



**Politecnico  
di Torino**

**Politecnico di Torino**

Master's Thesis in Biomedical Engineering

**Data-Driven Seizure Prediction  
Using EEG and ECG Signals**

**Candidate:**

Olivotto Matteo

**Supervisors:**

Prof. Mesin Luca

Academic Year 2024/25

# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Nervous System . . . . .	9
1.2	Epilepsy . . . . .	13
1.2.1	Diagnoses and consequences . . . . .	13
1.2.2	Classification . . . . .	13
1.2.3	Treatment . . . . .	15
1.3	EEG and ECG . . . . .	16
1.3.1	Electroencephalography (EEG) . . . . .	17
1.3.2	Electrocardiography (ECG) . . . . .	20
1.4	Machine Learning . . . . .	22
1.5	Neural Networks . . . . .	23
1.5.1	Long Short-Term Memory (LSTM) . . . . .	23
1.5.2	Convolutional Neural Networks (CNN) . . . . .	24
1.5.3	Recurrent Neural Networks with Gated Recurrent Units (RNN-GRU) . . . . .	24
1.5.4	Temporal Convolutional Networks (TCN) . . . . .	25
1.6	Neural Network Intelligence (NNI) . . . . .	25
<b>2</b>	<b>State of the art</b>	<b>27</b>
2.1	EEG based systems . . . . .	27
2.2	ECG based systems . . . . .	30
<b>3</b>	<b>Methods</b>	<b>33</b>
3.1	Dataset . . . . .	34
3.2	Preprocess . . . . .	35
3.2.1	Filtering . . . . .	35
3.2.2	Signal labeling . . . . .	36
3.2.3	Region mean calculation . . . . .	39
3.2.4	Signal windowing . . . . .	40
3.2.5	Feature extraction . . . . .	40
3.3	Dataset handling . . . . .	44
3.4	Classification training . . . . .	45
3.4.1	ML training . . . . .	45
3.4.2	NN training . . . . .	46
3.4.3	Final Classification . . . . .	47

<b>4</b>	<b>Results</b>	<b>49</b>
4.1	Patient specific system . . . . .	50
4.1.1	ECG signal . . . . .	50
4.1.2	EEG signal . . . . .	52
4.2	Generalised system . . . . .	55
4.2.1	ECG signal . . . . .	55
4.2.2	EEG signal . . . . .	56
4.3	Final discussion and conclusion . . . . .	58

# List of Figures

1.1	Nervous System anatomy. . . . .	9
1.2	Nervous System diagram. . . . .	10
1.3	Anatomy of the parasympathetic innervation. . . . .	11
1.4	Brain functional map. . . . .	12
1.5	Epilepsy classification diagram. . . . .	14
1.6	10-20 international standard map. . . . .	17
1.7	Example of 3-Hz (typical) generalized spike-wave IED. . . . .	19
1.8	Example of a patient who had a generalized axial myoclonic jerk . . .	20
1.9	12 leads electrodes placement. . . . .	20
1.10	Physiological ECG shape. . . . .	21
1.11	HRV calculated on an ECG signal. . . . .	21
3.1	PSD comparison for the raw and filtered ECG signals. . . . .	37
3.2	PSD comparison for the raw and filtered EEG signals. . . . .	38

# List of Tables

1.1	Comparison of amplitude and frequency ranges for ECG, EEG, and EMG techniques. . . . .	17
3.1	Filtering parameters . . . . .	36
3.2	Labeling system for event prediction based on alert horizon . . . . .	39
3.3	Hyperparameter Search Space . . . . .	47
4.1	Final metrics reported on the selected patients using ML on ECG signal. . . . .	50
4.2	Final metrics reported on the selected patients using an LSTM neural network on ECG signal. . . . .	51
4.3	Final metrics reported on the selected patients using ML on EEG signal. . . . .	52
4.4	Final metrics reported on patient 00 using NN on EEG signal. . . . .	53
4.5	Final metrics reported on patient 06 using NN on EEG signal. Note: FA = False Alarm. . . . .	53
4.6	Final metrics reported on patient 10 using NN on EEG signal. . . . .	54
4.7	Summary of the final metrics reported on each patient with the best NNs on EEG signal. . . . .	54
4.8	Final metrics reported on general dataset using ML on ECG signal. . . . .	55
4.9	Final metrics reported on general dataset using NN on ECG signal. . . . .	55
4.10	Final metrics reported on general dataset using ML on EEG signal. . . . .	56
4.11	Final metrics reported on general dataset using NN on EEG signal. . . . .	57





*"Insanity is doing the same thing over and over again  
and expecting different results."*

**Commonly attributed to Albert Einstein**

# Chapter 1

## Introduction

This section is designed to provide readers with the foundational knowledge necessary to fully comprehend the problem addressed in this thesis work and the system developed to solve it. Starting with a medical overview of the key themes involved, it proceeds with a description of the physiological signals utilized. Finally, it concludes with a general explanation of machine learning and deep learning techniques, with a particular emphasis on the systems used in the presented solution.

### 1.1 Nervous System

The nervous system is one of the most complex and developed systems within the human body, playing a crucial role in the acquisition, transmission, and processing of all endogenous and exogenous stimuli[27].

In humans, it is primarily divided into the Central Nervous System (CNS), responsible for processing and interpreting sensory information, and the Peripheral Nervous System (PNS), which serves as the communication lines that carry signals between the CNS and the rest of the body. The CNS is divided into an anterior part, the brain, and a posterior part, the spinal cord. The PNS is mainly composed by nerves, anatomical structures formed by long fibers called axons with a terminal part called dendrites; they are divided in cranial and spinal nerves and create an interconnected network that links the body periphery with its central parts[28].

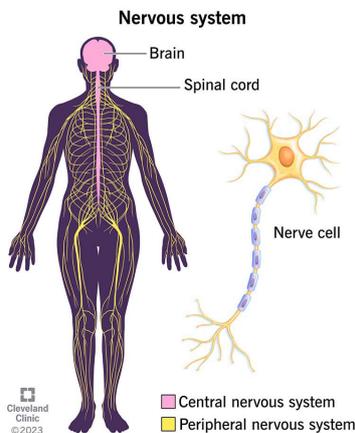


Figure 1.1: Nervous System anatomy. *Source:www.my.clevelandclinic.org.*

The PNS is further divided into the somatic nervous system, which controls voluntary movements and sensory input, and the autonomic nervous system, which regulates and controls involuntary body functions. To better cope with the overall complexity and diversity, the autonomic system is composed of two complementary branches: the sympathetic and parasympathetic systems. These systems are the result of the evolutionary mechanism of the fight-or-flight response, with the sympathetic part preparing the body to react to imminent danger by increasing readiness and activity levels, and the parasympathetic one focusing on conserving resources, restoring and maintaining normal activity levels. [11, 10, 21]. A summary diagram is shown in Figure 1.2.

As previously mentioned, nerves are essential for signal transmission within the body. They are divided into afferent sensory nerves, which transport sensory information from the periphery to the central nervous system, and efferent motor nerves, which convey movement commands from the brain and spinