RH). On the other hand, the "context availability theory" explains the effects of concreteness with greater semantic richness within a common modal semantic system. N400 results, however, have often been ambiguous: concrete words show a different distribution at the scalp level than abstract words, a result that seems to support the dual coding theory. However, these differences are attenuated for words that complete congruent sentences, a phenomenon that aligns more closely with the context availability theory, although not in the way exactly predicted. These data therefore suggest the need for a theory that integrates elements of both perspectives [21].

2.1 NEURONAL MECHANISMS OF SEMANTIC MEMORY

At the neurobiological level, semantic processing has been shown to involve a widespread network of brain areas, including the temporal lobes, the inferior parietal lobe, and parts of the frontal cortex. These areas are activated in a coordinated manner to support the encoding, retrieval, and integration of semantic information. Neuroimaging studies have shown the involvement of modality-specific sensory, motor, and emotional systems in language comprehension. This phenomenon is known as simulation theory, which suggests that semantic concepts are represented in a distributed manner across brain regions that directly process perceptual and motor information. For example, concepts related to movement involve motor regions, while visual concepts activate areas of the visual cortex. However, there are also brain regions that participate in semantic processing but are not tied to specific sensory modalities. These include the anterior temporal lobe and the inferior parietal lobe, which serve as hubs of multimodal integration, enabling abstract conceptual representations, which are essential for complex language comprehension and concept recognition. The goal is to identify the brain regions that influence the semantic component [22].

2.1.1 CONTEXT DEPENDENCE

Traditionally, semantic memory has been recognized as context-free, with static and immutable conceptual representations. Recent studies have questioned this staticity, suggesting dynamic realities that are strictly tied to the context in which they emerge. This new perspective finds support in the idea of a distributed architecture, which considers concepts not contained in a single brain area, but distributed among the different regions responsible for sensory, perceptive and motor functions [23].

This theory is supported by Barsalou's approach, according to which concepts are made up of two types of properties: **context-independent (CI)** properties and **context-dependent (CD)** properties. CI properties are always activated when one thinks of a concept, while CD properties emerge only when the concept is evoked in a specific context. In this way it is demonstrated that concepts are both stable and flexible [16]. To better understand Barsalou's study, let's take the word "piano" as an example. If we analyze it without a specific context, general characteristics emerge: "it has black and white keys" or "musical instrument". These are context-independent (CI) properties that are activated when we want to identify an object. If we imagine the same word placed within a context such as "moving", other properties emerge: "heavy". These are context-dependent (CD) properties, which are activated in situations relevant to the experience.

This distinction turns out to be very useful when a concept is associated with multiple pieces of information or has multiple meanings. Without the activation of contextual properties, the understanding of the concept may be less appropriate or incomplete [24].

2.2 BRAIN AREAS INVOLVED IN SEMANTIC FEATURE EXTRACTION

Brain areas commonly associated with semantic processing include:

- **Superior and middle temporal gyrus** : Involved in decoding language, these areas are activated when the brain processes words and sentences (Figure 1.2). Their activation is associated with the processing of words and sentences. Neuroimaging studies show a distinct activation during tasks requiring word recognition. The same areas are involved in face recognition and social cognition, demonstrating their versatility in processing semantic information [25].
- **Inferior parietal lobe (IPL)** : crucial for conceptual integration, it helps create abstract representations of concepts and semantic relations (Figure 2.3). It is involved in various tasks, including object recognition and understanding spatial relations between them. Its primary function is to maintain attention on the goals of the task at hand [26].
- **Inferior frontal cortex** : including Broca's area, which is not only essential for language production, but also participates in resolving semantic ambiguities and sorting relevant information (Figure 3.3).

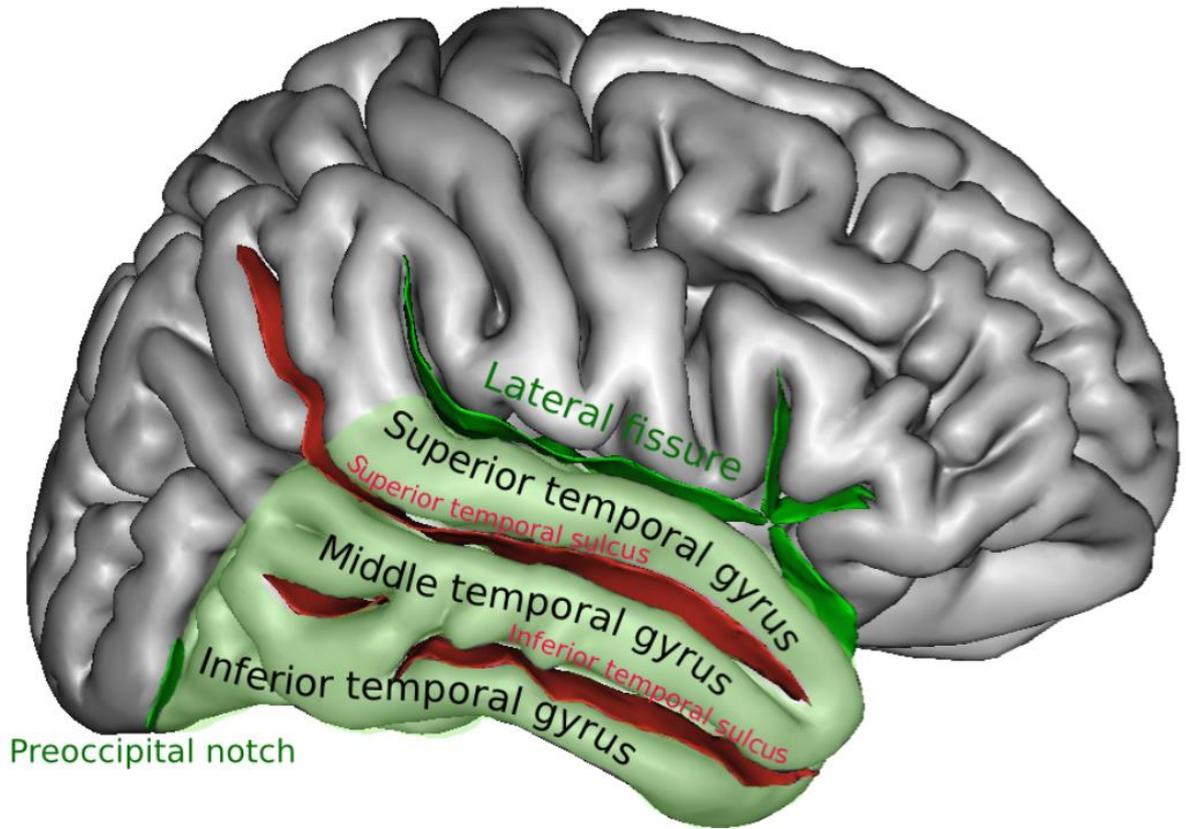


Figure 1.2: Location of the inferior temporal gyrus in the brain.

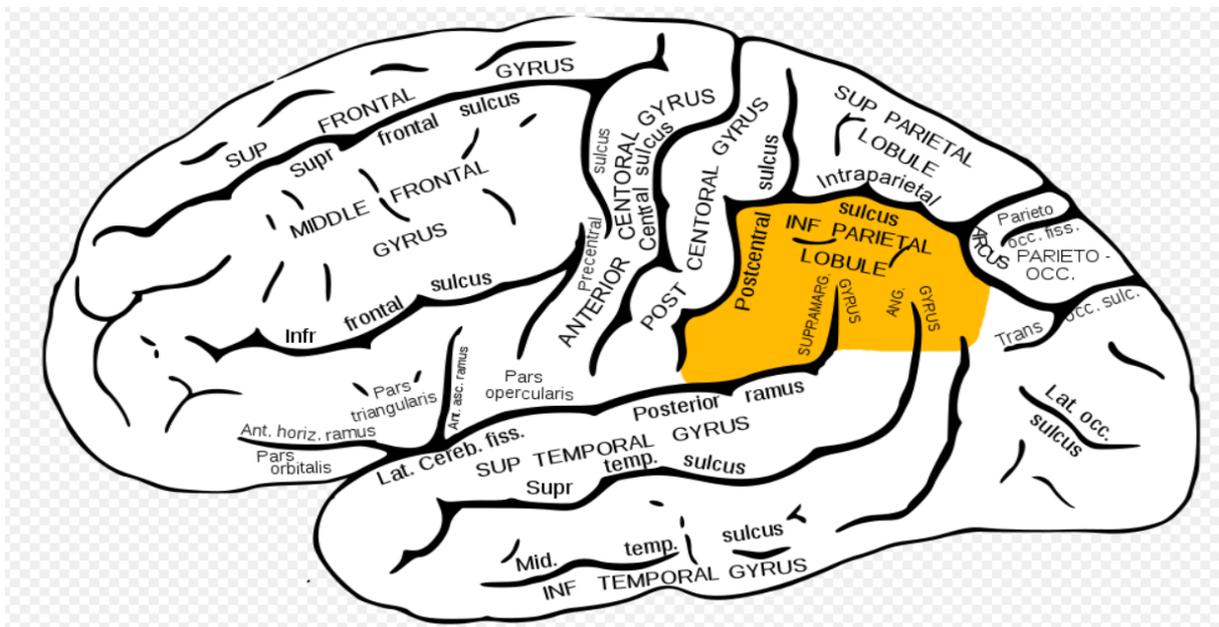


Figure 2.2: Location of the inferior parietal lobule in the brain.

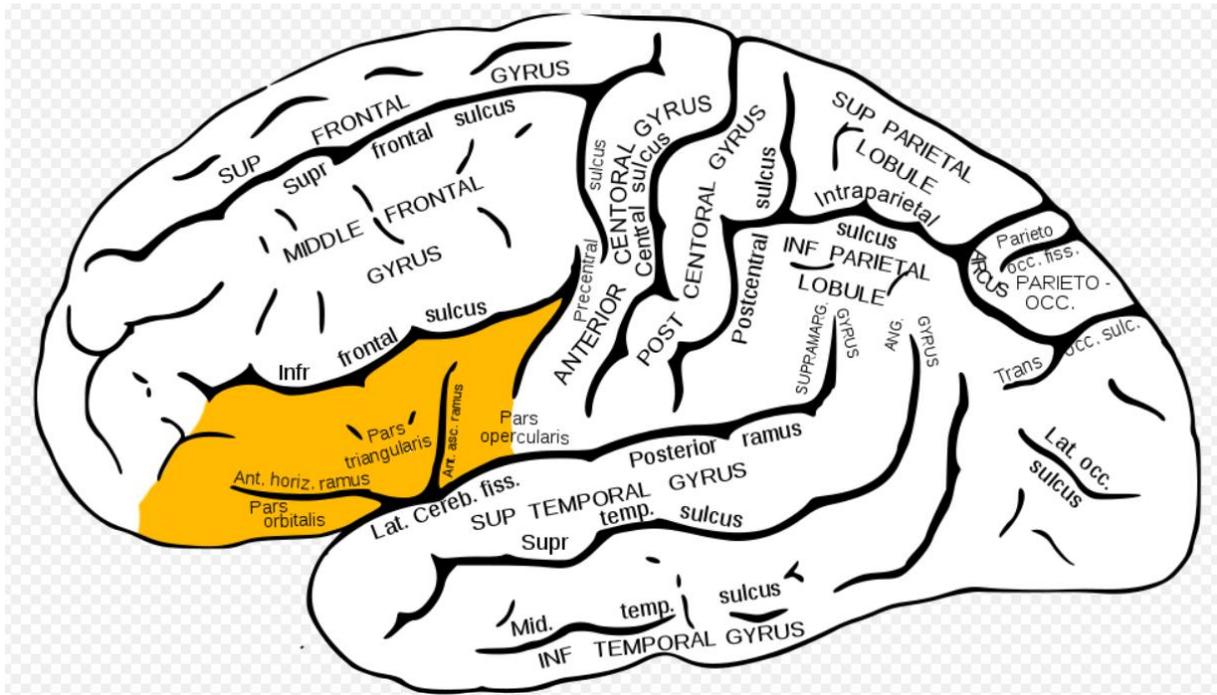


Figure 3.3: Location of the inferior frontal cortex in the brain.

These regions collaborate with each other through a distributed process, allowing the brain to efficiently manage complex concepts and semantic categories [21].

Clarification is now needed regarding the activation of areas involved in abstract and concrete concepts. These regions are the posterior part of the middle temporal gyrus in the left hemisphere and the inferior occipital gyrus in the right hemisphere.

The middle temporal gyrus in the left hemisphere is part of a network that sorts semantic information into categories. This region has been shown to be statistically more active for abstract concepts than for concrete ones.

The inferior occipital gyrus of the right hemisphere has a greater activation for concrete concepts, given their greater imageability. These results strengthen the hypothesis of having two distinct networks for the processing of abstract and discrete concepts [27].

2.3 EVENT RELATED POTENTIALS (ERPs) IN SEMANTIC PROCESSING

Event-related potentials (ERPs) measured through electrodes placed on the scalp represent the electrical activity of the brain, which mainly reflects the postsynaptic potentials of simultaneously activated pyramidal neurons in the cortex. This brain activity is activated by specific events, such as exposure to words or images. To obtain detailed information about sensory, perceptual, cognitive and motor processes, it is necessary to analyze the changes in magnitude, timing and distribution of ERP activity. In this way, information is obtained with temporal precision of the order of milliseconds, involving various brain districts.

ERP events are short-lived, about 500 milliseconds, and several trials must be conducted to obtain a clear signal. This is because the ERP waveform is calculated by averaging all trials of the same stimulus, allowing to eliminate the background noise of the EEG signal and to highlight the specific response of the brain. A great advantage of ERPs is the ability to link different components of the

waveform, both positive and negative, with variable latency with respect to the onset of the stimulus, to specific cognitive processes. However, a disadvantage is that to obtain reliable data a relatively large number of trials must be performed [28].

ERPs are particularly powerful in studying language comprehension because a negative component, known as the N400, occurs approximately 400 milliseconds after the stimulus onset. This component has been associated with systemic changes in semantic information processing, highlighting how the brain reacts in real time to linguistic stimuli.

N400 amplitude shows remarkable sensitivity to the immediate context in which the word appears. Amplitude changes based on semantic relationships between words in a list. The presence of a related word tends to reduce N400 amplitude. When both sentence context and word association information are available, they affect N400 amplitude partially independently, suggesting context conditioning.

language comprehension, the N400 amplitude also appears to reflect the ease with which information can be retrieved from long-term memory. For example, N400 amplitudes for words presented without specific context can vary based on non-semantic factors, such as the frequency of use of the words themselves. These factors indicate how easily information related to these perceptual forms can be extracted from lexical memory. Furthermore, the amplitude decreases following repetitions, with variations depending on both the number of times the word was repeated and the time between repetitions. Factors that influence recognition memory also impact this amplitude. The link between the N400 and access to long-term memory is further supported by studies with intercranial recordings, which demonstrate that part of the N400 source, measured on the scalp, comes from specific brain areas such as the medial temporal lobe, which is widely recognized for long-term memory processes [27].

Another interesting aspect regarding the N400 is the dissociation that can exist between measures of reaction times (RT) and N400 responses. These measures do not always behave similarly, for example a high reaction time does not necessarily imply a large amplitude N400 response. This dissociation is common in other cognitive ERP measures, suggesting that ERP components capture only a portion of the processes that contribute to behavioral reactions. Thus, using the N400 response provides insights not only into how the brain manages meaningful information, but also raises questions about how semantic memory is organized and retrieved [29].

In addition to the traditional N400-based approach, recent studies have introduced machine learning and deep learning techniques to identify complex semantic patterns directly from brain signals. Artificial neural networks have been employed to analyze EEG activity during reading or language comprehension tasks, allowing to identify and classify conceptual representation stimuli, demonstrating that the human brain encodes semantic relations in a coherent and identifiable way [30].

An innovative approach in this field involves the use of natural techniques language processing (NLP), such as word embedding models (Word2Vec, GloVe), to create semantic representations of words or sentences. These representations can then be correlated with EEG data, resulting in a mapping of the semantic space that allows prediction of brain activity associated with semantically similar words.

2.3.1 WORD2VEC

Word2Vec has attracted attention for its flexibility and ability to semantically represent words in a vector space. Word2Vec can be considered both supervised and unsupervised. It is supervised because the model learns from data using architectures such as continuous bag-of-words or the continuous skip-gram, but it is also unsupervised because it can be applied to any large corpus without specific labels. The limitation of Word2Vec is that it treats all words in the same way, which makes it difficult to identify which words are most relevant to a given document. To solve this limitation, a project has been developed that allows to add and subtract vectors of similar words using the Word2Vec algorithm, in order to avoid ambiguous terms and create a search system that allows to add words as vectors. Another interesting approach is to calculate weighted sums of word vectors, to obtain more accurate representations of documents. In this case, the skip-gram model is chosen, which despite having longer processing times, offers greater semantic accuracy than other architectures.

Another key aspect concerns the importance of categories. For example, topics such as sports and agriculture tend to generate higher scores because their contents do not overlap semantically, unlike categories such as agriculture and nutrition. The set used to train the models can significantly influence the results [31].

2.3.2 GLOVE

GloVe is an unsupervised learning algorithm designed to create vector representations for words based on their occurrence. Unlike the previously mentioned Word2Vec, it combines global and local context information, exploiting the idea that words with similar meanings tend to appear together frequently.

GloVe constructs a co-occurrence matrix. Within this matrix each entry represents the frequency of word pairs occurring at the same time. This matrix is factorized via singular value decomposition to obtain word vectors. These vectors are useful for encoding both semantic and syntactic information. The GloVe model has several advantages including efficiency, making it faster to train than Word2Vec, interpretability, aiding in the interpretation of semantic relations, and scalability, as it scales well to large corpora, making it suitable for big data processing [32].

2.4 OSCILLATORY ANALYSIS

The analysis of brain oscillations offers another window to understand semantic processing. Gamma (30-100Hz), alpha (8-12Hz) and theta (4-8Hz) oscillations have been identified as particularly relevant in the context of semantic processing. Gamma oscillations are associated with neuronal synchronization and information integration, alpha oscillations are often related to attention modulation, while theta oscillations are useful for semantic memory retrieval and connectivity between cortical areas [33].

2.4.1 GAMMA BAND OSCILLATION IN SEMANTIC FEATURE EXTRACTION

Gamma wave activity can be associated with specific stimulus processes, such as the perception of visual or acoustic stimuli, and also with more complex cognitive processes such as facial recognition or decision making. Some researchers have hypothesized that this activity is based on the activation of cell assemblies, that is, sets of neural cells that are activated simultaneously during the processing of concepts. According to this study, each linguistic process has its own cell assembly, which is activated when the stimulus is processed, so it is expected that the high-frequency responses of the gamma band are linked to linguistic processes [34]. To analyze this neuronal response, experiments were performed for the recognition of real words and pseudowords, that is, sequences of letters without meaning. A reduction in power was observed in the alpha (7.5-12Hz) and beta (12.5-17.5Hz) bands for both types of stimuli, while the gamma band shows a different behavior for the two types of stimuli. In the case of real words, no significant changes are observed, while in the case of pseudowords, a reduction in power around 30Hz is observed. The explanation for this type of response is given by the fact that pseudowords are not represented cortically as defined linguistic concepts, and therefore they are unable to activate a specific cell assembly [35]. This study, performed in the first half of the 90s, was re-performed with an MEG study that confirmed the reduction of gamma activity for pseudowords. These results make it clear how gamma band activity is specific for word processing, as they can be used as indicators of cell assembly activity, closely linked to semantic processes and linguistic representation [36].

Although gamma activity appears to be of considerable importance for linguistic processes, some notes must be made:

1. Differential gamma responses are observed only in studies where single words are processed. These studies suggest two types of gamma band activation, an early “ignition” and a late “reverberation”. These activation stages could be associated with different semantic networks, as in the case of early activation, or with active memory as in the case of late activation [37].
2. Gamma responses are assumed to be a sign of activation of neuronal assemblies when processing single words, so the problem remains of understanding how these responses influence the construction of syntactic structures. In other words, the challenge still open is to connect the results based on oscillations with those obtained by techniques such as event-related potentials (ERPs) and functional magnetic resonance imaging (fMRI).
3. The larger amplitude gamma oscillations might occur because meaningful words attract attention, causing greater neural activation, which would explain the poor arousal in the presence of pseudowords [38].

In conclusion, it can be stated that semantic feature extraction in the context of EEG represents a rapidly growing field of research. The use of the aforementioned techniques is allowing for more precise mapping of the neural processes involved in the comprehension and retrieval of semantic information.

3 METHODOLOGY

This study aims to correlate the results of EEG classification with VVIQ (Vividness of Visual Imagery Questionnaire) and BAIS-V (Bucknell Auditory Imagery Scale - Vividness) scores. This significant correlation provides insights into how cognitive abilities related to self-reported imagery scores reflect on neurophysiological measures. Understanding these relationships can contribute to developing personalized classification models and improve our comprehension of the cognitive processes underlying imagination and perception.

Analyzing EEG signals may reveal connections between brain activity patterns and self-reported imagery scores, shedding light on the neural basis of imagination and perception. This knowledge can have applications in cognitive neuroscience and BCI, in which medical interest is growing significantly.

3.1 DATASET DESCRIPTION

This research draws on the EEG-based BCI dataset of semantic concepts for imagination and perception tasks, as outlined in [39]. The dataset under consideration is deemed suitable for the analysis in question, as it provides pre-processed electroencephalogram (EEG) recordings combined with self-reported VVIQ and BAIS-V scores, thus allowing for a comprehensive investigation of the relationship between EEG features and image scores.

The following information is to be provided by the dataset:

- **Participants:** The dataset includes recordings from 15 participants.
- **EEG channels:** Data were collected using a 124-channel EEG system.
- **Conditions:** Participants were tasked with performing imagination and perception tasks, which were divided into three stimulus modes: auditory, orthographic and pictorial. Each participant was instructed to view images of animals (penguins, flowers and guitars), as well as images of the words “penguin”, “flower” and “guitar”. They were also instructed to listen to the sounds of the words “penguin”, “flower” and “guitar”.
- **Stimuli:** 22,248 images relating to 1,854 semantic concepts were used. In this manner, three categories of stimuli were identified: orthographic, pictorial and sound.

Participants were asked to report the vividness of their imagery using the VVIQ and BAIS-V questionnaires prior to the EEG recording sessions. These self-reports form part of the investigation into potential correlations with EEG features.

3.1.1 VVIQ

This is a questionnaire created in 1973 by the British psychologist David Marks that allows us to explore the vividness of the visual imagination of the patient to whom it is submitted [40]. It is an accurate test of vividness that tries to form mental images of people, objects, and settings and evaluates the vividness of the imagination with a score from 1 to 5. Without visual imagination, vividness will be assessed as 1, while the maximum score of 5 is attributed only to vivid images, such as real sight. Table 1.3 shows the relationship between the scores and the type of imagination.

Scores	1	2	3	4	5
--------	---	---	---	---	---

Imagination	No image, the subject is simply thinking about the object	Dim and vague image	Moderately realistic and vivid	Realistic and vivid reasonably	As vivid as the real view

Table 1.3: The table shows how scores are assigned based on the type of image the subject can construct.

Vividness is a construct related to the level of detail and clarity and the intensity with which sensory stimuli are perceived compared to real vision. It depends on various factors that can increase or decrease it.

The questionnaire can be exemplified as follows: The subject is asked to think of a friend or relative he sees frequently and to evaluate the following requests.

1. The exact contours of the face, head, shoulders and body.
 - No image, you “know” you are thinking about the person in question
 - Dark and vague
 - Moderately clear and lively
 - Clear and lively
 - Perfectly clear and vivid

2. Characteristic head poses, body attitudes.
 - No image, you “know” you are thinking about the person in question
 - Dark and vague
 - Moderately clear and lively
 - Clear and lively
 - Perfectly clear and vivid

3. The bearing, the length of the step when walking
 - No image, you “know” you are thinking about the person in question
 - Dark and vague
 - Moderately clear and lively
 - Clear and lively
 - Perfectly clear and vivid

4. The colors worn in some family clothes
 - No image, you “know” you are thinking about the person in question
 - Dark and vague
 - Moderately clear and lively
 - Clear and lively

- Perfectly clear and vivid

This test is easy to administer even to elderly people, but it is essential to remember that there is no reliable, objective measure that can express how vividly a person can imagine. Visual imagery is a self-report measure, so it is possible that subjects affected by neurovegetative cognitive diseases may not be reliable in expressing how vivid their visual imagery is [41].

3.1.2 BAIS

The BAIS is a measurement scale developed to improve the VVIQ. This is still a self-assessment of vividness, but it focuses on potential differences between the auditory and visual domains using subscales. Participants are described in a situation accompanied by a relevant sound. The BAIS-V is a subscale of the BAIS in which participants must rate the vividness of their images of the sound on a scale of 1 to 7. As in the previous case, 1 indicates a lack of imagination, while the highest score is given when the image is as vivid as the actual sound. The other subscale used is the BAIS-C, in which participants have again described the situation, but the task is to rate how easily they can change the image of the original sound into a new sound. Again, a scale of 1 to 7 is used, where 1 indicates no imagination, 4 indicates that the image could be changed but with difficulty, and 7 is a sign of extreme ease in changing the image. An example of a questionnaire is presented below: in the BAIS-V, participants are asked to imagine a specific song indicated by the operators and to rate the vividness of an instrument that plays at the beginning of this song, such as a trumpet. In the BAIS-C, they are asked to rate how the sound changes from a trumpet to a violin.

The BAIS scale appears to have good reliability, and although people tend to be skeptical about self-rating scales, it shows that participants can assess their own auditory imagery [42].

3.2 DATA PROCESSING AND FEATURE EXTRACTION

- 1. Preprocessing:** The study utilised a dataset comprising both raw and pre-processed subsets. For the purposes of this study, the latter were employed. The implementation of independent component analysis (ICA) and filtering to eliminate power-line noise served to streamline the preprocessing pipeline and ensure data quality. Furthermore, additional preprocessing steps were incorporated beyond those applied by the original authors to align the data with the specific approach and methodology employed in this study. Events were re-labeled and segmented into distinct conditions (perception vs. imagination). To aim this purpose, event markers were extracted to identify perception and imagination tasks, and then the data was epoched to isolate trials of interest using a 4-second window by preserving the signal dynamics.

A subset of epochs was selected for the purpose of computing rejection thresholding, with the objective of excluding noisy trials. A detailed mapping of events was performed in order to categorise stimuli based on type (auditory, pictorial, orthographic) and task (perception or imagination).

- 2. Feature Extraction:** Relevant EEG features were extracted, including:
 - **Power Spectral Density (PSD):** The PSD was computed using the Welch method to estimate the power distribution across frequencies. The epochs were loaded and concatenated, and the

PSD was plotted, focusing on a range from 0.5 to 100 Hz. The frequency bands of interest were defined, and the average power for each band was calculated. In this way, plotting a PSD graphic in dB was possible, emphasizing spatial distributions of power.

- **Frequency Band Analysis:** EEG signals were divided into cortical regions (prefrontal, temporal, parietal, occipital), and band power was computed for each region. Available channels were identified for each cortical region. The band power was calculated using a custom function to integrate the PSD over specified frequency ranges. Then, the results were aggregated across participants to derive the mean power for each band and cortical region. In this way, it was possible to identify how the frequency bands are activated within different regions.

3.3 CLASSIFICATION

Since the dataset contains only 15 subjects, it is relatively small for deep learning (DL) models, which inherently require larger datasets for effective training. Additionally, to preserve the integrity of the raw data, we chose not to apply any data augmentation techniques to increase the dataset size. As a result, DL models were not utilized in this study. Four traditional machine learning (ML) models were trained to analyze the relationship between EEG features and self-reported imagery scores:

1. Support Vector Machine (SVM) is used for its robustness in binary classification.
2. Random Forest (RF) chosen for its capability to handle complex datasets with non-linear relationship.
3. Logistic Regression (LR) used to analyze binary outcomes and reduce bias between groups.
4. Multilayer Perceptron (MLP) applied for its ability to capture non-linear pattern.

The operation on these classifiers is described below.

3.3.1 SUPPORT VECTOR MACHINE

SVM is a supervised learning model associated with classification learning algorithms. Given a set of examples labeled with the class they belong to between two possible classes, this algorithm builds a model that assigns the new examples to one of the two classes, obtaining a binary classifier. The SVM ensures that the examples belonging to the two categories are separated by as much space as possible.

It uses the linear kernel to define a line that evenly divides the two classes. There will be new unlabeled elements classified as:

- Belonging to the blue class, if they are above the red line
- Belonging to the green class if they are below the red line

Different kernels can be used, including polynomial, exponential or even a custom kernel can be created, to exploit mathematically more complex functions than a straight line that allow the instances of a dataset to be divided. In fact, a linear kernel does not always allow a correct division of the instances of a dataset.

3.3.2 RANDOM FOREST

RF is a supervised learning algorithm that can be used for both classification and regression. It is part of the ensemble methods, that is, models that include simpler models within them. Decision trees are independent, they do not influence each other with the advantage of avoiding some errors, such as instability deriving from their way of making predictions. The independence of the trees is strengthened by the peculiarity of always using random samples as variables for internal training, which makes them suitable for different specializations. It has a set of hyperparameters that can be used to improve the performance of the model.

3.3.3 LOGISTIC REGRESSION

LR classifier aims to analyze binary outcomes. It uses continuous or categorical predictors. This is useful when it is necessary to reduce the bias between two compared groups [43].

The most used model uses a multiple or multivariable LR and this is used when there are multiple predictors. The procedure is like multiple linear regression, but the response is binomial. The most relevant advantage is that it is possible to use more than one variable simultaneously. The struggle is to understand what variables are significant, so two situations increase. The first one is that there aren't study that assay that there is a link between significant variables and the event of interest; the second one is that when a classifier is characterized by more than one variable, it has less statistical power. So, this kind of classifier is a powerful tool, allowing multiple variables, but it is necessary to pay attention to avoid feeding software with raw data [44].

3.3.4 MULTILAYER PERCEPTRON

MLP is a powerful classifier made up of a certain number of neurons. To activate this architecture, threshold functions can be used, which however make the network more limiting; to create a multilayer model perceptron we choose to activate the algorithm with the logistic function, so that we can use the rules of differential calculus in the learning algorithm. To teach the algorithm the relationship between input and output, we give it a set of input and target vectors. This way when we insert an input after training, the output will be very close to the target [45]. Table 2.3 shows the hyperparameters for all models.

Models	Hyperparameters
Random Forest	n_estimators=100 random_state=42
Logistic Regression	solver="liblinear"
Support Vector Machine	probability=True kernel="rbf" random_state=42
Multilayer Perceptron	hidden_layer_sizes=(100,) max_iter=500 random_state=42

Table 2.3: hyperparameters for all classification methods.

An explanation about this hyperparameters is viewable below:

- n_estimators=100: indicates the number of trees.
- random_state=42: set a specific value for the random number generation. In this way the outcomes are deterministic and they can be replicated in the exact way of the same script.
- solver="liblinear": is an optimization used for Logistic Regression.
- probability=True: calculates probabilities based on the model.
- kernel="rbf": explains that the kind of kernel is a radial basis function.
- hidden_layer_sizes=(100,): is referred to the number of neurons in hidden layers.
- max_iter=500: represents the maximum number of iterations for the optimization model.

The evaluation of the models was conducted through the calculation of accuracy and the Area Under the Curve (AUC), with the subsequent plotting of the Receiver Operating Characteristic (ROC) curve. In addition to this, a series of statistical analyses were conducted with the aim of assessing the correlation between the results obtained and the self-reported VVIQ and BAIS-V scores. This approach enabled the evaluation of not only the accuracy of the models, but also the sensitivity and specificity of each classifier, thereby providing insights into their predictive capability. To mitigate the risk of overfitting, a stratified K-fold cross-validation with five folds was employed to ensure the robustness of the model evaluation.

3.4 STATISTICAL ANALYSIS

The following statistical analyses were conducted:

1. **Correlation Analysis:** Pearson and Spearman correlations were computed to examine the relationship between classification metrics (accuracy, AUC) and self-reported scores (VVIQ, BAIS-V). A Pearson correlation is performed by identifying linear relationships between the accuracy and AUC with the self-reported values for VVIQ and BAIS-V. This test provides a correlation coefficient, therefore how much the variables are correlated and a p-value, or how significant this relationship is. Through heatmap graphs, the relationship between the self-reported values and the results is visually shown. To better interpret how much a variable influences the dependent variable (Accuracy or AUC), a regression plane calculation is implemented based on a grid of values. From this 3D graph, the slope of the plane is observed.

A plane strongly inclined towards one axis rather than another suggests which self-reported value is a strong predictor of one or the other dependent variable.

2. **Group Comparisons:** Participants were divided into high and low VVIQ/BAIS-V groups to assess differences in classification performance.
3. **Regression Analysis:** Regression planes were calculated to visualize how self-reported scores influenced classification metrics. Heatmaps and 3D graphs were used to illustrate these relationships.

All these steps that make up the methodology of this study ensures a comprehensive investigation of the relationship between EEG features and cognitive abilities.

4 RESULTS

This section presents the results obtained from the four ML models and the corresponding statistical analysis.

The results of the Accuracy (Acc), AUC, ROC curves, and statistical analysis are presented below, categorized by each modality.

4.1 AUDIO MODALITY

Table 3.4 shows the accuracy and AUC results for each patient and each classification method for what concern the audio modality.

Subject	RF		LR		SVM		MLP	
	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
3-3	53	0.56	49	0.51	54	0.55	60	0.55
8-3	58	0.60	59	0.63	60	0.62	56	0.58
10-1	54	0.59	54	0.57	55	0.56	53	0.52
11-1	52	0.52	56	0.58	50	0.59	51	0.52
12-1	56	0.60	60	0.63	60	0.62	62	0.64
12-2	47	0.48	53	0.55	51	0.52	52	0.51
13-1	65	0.69	63	0.69	64	0.67	62	0.62
14-1	56	0.60	59	0.63	57	0.59	58	0.59
14-2	48	0.50	57	0.60	49	0.51	52	0.54
15-1	51	0.55	57	0.58	53	0.59	55	0.57
15-2	63	0.69	63	0.64	61	0.66	58	0.59
16-1	60	0.64	62	0.66	64	0.68	58	0.59
17-1	75	0.83	73	0.84	75	0.83	65	0.70
18-1	49	0.50	49	0.54	53	0.55	50	0.52
19-1	77	0.86	70	0.76	73	0.85	64	0.66
Average	57.6	0.61	59.0	0.63	58.6	0.63	57.1	0.58

Table 3.4: Accuracy and AUC for audio modality.

ROC curves for audio modality are shown in Figure 1.4. The curves are plotted for each subject and classification method. The various classifiers are distinguished by different colors: blue for RF, orange for LR, green for SVM, and red for MLP.

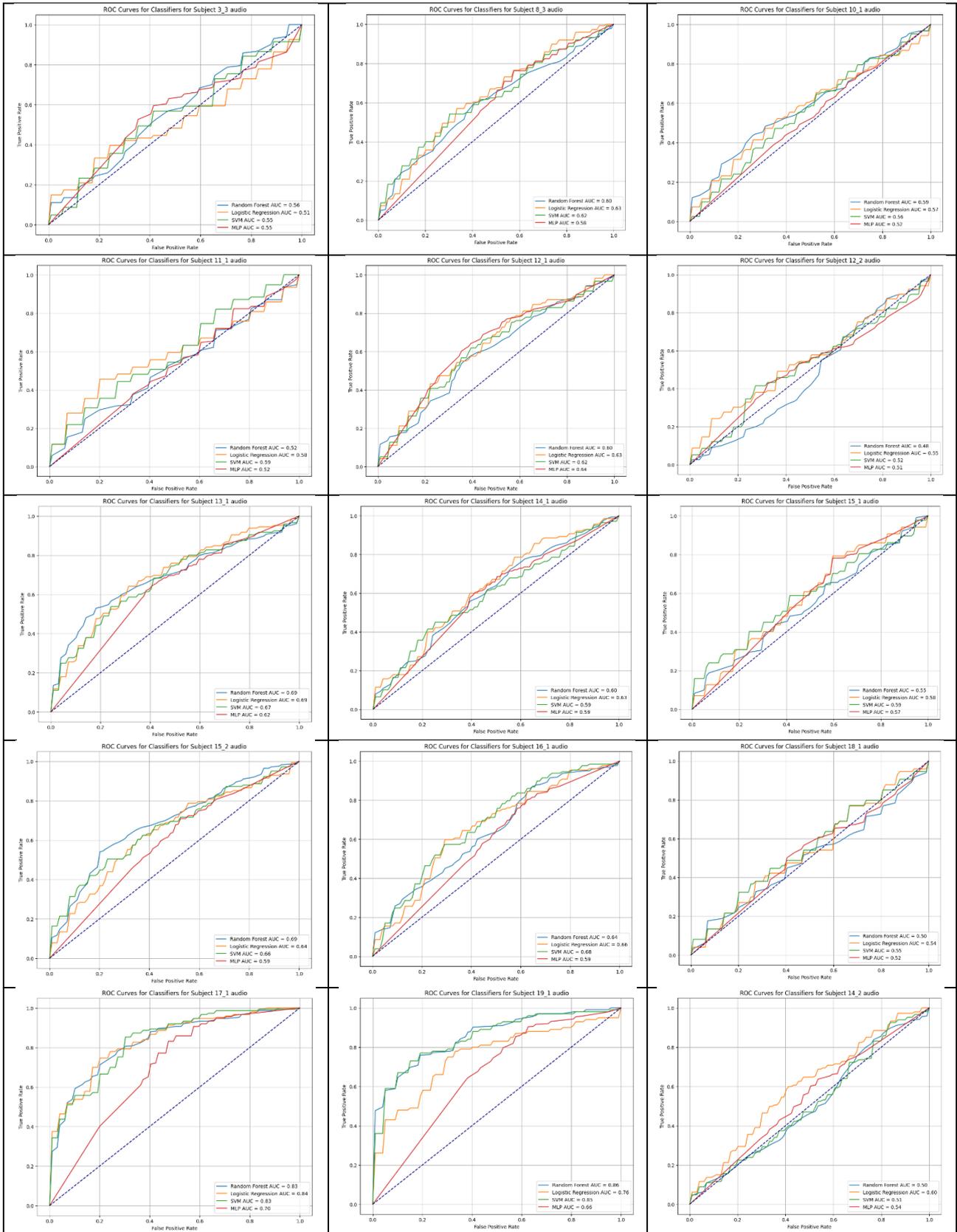


Figure 1.4: ROC curves for audio modality for each subject and classifier.

Next, the regression planes for each individual classifier, the results of the Pearson correlations and the comparison between combined groups will be shown.

Figure 2.4 shows the regression plane for the accuracy and AUC on the audio modality for the RF classifier.

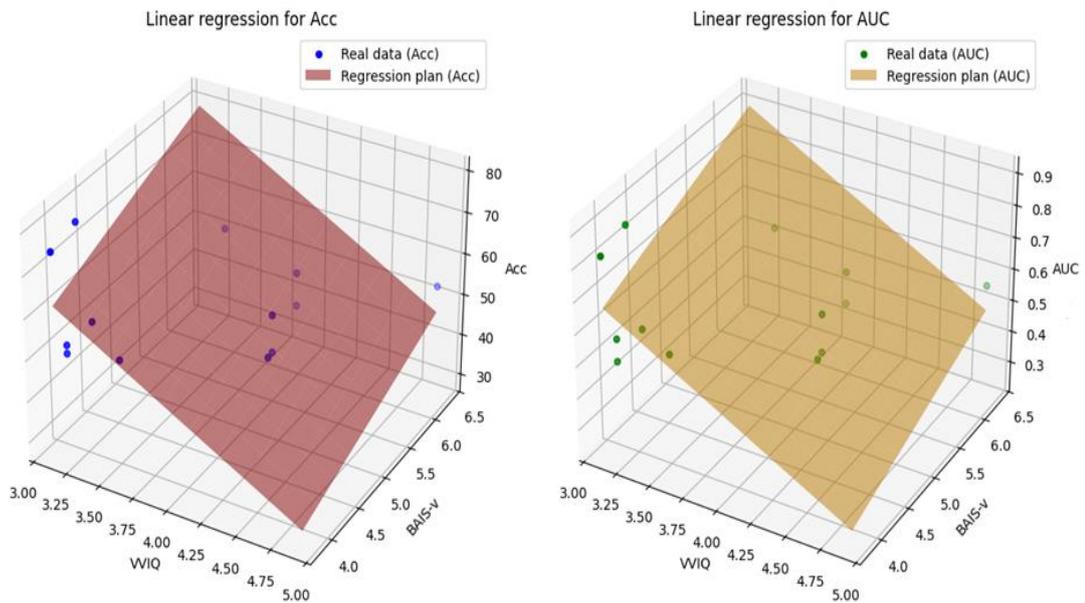


Figure 2.4: Regression planes on the RF classifier for audio modality on accuracy and AUC.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-18.31	6.85
AUC	-0.23	0.09

Table 4.4: Regression plane inclination for RF classifier and audio modality.

Table 5.4 shows the Pearson correlation.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.5653	0.0441	-0.3039	0.3128
AUC	-0.5495	0.0517	.0.2801	0.3541

Table 5.4: Pearson correlation with VVIQ and BAIS-V for the RF classifier.

Figure 3.4 shows the comparison results between combined groups for Acc and AUC.

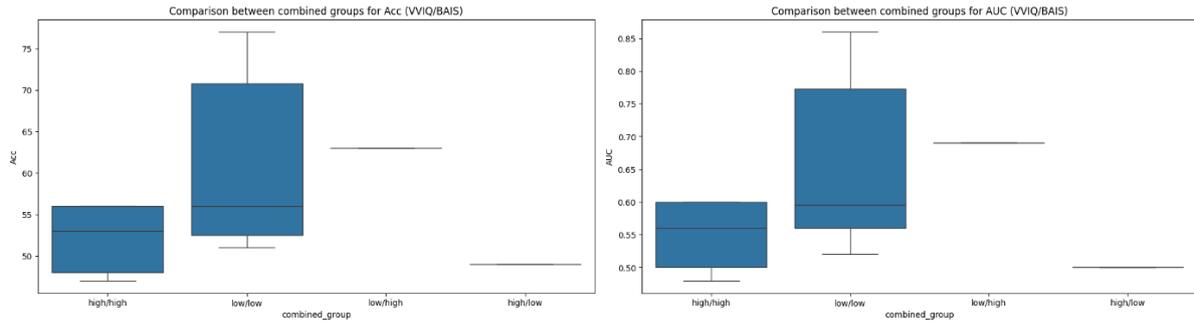


Figure 3.4: Comparison between combined groups for Acc and AUC for RF classifier. T-test Acc (VVIQ high vs low): $t = -2.1258$, $p = 0.0570$). T-test AUC (BAIS high vs low): $t = -0.9550$, $p = 0.3601$.

Figure 4.4 shows the regression plane for the accuracy and AUC for the audio modality for the LR classifier.

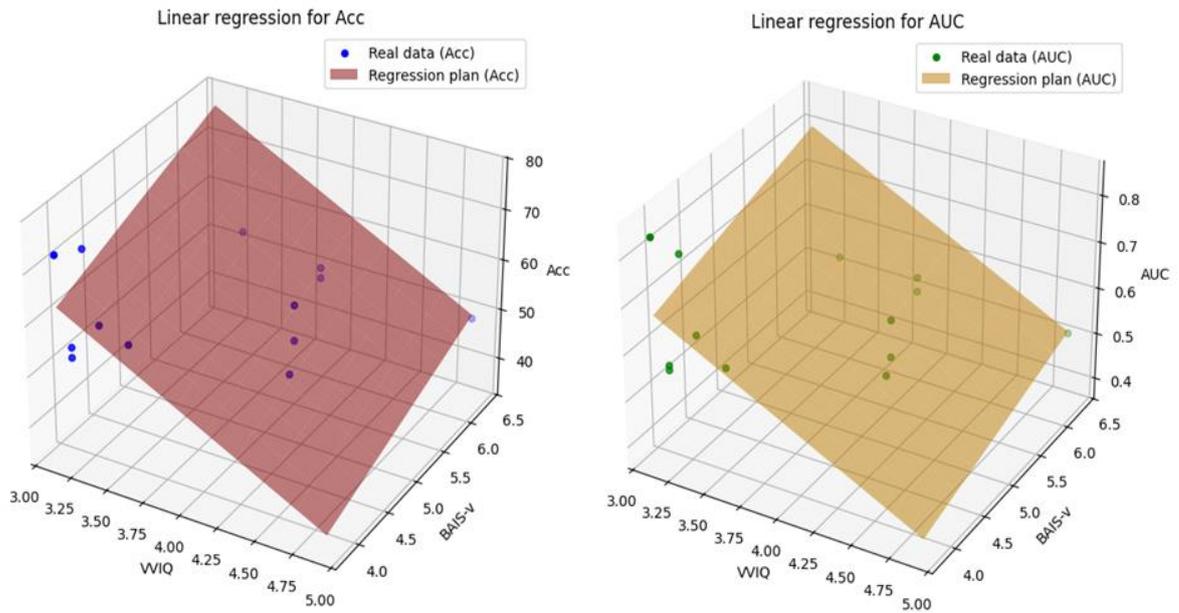


Figure 4.4: Regression planes on the LR classifier for audio modality on accuracy and AUC.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-15.21	5.61
AUC	-0.16	0.05

Table 6.4: Regression plane inclination for LR classifier and audio modality.

Table 7.4 shows the Pearson correlation.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.6404	0.018	-0.3504	0.2404
AUC	-0.6078	0.0276	-0.3804	0.1997

Table 7.4: Pearson correlation with VVIQ and BAIS-V for the LR classifier.

Figure 5.4 shows the comparison results between combined groups for Acc and AUC (audio modality and LR classifier).

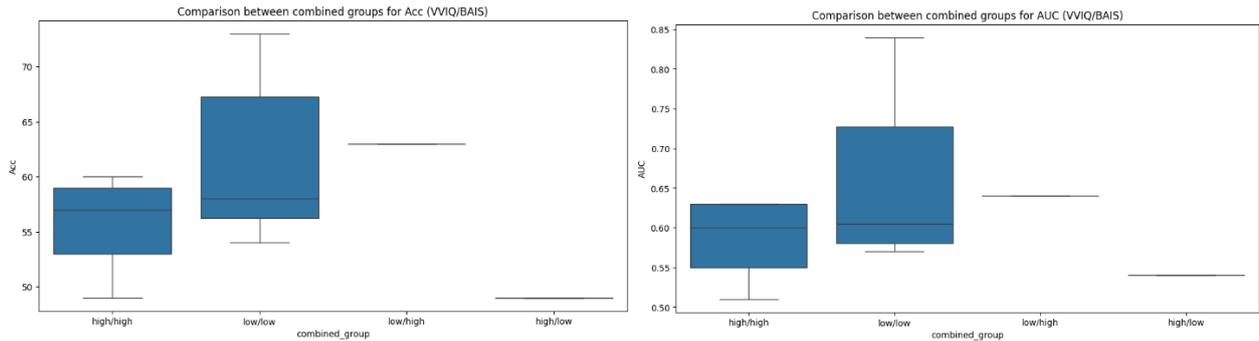


Figure 5.4: Comparison between combined groups for Acc and AUC for LR classifier. T-test Acc (VVIQ high vs low): $t = -2.0528$, $p = 0.0647$. T-test AUC (BAIS high vs low): $t = -0.9826$, $p = 0.3469$.

Figure 6.4 shows the regression plane for the accuracy and AUC for the audio modality for the SVM classifier.

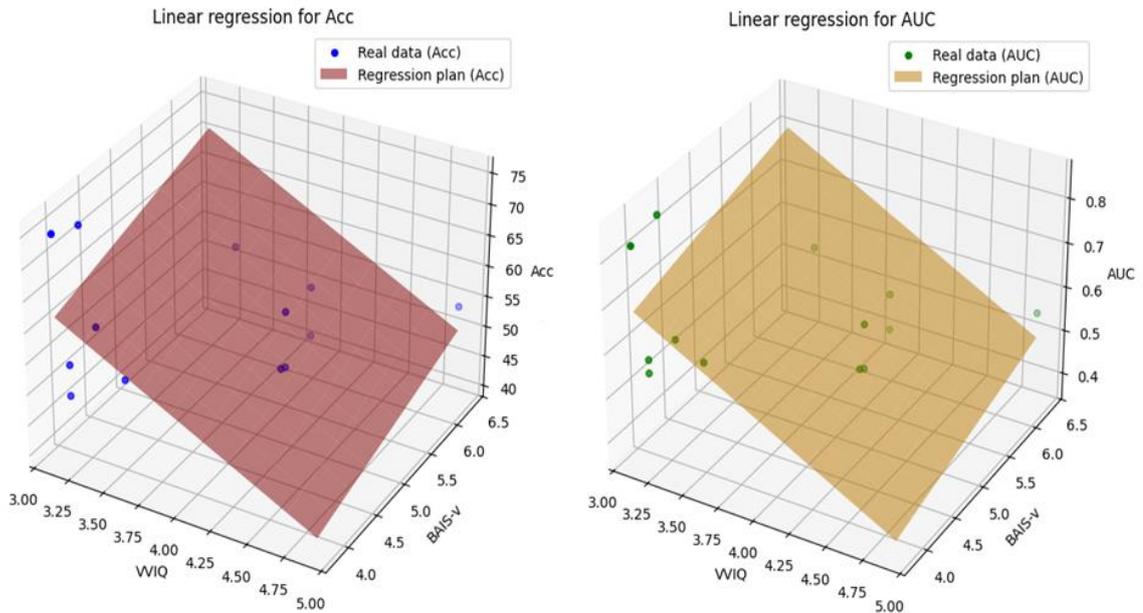


Figure 6.4: Regression planes on the SVM classifier for audio modality on accuracy and AUC.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-11.78	3.68
AUC	-0.17	0.04

Table 8.4: Regression plane inclination for SVM classifier and audio modality.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.4920	0.0876	-0.3071	0.3074
AUC	-0.6040	0.0288	-0.4047	0.1701

Table 9.4: Pearson correlation with VVIQ and BAIS-V for the SVM classifier.

Figure 7.4 shows the comparison results between combined groups for Acc and AUC (audio modality and SVM classifier).

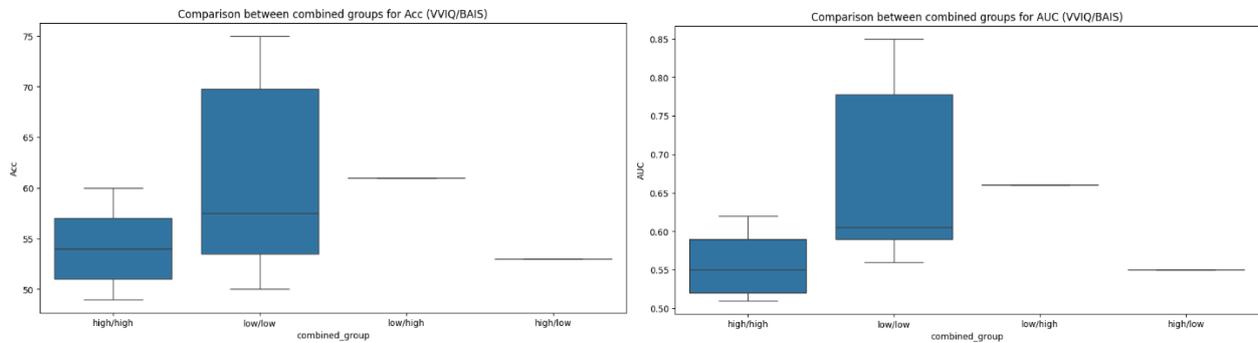


Figure 7.4: Comparison between combined groups for Acc and AUC for SVM classifier. T-test Acc (VVIQ high vs low): $t = -1.6469$, $p = 0.1278$. T-test AUC (BAIS high vs low): $t = -1.4145$, $p = 0.1849$.

Figure 8.4 shows the regression plane for the accuracy and AUC for the audio modality for the MLP classifier.

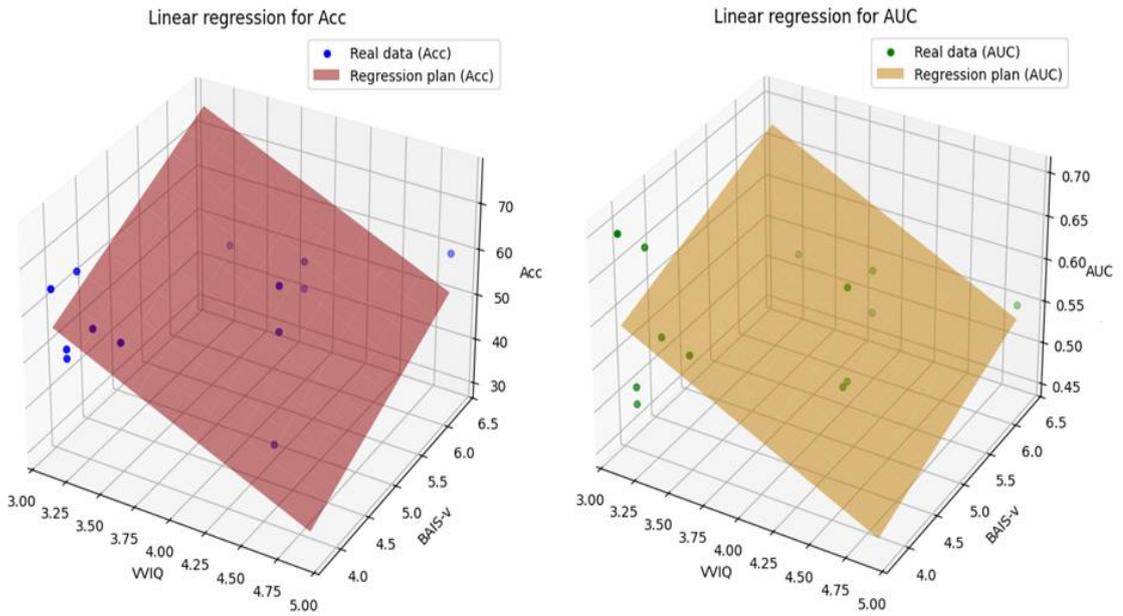


Figure 8.4: Regression planes on the MLP classifier for audio modality on accuracy and AUC.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-13.84	7.75
AUC	-0.07	0.03

Table 10.4: Regression plane inclination for MLP classifier and audio modality.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.2446	0.4206	0.0075	0.9805
AUC	.03773	0.2038	-0.1862	0.5425

Table 11.4: Pearson correlation with VVIQ and BAIS-V for the MLP classifier.

Figure 9.4 shows the comparison results between combined groups for Acc and AUC (audio modality and MLP classifier).

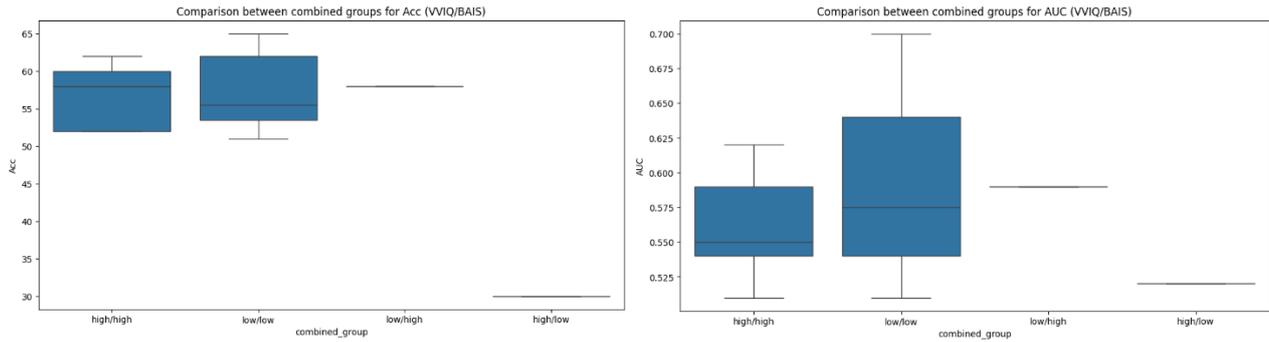


Figure 9.4: Comparison between combined groups for Acc and AUC for MLP classifier. T-test Acc (VVIQ high vs low): $t = -1.0400$, $p = 0.3207$. T-test AUC (BAIS high vs low): $t = -0.4399$, $p = 0.6685$.

4.2 ORTHOGRAPHIC MODALITY

Table 12.4 shows the accuracy and AUC results for each subject and classification method for the orthographic modality.

Subject	RF		LR		SVM		MLP	
	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
3-3	55	0.55	57	0.62	52	0.46	62	0.61
8-3	65	0.71	65	0.71	61	0.68	56	0.60
10-1	85	0.91	75	0.84	73	0.82	70	0.74
11-1	70	0.80	77	0.85	56	0.75	64	0.66
12-1	74	0.80	79	0.85	68	0.74	74	0.78
12-2	73	0.78	70	0.76	63	0.71	66	0.68
13-1	92	0.98	82	0.91	85	0.94	75	0.79
14-1	76	0.85	65	0.70	72	0.78	60	0.62
14-2	74	0.85	71	0.79	65	0.73	63	0.65
15-1	88	0.94	78	0.88	75	0.84	72	0.72
15-2	80	0.89	82	0.89	74	0.82	72	0.75
16-1	75	0.82	84	0.87	78	0.84	78	0.80
17-1	96	0.98	92	0.98	91	0.97	80	0.85
18-1	62	0.69	66	0.68	61	0.66	68	0.68
19-1	89	0.96	77	0.87	78	0.88	76	0.79
Average	76.9	0.83	74.7	0.81	70.1	0.77	69.0	0.71

Table 12.4: Accuracy and AUC for orthographic modality.

ROC curves for orthographical modality are shown in Figure 10.4.

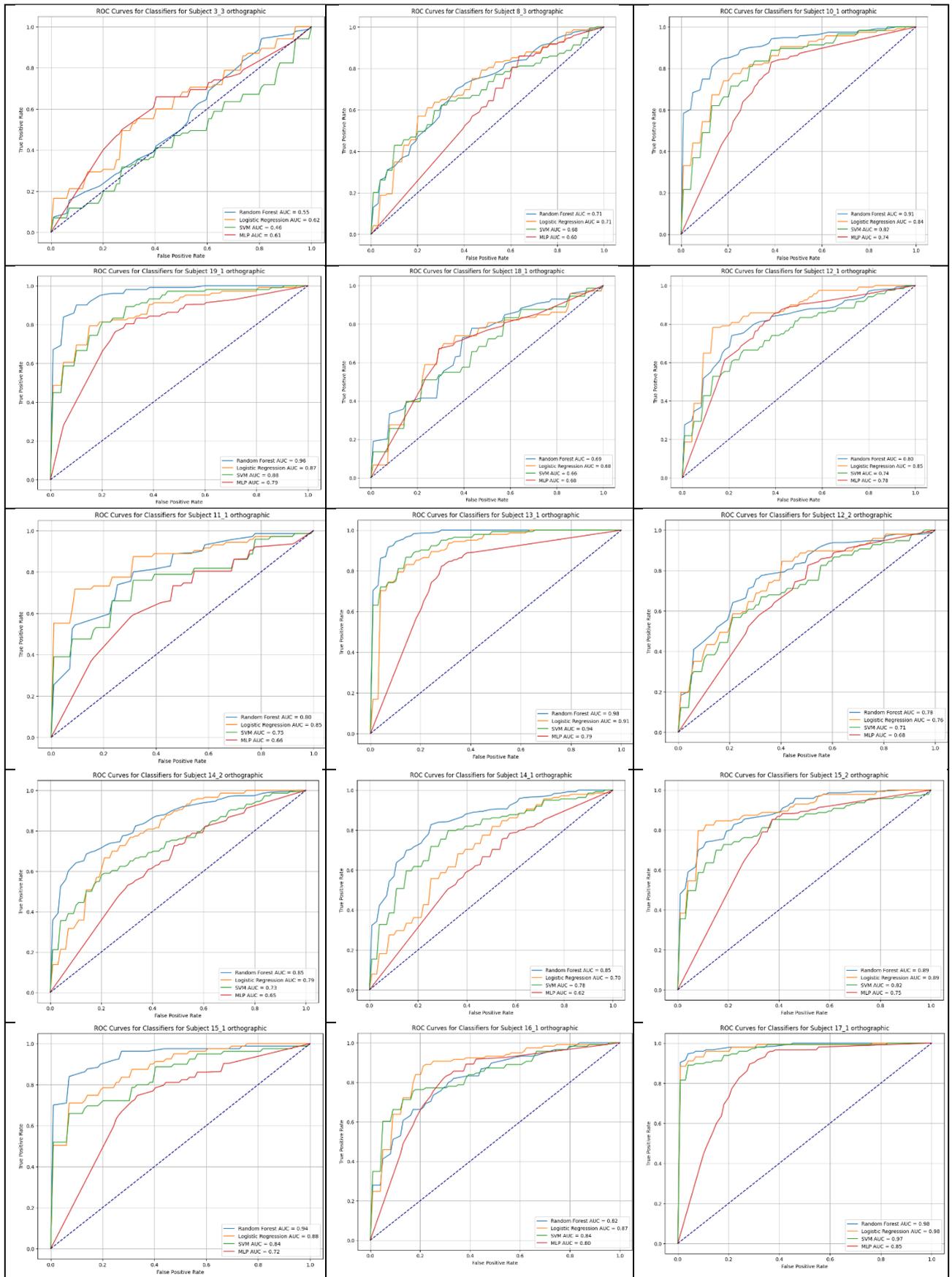


Figure 10.4: ROC curves for orthographic modality. Colors distinguish classifiers.

Figure 11.4 shows the regression plane for the accuracy and AUC on the audio orthographic for the RF classifier.

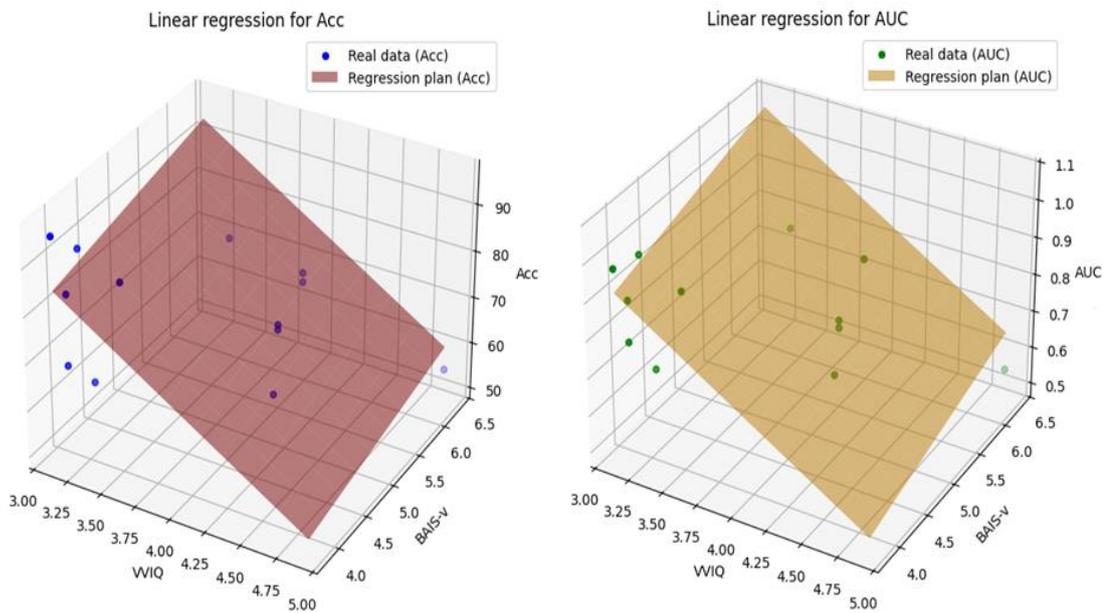


Figure 11.4: Regression planes on the RF classifier for orthographic modality on accuracy and AUC.

The regression plane inclination is shown in Table 13.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-18.74	3.56
AUC	-0.23	0.06

Table 13.4: Regression plane inclination for RF classifier and orthographic modality.

Pearson correlation is shown in Table 14.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.6971	0.0081	-0.5166	0.0707
AUC	-0.7208	0.0054	-0.4928	0.0871

Table 14.4: Pearson correlation with VVIQ and BAIS-V for the RF classifier.

Figure 12.4 shows the comparison results between combined groups for Acc and AUC (orthographic modality and RF classifier).

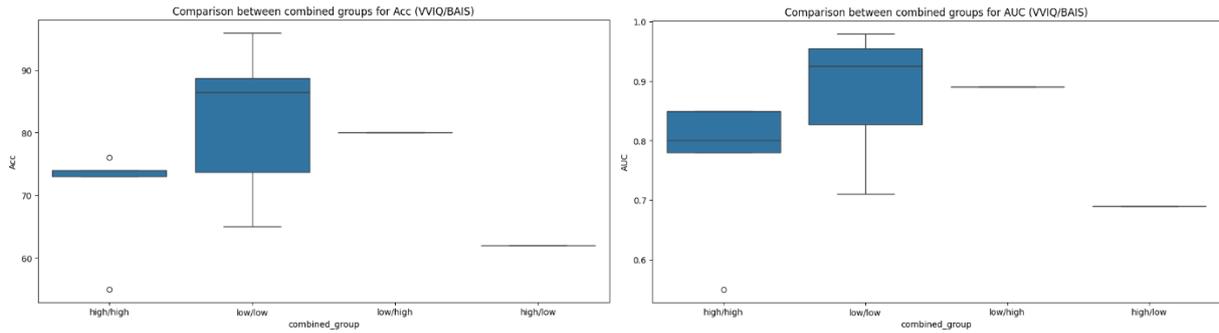


Figure 12.4: Comparison between combined groups for Acc and AUC for RF classifier. T-test Acc (VVIQ high vs low): $t = -2.3249$, $p = 0.0402$. T-test AUC (BAIS high vs low): $t = -1.0192$, $p = 0.3300$.

Figure 13.4 shows the regression plane for the accuracy and AUC on the orthographic modality for the LR classifier.

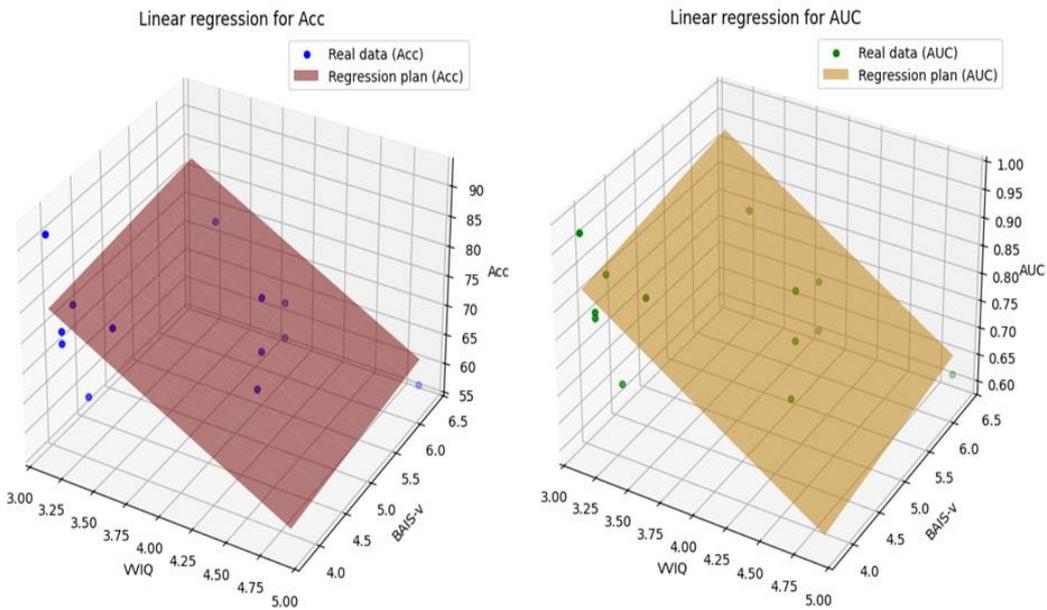


Figure 13.4: Regression planes on the LR classifier for orthographic mode on accuracy and AUC.

The regression plane inclination is shown in Table 15.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-12.07	1.26
AUC	-0.15	0.02

Table 15.4: Regression plane inclination for LR classifier and orthographic modality.

Pearson correlation is shown in Table 16.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.6619	0.0137	-0.5277	0.0638
AUC	-0.7193	0.0056	-0.5601	0.0465

Table 16.4: Pearson correlation with VVIQ and BAIS-V for the LR classifier.

Figure 14.4 shows the comparison results between combined groups for Acc and AUC (orthographic modality and LR classifier).

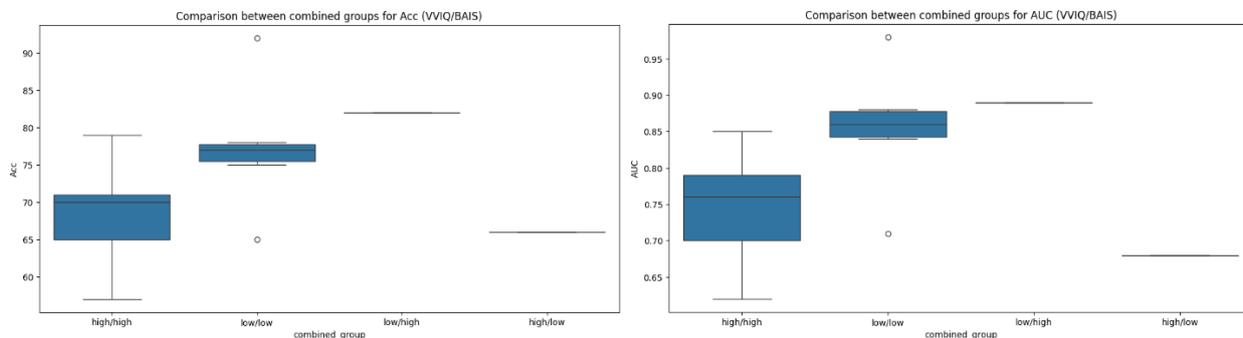


Figure 14.4: Comparison between combined groups for Acc and AUC for LR classifier. T-test Acc (VVIQ high vs low): $t = -2.3205$, $p = 0.0405$. T-test AUC (BAIS high vs low): $t = -1.0951$, $p = 0.2969$.

Figure 15.4 shows the regression plane for the accuracy and AUC on the orthographic modality for the SVM classifier.

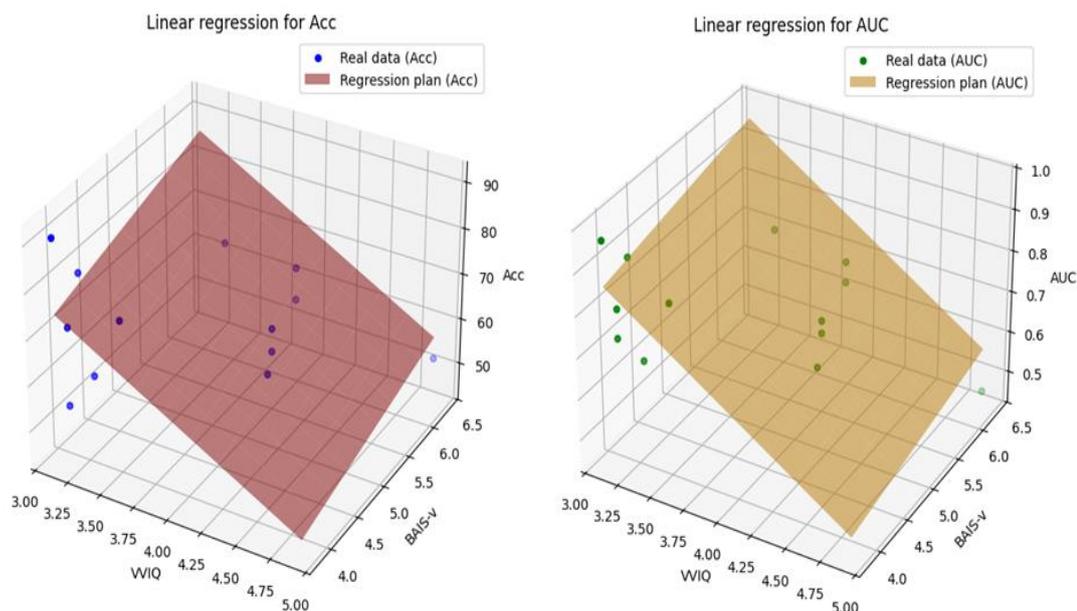


Figure 15.4: Regression planes on the SVM classifier for orthographic modality on accuracy and AUC.

The regression plane inclination is shown in Table 17.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-16.49	4,61
AUC	-0.21	0.03

Table 17.4: Regression plane inclination for SVM classifier and orthographic modality.

Pearson correlation is shown in Table 18.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.5779	0.0386	-0.3821	0.1976
AUC	-0.7819	0.0016	-0.5897	0.0339

Table 18.4: Pearson correlation with VVIQ and BAIS-V for the SVM classifier.

Figure 16.4 shows the comparison results between combined groups for Acc and AUC (orthographic modality and SVM classifier).

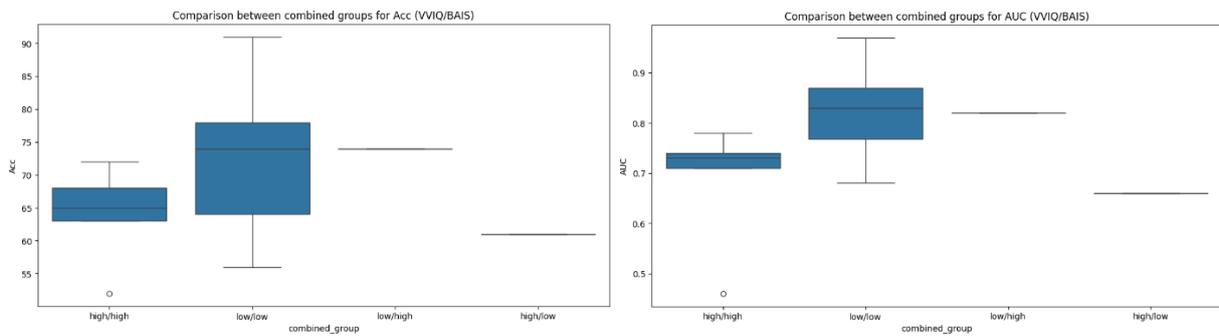


Figure 16.4: Comparison between combined groups for Acc and AUC for SVM classifier. T-test Acc (VVIQ high vs low): $t = -1.7108$, $p = 0.1151$. T-test AUC (BAIS high vs low): $t = -1.4154$, $p = 0.1846$.

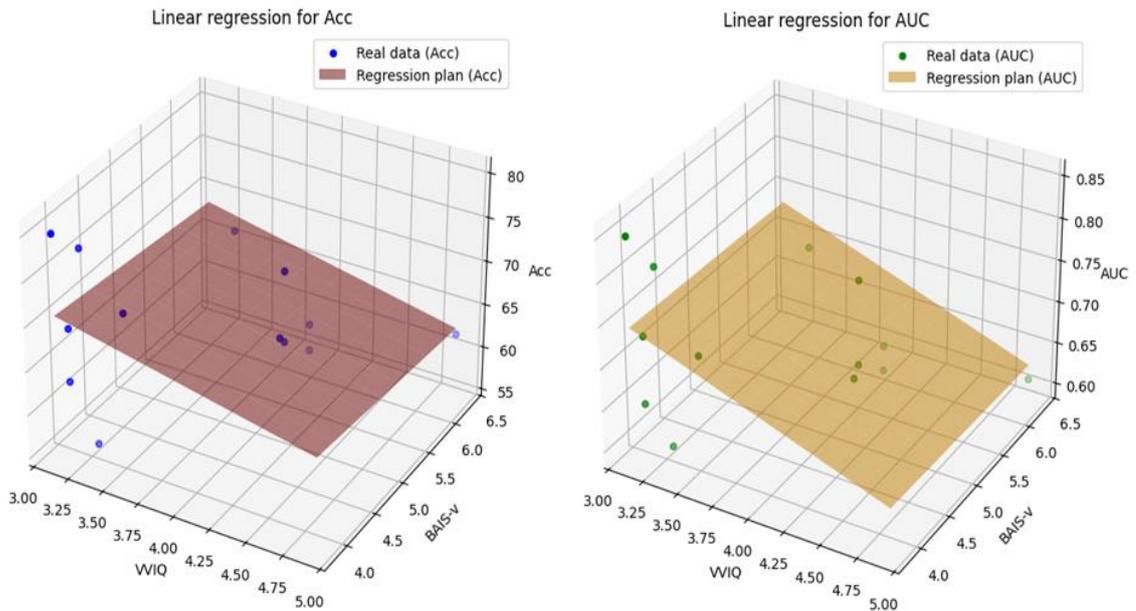


Figure 17.4: Regression planes on the MLP classifier for orthographic modality on accuracy and AUC.

The regression plane inclination is shown in Table 19.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-3.35	-0.86
AUC	-0.06	-0.003

Table 19.4: Regression plane inclination for MLP classifier and orthographic modality.

Pearson correlation is shown in Table 20.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.3740	0.2081	-0.3485	0.2432
AUC	-0.4824	0.0950	-0.4206	0.1524

Table 20.4: Pearson correlation with VVIQ and BAIS-V for the MLP classifier.

Figure 18.4 shows the comparison results between combined groups for Acc and AUC (orthographic modality and MLP classifier).

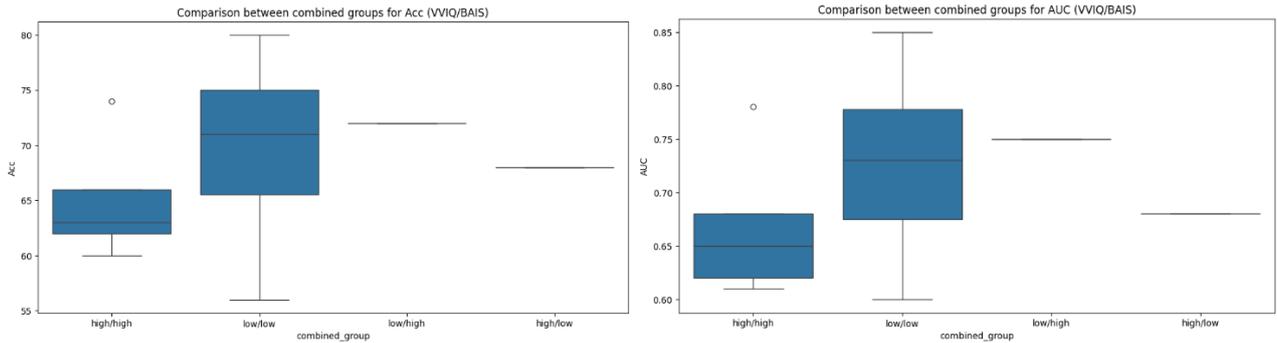


Figure 18.4: Comparison between combined groups for Acc and AUC for MLP classifier. T-test Acc (VVIQ high vs low): $t = -1.1955$, $p = 0.2570$. T-test AUC (BAIS high vs low): $t = -0.8890$, $p = 0.3931$.

4.2 PICTORIAL MODALITY

Table 21.4 shows the accuracy and AUC results for each subject and the classification method concerning the pictorial modality.

Subject	RF		LR		SVM		MLP	
	Acc	AUC	Acc	AUC	Acc	AUC	Acc	AUC
3-3	69	0.71	68	0.73	62	0.64	67	0.71
8-3	59	0.66	69	0.75	61	0.67	66	0.67
10-1	83	0.88	78	0.84	76	0.85	69	0.73
11-1	85	0.91	73	0.81	66	0.80	70	0.74
12-1	76	0.82	73	0.83	71	0.76	75	0.78
12-2	66	0.77	68	0.75	62	0.69	66	0.69
13-1	90	0.95	82	0.90	82	0.91	78	0.83
14-1	83	0.90	65	0.71	71	0.80	58	0.61
14-2	71	0.78	72	0.79	63	0.73	65	0.65
15-1	91	0.97	90	0.98	86	0.93	85	0.86
15-2	84	0.91	83	0.91	81	0.89	77	0.81
16-1	71	0.78	76	0.85	67	0.79	72	0.72
17-1	93	0.97	89	0.96	88	0.95	83	0.86
18-1	67	0.81	73	0.79	71	0.78	68	0.69
19-1	87	0.95	86	0.92	82	0.90	80	0.83
Average	78.3	0.85	76.3	0.83	72.6	0.81	71.9	0.75

Table 21.4: Accuracy and AUC for pictorial modality.

ROC curves for orthographical modality are shown in Figure 19.4.

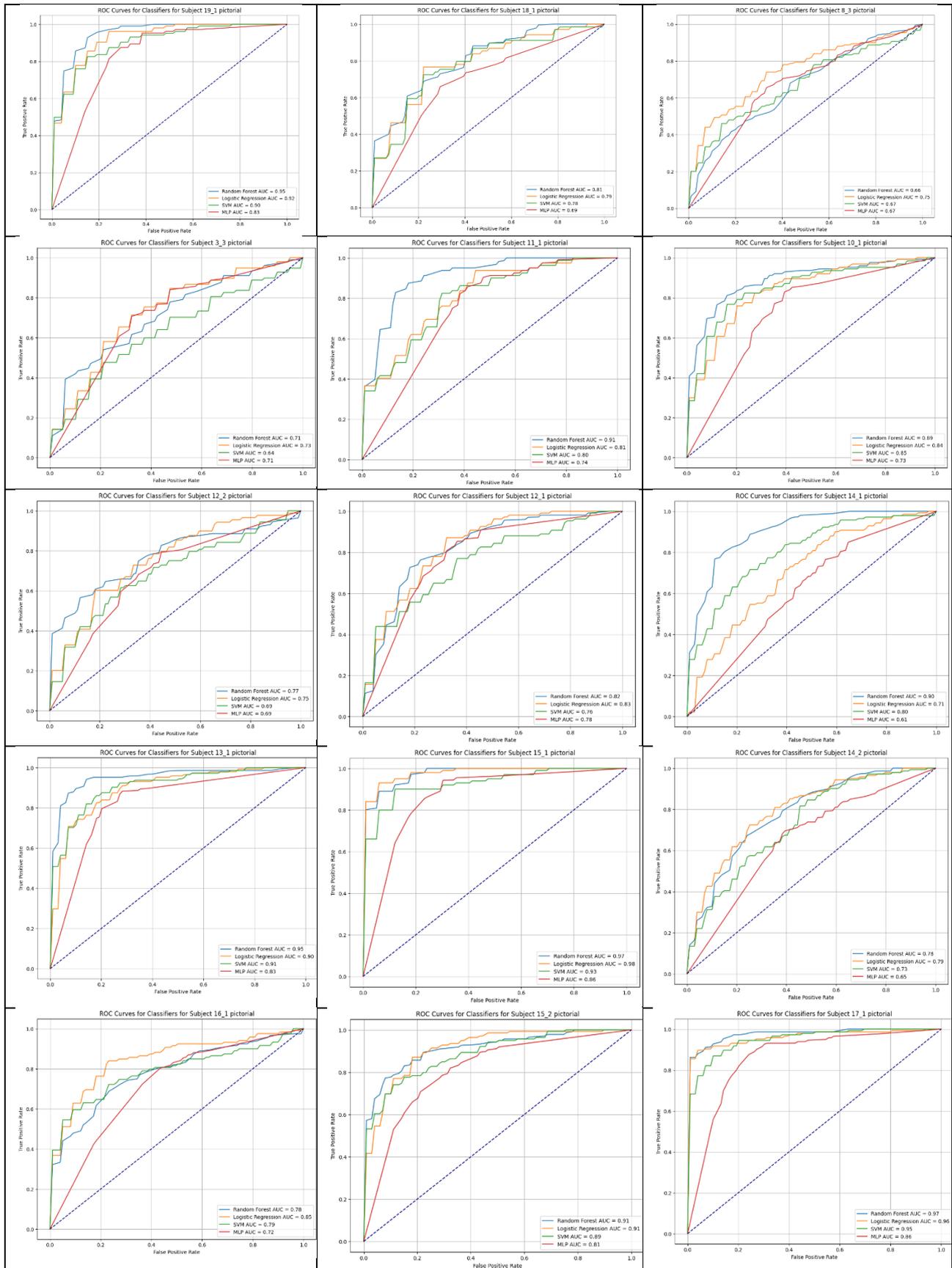


Figure 19.4: ROC curves for pictorial modality.

Figure 20.4 shows the regression plane for the accuracy and AUC on the pictorial modality for the RF classifier.

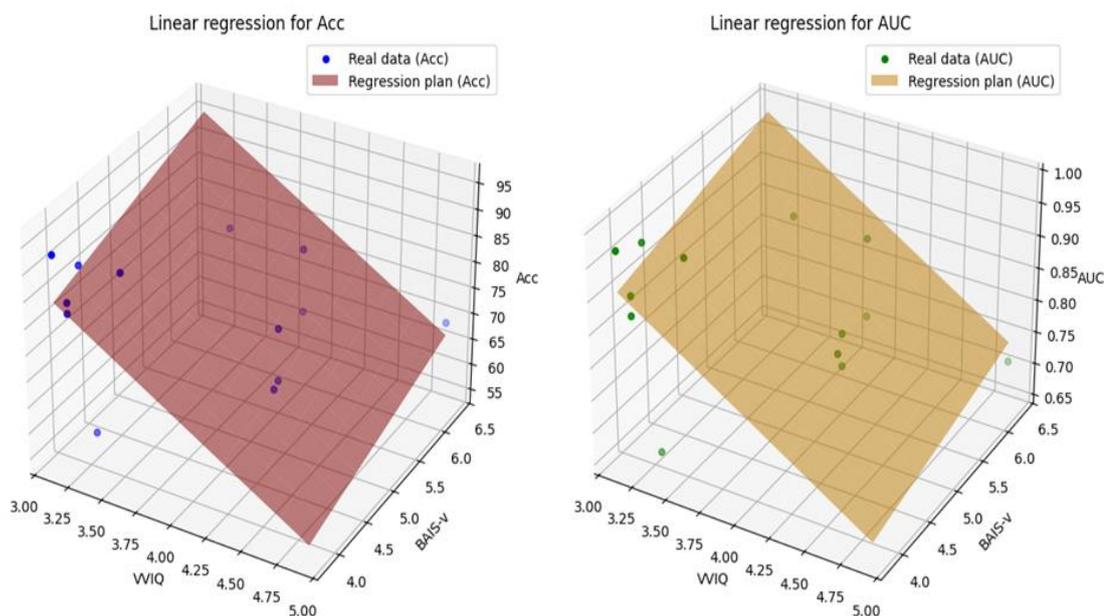


Figure 20.4: Regression planes on the RF classifier for pictorial mode on accuracy and AUC.

The regression plane inclination is shown in Table 22.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-16.13	4.51
AUC	-0.13	0.030

Table 22.4: Regression plane inclination for RF classifier and pictorial modality.

Pearson correlation is shown in Table 23.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.5519	0.0505	-0.2650	0.2201
AUC	-0.5589	0.0471	-0.4004	0.1752

Table 23.4: Pearson correlation with VVIQ and BAIS-V for the RF classifier.

Figure 21.4 shows the results of the comparison between combined groups for Acc and AUC (pictorial modality and RF classifier).

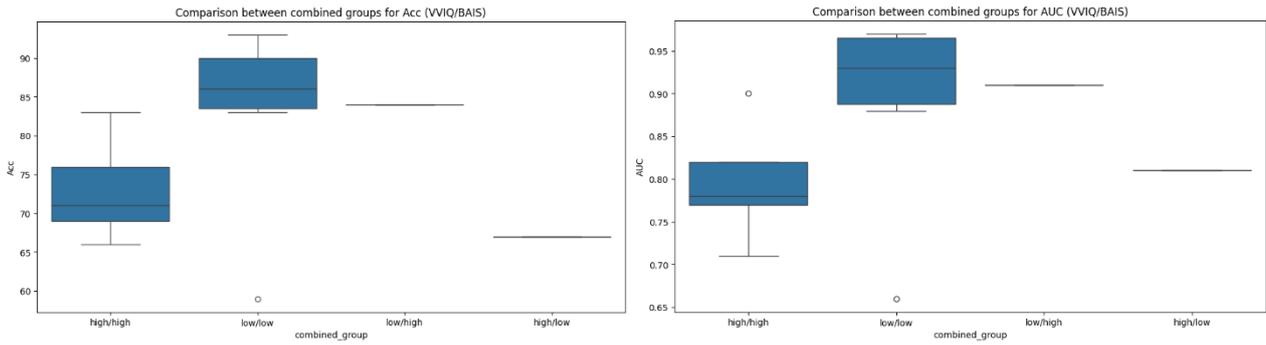


Figure 21.4: Comparison between combined groups for Acc and AUC for RF classifier. T-test Acc (VVIQ high vs low): $t = -2.1341$, $p = 0.0562$. T-test AUC (BAIS high vs low): $t = -1.1649$, $p = 0.2687$.

Figure 22.4 shows the regression plane for the accuracy and AUC on the pictorial modality for the LR classifier.

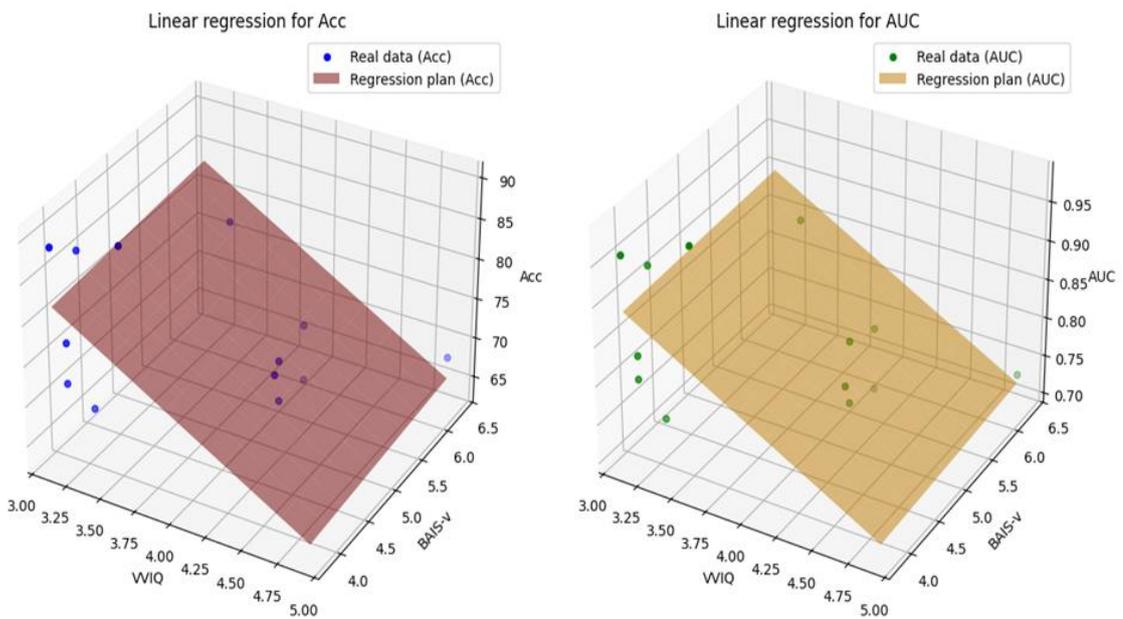


Figure 22.4: Regression planes on the LR classifier for pictorial mode on accuracy and AUC.

The regression plane inclination is shown in Table 24.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-10.13	0.66
AUC	-0.10	0.004

Table 24.4 : Regression plane inclination for LR classifier and pictorial modality.

Pearson correlation is shown in Table 25.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.6293	0.0212	-0.5150	0.0717
AUC	-0.6170	0.0247	-0.5117	0.0739

Table 25.4: Pearson correlation with VVIQ and BAIS-V for the LR classifier.

Figure 23.4 shows the comparison results between combined groups for Acc and AUC (pictorial modality and LR classifier).

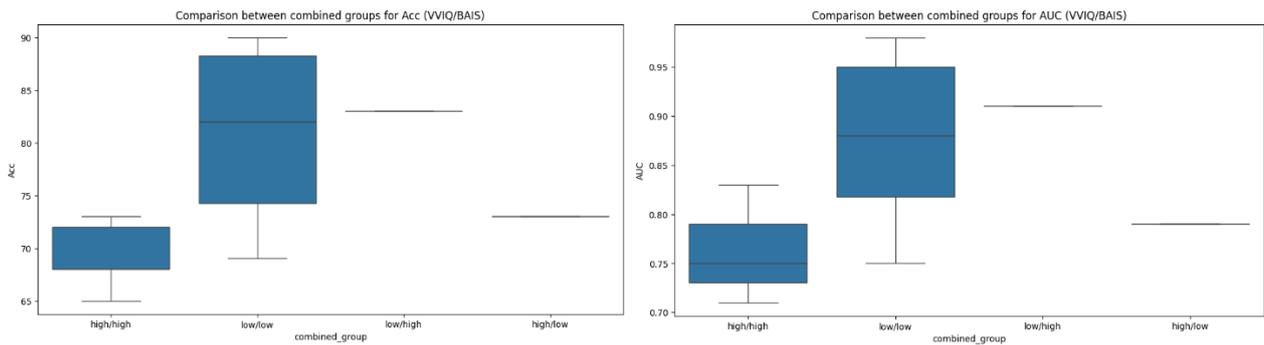


Figure 23.4: Comparison between combined groups for Acc and AUC for LR classifier. T-test Acc (VVIQ high vs low): $t = -3.1936$, $p = 0.0086$. T-test AUC (BAIS high vs low): $t = -1.6862$, $p = 0.1199$.

Figure 24.4 shows the regression plane for the accuracy and AUC on the pictorial modality for the SVM classifier.

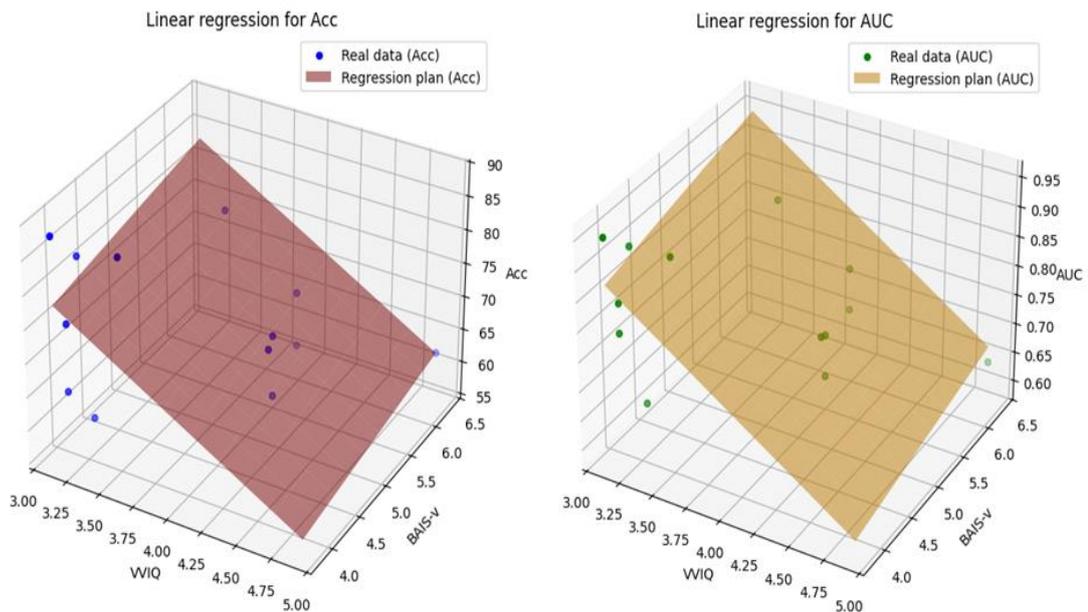


Figure 24.4: Regression planes on the SVM classifier for pictorial modality on accuracy and AUC.

The regression plane inclination is shown in Table 26.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-11.89	2.11
AUC	-0.15	0.02

Table 26.4: Regression plane inclination for SVM classifier and pictorial modality.

Pearson correlation is shown in Table 27.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.5505	0.0512	-0.4127	0.1611
AUC	-0.6682	0.0125	-0.4955	0.0851

Table 27.4: Pearson correlation with VVIQ and BAIS-V for the SVM classifier.

Figure 25.4 shows the results of the comparison between combined groups for Acc and AUC (pictorial modality and SVM classifier).

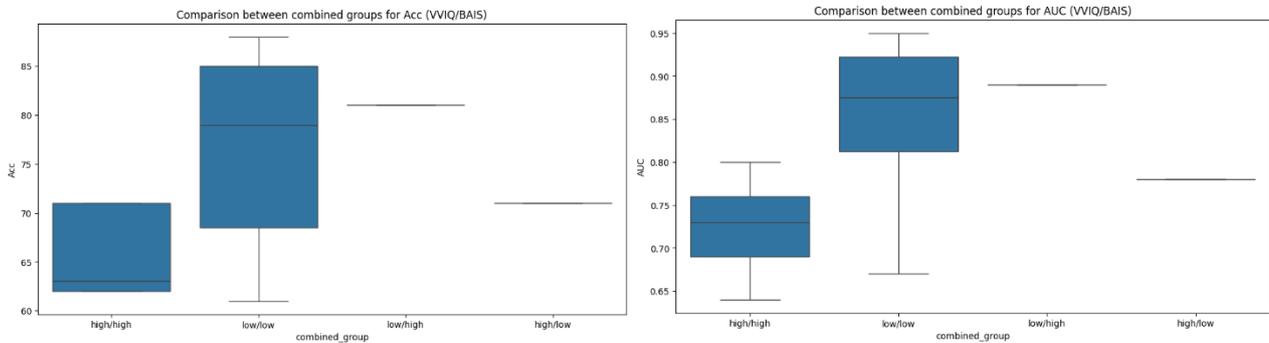


Figure 25.4: Comparison between combined groups for Acc and AUC for combined SVM classifier. T-test Acc (VVIQ high vs low): $t = -2.3049$, $p = 0.0417$. T-test AUC (BAIS high vs low): $t = -1.6971$, $p = 0.1178$.

Figure 26.4 shows the regression plane for the accuracy and AUC on the pictorial modality for the MLP classifier.

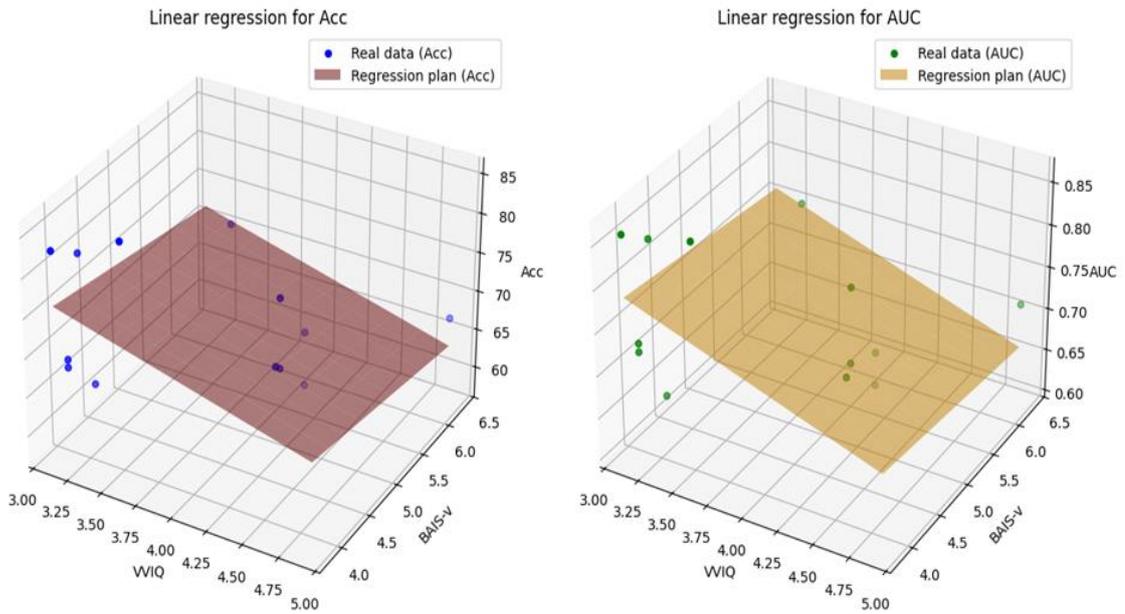


Figure 26.4: Regression planes on the MLP classifier for pictorial modality on accuracy and AUC.

The regression plane inclination is shown in Table 28.4.

Metric	Inclination to VVIQ	Inclination to BAIS-V
Accuracy	-4.72	-1.67
AUC	-0.056	-0.011

Table 28.4: Regression plane inclination for MLP classifier and pictorial modality.

Pearson correlation is shown in Table 29.4.

Metric	Correlation with VVIQ	P-value (VVIQ)	Correlation with BAIS-V	P-value (BAIS-V)
Accuracy	-0.5001	0.0818	-0.4766	0.0997
AUC	-0.5015	0.0808	-0.4584	0.1151

Table 29.4: Pearson correlation with VVIQ and BAIS-V for the MLP classifier.

Figure 27.4 shows the results of the comparison between combined groups for Acc and AUC (pictorial modality and MLP classifier).

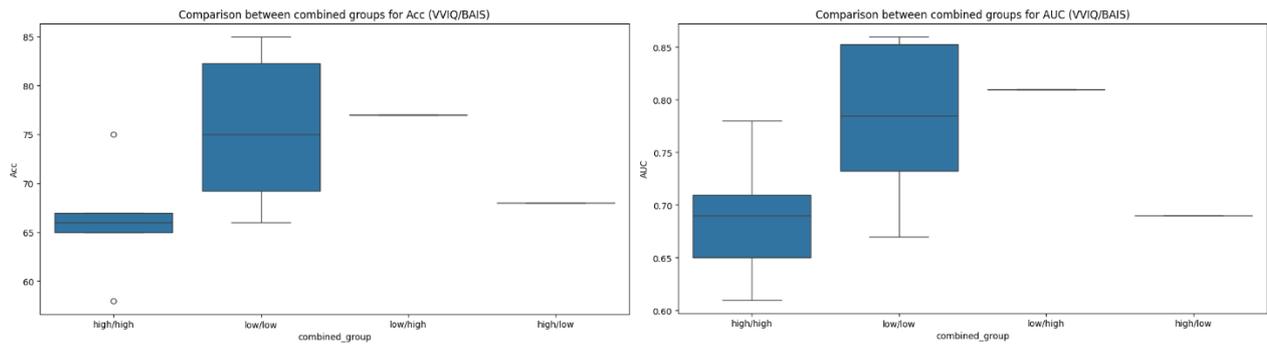


Figure 27.4: Comparison between combined groups for Acc and AUC for MLP classifier. T-test Acc (VVIQ high vs low): $t = -2.5048$, $p = 0.0293$. T-test AUC (BAIS high vs low): $t = -1.3815$, $p = 0.1945$.

5 DISCUSSION

The results presented in the previous section underscore a substantial relationship between self-reported vividness scores (VVIQ and BAIS-V) and the performance of the four classification models employed for the various stimulus modalities (audio, orthographic, and pictorial). Tables 3.4, 12.4, and 21.4 demonstrate that VVIQ scores exert a more significant influence than BAIS-V scores in determining accuracy and AUC metrics. This finding aligns with the logical hypothesis that the vividness measured by VVIQ plays a more substantial role in imagination and perception tasks compared to auditory vividness measured by BAIS-V. The following subsections will specifically discuss the results for each stimulus modality.

5.1 AUDIO MODALITY

For the audio modality, as illustrated in subsection 4.1, Figure 2.4 demonstrates that the regression plane for the RF model exhibits a stronger inclination toward VVIQ (-18.31) compared to BAIS-V. Table 5.4 indicates significant relationships between VVIQ and accuracy, with a Pearson correlation of -0.5636 ($p = 0.0441$). This suggests that higher VVIQ scores are associated with reduced accuracy. Conversely, the relationship with BAIS-V is not significant, as indicated by a Pearson correlation of $p = 0.3128$. A similar trend is observed for AUC, where BAIS-V does not yield significant results ($p = 0.3541$). Conversely, the correlation between AUC and VVIQ remains negative and significant, albeit slightly weaker (-0.5495, $p = 0.0517$). Figure 3.4 demonstrates that a T-test reveals a marginally significant difference in accuracy between high and low VVIQ groups but not for AUC.

In a similar vein, Figure 4.4 underscores a pronounced tendency of the LR classifier toward VVIQ, as evidence by its superior performance in terms of accuracy (-15.21) and AUC (-0.16) when compared to BAIS-V (5.61 and 0.05). The correlation between VVIQ and accuracy in Table 7.4 is more significant than that of the RF model, indicating a relevant relationship (-0.6404, $p = 0.018$). AUC is also strongly correlated with VVIQ (-0.6078, $p = 0.0276$). As in the previous case, the correlation with BAIS-V does not produce appreciable results for either metric. As established in Figure 5.4, a statistically big discrepancy in accuracy is discovered among the excessive and coffee VVIQ companies ($p = 0.0647$). However, no statistically big variations are discovered among the excessive and coffee BAIS-V companies. Utilising the SVM classifier, Figure 6.4 and Table 8.4 suggest a discounted effect of VVIQ, with dispositions of -11.seventy eight for accuracy and -0.17 for AUC. A comparable decline is discovered for BAIS-V, suggesting a minimum impact on each metrics. Table 9.4 confirms that the correlation among VVIQ and accuracy is weaker as compared to the preceding models. Figure 7.4 suggests that t-test do not display big variations among high and low VVIQ or BAIS-V groups for both accuracy or AUC.

For the MLP version, comparable conclusions to the SVM version may be drawn. The findings suggest that VVIQ and BAIS-V exert minimum impact on, showing simplest modest and non-big correlations for each metrics, and that the accuracy and AUC among those groups do not yield applicable results.

5.2 ORTHOGRAPHIC MODALITY

In this case as well, even when a different kind of stimulus is used, it should be noted that the regression plane is significantly more influenced by VVIQ (-18.74) than by BAIS-V (3.56), as illustrated in Figure 11.4 and Table 13.4. AUC shows a similar trend, with a greater inclination toward VVIQ (-0.23), albeit less pronounced than for accuracy. Table 14.4 corroborates this, showing an appreciable correlation between accuracy and VVIQ ($p = 0.0081$), which becomes even stronger for AUC ($p = 0.0054$). Figure 12.4 highlights a significant difference in accuracy by comparing high and low VVIQ groups. Across all analyses, the role of BAIS-V in RF model is marginal.

The regression plane inclinations of the LR model in Figure 13.4 demonstrate that VVIQ exerts a more substantial influence on accuracy (-12.07) and AUC (-0.15) compared to BAIS-V. Table 16.4 corroborates the observations by VVIQ participants achieving a higher degree of accuracy as shown in Table 16.4, which has a statistically significant p value of 0.0137; however, that p -value concerning the correlation with BAIS-V is close to zero. These results suggest that both self-reported scales, that is VVIQ and BAIS-V, do reasonably affect models' discriminative power as the AUCs are highly correlated with VVIQ and BAIS-V. Further comparison between the groups demonstrates these differences and adds to those who argue that VVIQ is much more pronounced in effect than BAIS V. The superiority of VVIQ over BAIS-V is consistently observed in both metrics when employing the SVM model. The findings in Table 18.4 indicate a substantial negative correlation between AUC and VVIQ (-0.7819, $p = 0.0016$) and a significant positive correlation with BAIS-V. This observation indicates that, in contrast to previous cases, BAIS-V exerts a more substantial influence on the performance of the classifiers. The group comparison does not strengthen our hypothesis that individuals with low versus high VVIQ and BAIS-V scores will perform differently. Similar to audio modality, the effectiveness of MLP model also decreases. For example in Figures 17.4 and 18.4, VVIQ or BAIS-V do not have any influence on model performance. The results in Tables 19.4 and 20.4 illustrate the classifier deficiencies as indicated by the p -value (VVIQ) of 0.2081 and 0.0950, respectively, for accuracy.

5.3 PICTORIAL MODALITY

Figure 20.4 reveals a similar trend for VVIQ and BAIS-V, with both showing marginally significant results for accuracy. Tables 22 and 23 highlight a weak correlation with VVIQ for both accuracy and AUC (-0.5519, $p = 0.0505$; -0.5589, $p = 0.0471$). The correlation values for BAIS-V are not appreciable. As illustrated in Figure 18.4, a significant discrepancy is observed between the VVIQ and BAIS-V groups.

The RF classifier proves quite effective in aligning VVIQ with performance of models, something that is not exhibited by the BAIS-V group. In addition, Tables 24.4 and 25.4, along with Figure 22.4, highlight the dominance of VVIQ over BAIS-V. This trend is further reinforced by the observation in Figure 23, where there is a significant difference in accuracy between high VVIQ and low VVIQ groups ($p=0.0086$) but where that difference cannot be statistically established for AUC.

The LR model aligns with the RF model, indicating that VVIQ has a substantial impact on classifier performance, while BAIS-V exhibits minimal influence.

The SVM model further substantiates the notion that VVIQ is the predominant self-reported score in influencing classifier performance. As illustrated in Figure 24.4, there is a pronounced tendency for VVIQ, as evidenced by the negative values for accuracy and AUC. Table 26.4 underscores a substantial correlation between AUC and VVIQ, while the association between BAIS-V and both AUC and accuracy is negligible.

As in previous instances, the MLP model consistently yields unsatisfactory outcomes across all statistical analyses. The regression plane inclinations depicted in Figure 26.4 are notably low. Tables 28.4 and 29.4 underscore the absence of significant correlations for both self-reported scores and metrics.

6 CONCLUSIONS

The findings of this study indicate that visual and imaginative abilities are superior predictors of performance when compared to auditory imagined abilities, as evidenced by the superior performance on VVIQ in comparison to the other. However, it is also evident that VVIQ exerts a significant influence on the ascending performance of all models, with the exception of the Multilayer Perceptron, which is based on the dependency between self-report scores and EEG signal characteristics. The correlations further suggest that participants with a higher VVIQ may impede the differentiation among the neural models.

One potential explanation for this phenomenon is that strong VVIQ scores are associated with more complex neural patterns, particularly in the orthographic and pictorial modalities, where VVIQ emerges as a robust predictor of accuracy and AUC metrics. There are a number of limitations to this study. First and foremost, the dataset contains only 15 participants, which produce results with limited statistical power and makes them. Additionally, the scores obtained from self-reported questionnaire are contaminated by external factors, such as participant focus or their personal experiences.

The differing results observed across stimulus modalities could stem from confounding factors, such as the complexity of the stimuli and their activation of specific cortical areas. Future studies should aim to expand the dataset to improve generalizability. It is imperative to acknowledge the subjective nature of questionnaire-based measures and to introduce objective metrics that account for individual participant characteristics. Furthermore, exploring new EEG signal features to determine their impact and managing confounding factors during test preparation, such as stimulus type and familiarity, could provide a more comprehensive evaluation and enhance the predictive capabilities of the models.

REFERENCES

- [1] Snell, R. S. (2010). *Clinical neuroanatomy*. Lippincott Williams & Wilkins.
- [2] Schmidt, R. F., Dudel, J., Jaenig, W., & Zimmermann, M. (2012). *Fundamentals of neurophysiology*. Springer Science & Business Media.
- [3] Snell, R. S. (2010). *Clinical neuroanatomy*. Lippincott Williams & Wilkins.
- [4] Blinowska, K., & Durka, P. (2006). Electroencephalography (eeg). *Wiley encyclopedia of biomedical engineering*, 10, 9780471740360.
- [5] Teplan, M. (2002). Fundamentals of EEG measurement. *Measurement science review*, 2(2), 1-11.
- [6] Bronzino, J. D., & Peterson, D. R. (2006). Principles of electroencephalography. In *Biomedical Engineering Fundamentals* (pp. 445-456). CRC press.
- [7] Singh, S. P. (2014). Magnetoencephalography: basic principles. *Annals of Indian Academy of Neurology*, 17(Suppl 1), S107-S112.
- [8] De Gennaro, L., Scarpelli, S., & Bartolacci, C. (2017). L'attività elettrica cerebrale (EEG) predice la presenza del ricordo dei sogni?. *Rivista sperimentale di freniatria: la rivista dei servizi di salute mentale: CXLI*, 2, 2017, 79-99.
- [9] Lega Italiana Contro l'Epilessia. (2023). *Guida alle epilessie*. https://www.lice.it/pdf/web_2023_guida_alle_epilessie.pdf
- [10] Värbu, K., Muhammad, N., & Muhammad, Y. (2022). Past, present, and future of EEG-based BCI applications. *Sensors*, 22(9), 3331.
- [11] Al-Saegh, A., Dawwd, S. A., & Abdul-Jabbar, J. M. (2021). Deep learning for motor imagery EEG-based classification: A review. *Biomedical Signal Processing and Control*, 63, 102172.
- [12] EEG based brain computer interfaces using motor imagery: techniques and challenges. Natasha Padfield, Jaime Zabalza, Huimin Zhao.
- [13] Urigüen, J. A., & Garcia-Zapirain, B. (2015). EEG artifact removal—state-of-the-art and guidelines. *Journal of neural engineering*, 12(3), 031001.
- [14] Jiang, X., Bian, G. B., & Tian, Z. (2019). Removal of artifacts from EEG signals: a review. *Sensors*, 19(5), 987.
- [15] Nuwer, M. R. (1998). Fundamentals of evoked potentials and common clinical applications today. *Electroencephalography and clinical neurophysiology*, 106(2), 142-148.
- [16] Tandle, A., Jog, N., D'cunha, P., & Chheta, M. (2015). Classification of artefacts in EEG signal recordings and overview of removing techniques. *International Journal of Computer Applications*, 975, 8887.
- [17] Al-Fahoum, A. S., & Al-Fraihat, A. A. (2014). Methods of EEG Signal Features Extraction Using Linear Analysis in Frequency and Time-Frequency Domains. *International Scholarly Research Notices*, 2014(1), 730218.

- [18] Frisby, S. L., Halai, A. D., Cox, C. R., Ralph, M. A. L., & Rogers, T. T. (2023). Decoding semantic representations in mind and brain. *Trends in cognitive sciences*, 27(3), 258-281.
- [19] Binder, J. R., & Desai, R. H. (2011). The neurobiology of semantic memory. *Trends in cognitive sciences*, 15(11), 527-536.
- [20] Coronel, J. C., & Federmeier, K. D. (2016). The N400 reveals how personal semantics is processed: Insights into the nature and organization of self-knowledge. *Neuropsychologia*, 84, 36-43.
- [21] Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual review of psychology*, 62(1), 621-647.
- [22] Binder, J. R., Desai, R. H., Graves, W. W., & Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral cortex*, 19(12), 2767-2796.
- [23] Yee, E., & Thompson-Schill, S. L. (2016). Putting concepts into context. *Psychonomic bulletin & review*, 23, 1015-1027.
- [24] Barsalou, L. W. (1982). Context-independent and context-dependent information in concepts. *Memory & cognition*, 10(1), 82-93.
- [25] Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *science*, 303(5664), 1634-1640.
- [26] O'Connor, A. R., Han, S., & Dobbins, I. G. (2010). The inferior parietal lobule and recognition memory: expectancy violation or successful retrieval?. *Journal of neuroscience*, 30(8), 2924-2934.
- [27] Hnazaee, M. F., & Van Hulle, M. M. (2017, May). Typicality effect on N400 ERP in categories despite differences in semantic processing. In *2017 International Joint Conference on Neural Networks (IJCNN)* (pp. 4379-4386). IEEE.
- [28] Bell, M. A., & Cuevas, K. (2012). Using EEG to study cognitive development: Issues and practices. *Journal of cognition and development*, 13(3), 281-294.
- [29] Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in cognitive sciences*, 4(12), 463-470.
- [30] Anderson, A. J., Binder, J. R., Fernandino, L., Humphries, C. J., Conant, L. L., Aguilar, M., ... & Raizada, R. D. (2017). Predicting neural activity patterns associated with sentences using a neurobiologically motivated model of semantic representation. *Cerebral Cortex*, 27(9), 4379-4395.
- [31] Lilleberg, J., Zhu, Y., & Zhang, Y. (2015, July). Support vector machines and word2vec for text classification with semantic features. In *2015 IEEE 14th International Conference on Cognitive Informatics & Cognitive Computing (ICCI* CC)* (pp. 136-140). IEEE.
- [32] Pennington, J., Socher, R., & Manning, C. D. (2014, October). Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)* (pp. 1532-1543).

- [33] Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Front Biosci*, 14(14), 321-332.
- [34] Freeman, W. J., & van Dijk, B. W. (1987). Spatial patterns of visual cortical fast EEG during conditioned reflex in a rhesus monkey. *Brain research*, 422(2), 267-276.
- [35] Lutzenberger, W., Pulvermüller, F., & Birbaumer, N. (1994). Words and pseudowords elicit distinct patterns of 30-Hz EEG responses in humans. *Neuroscience letters*, 176(1), 115-118.
- [36] Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and brain sciences*, 22(2), 253-279.
- [37] Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in cognitive sciences*, 5(12), 517-524.
- [38] Kirschfeld, K. (1995). Neuronal oscillations and synchronized activity in the central nervous system: functional aspects. *Psychology*, 6(36), 1-6.
- [39] «Wilson, H., Golbabaee, M., Proulx, M.J., & Charles, S. (2023). EEG-based BCI Dataset of Semantic Concepts for Imagination and Perception Tasks, *Scientific Data*, 10(1), 1-11. <https://doi.org/10.18112/openneuro.ds004306.v1.0.1>».
- [40] Aphantasia Network. Vividness of Visual Imagery Questionnaire (VVIQ). Aphantasia Network. <https://aphantasia.com/study/vviq/>
- [41] McKelvie, S. J. (1995). The VVIQ as a psychometric test of individual differences in visual imagery vividness: a critical quantitative review and plea for direction. *Journal of Mental imagery*.
- [42] Tabi, Y. A., Maio, M. R., Attaallah, B., Dickson, S., Drew, D., Idris, M. I., ... & Husain, M. (2022). Vividness of visual imagery questionnaire scores and their relationship to visual short-term memory performance. *Cortex*, 146, 186-199.
- [43] Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, 25(1), 37.
- [44] Jurafsky, D. (2000). Speech and language processing.
- [45] Da Broi, G. (2012). Modelli di reti neurali: multilayer perceptron e radial basis function.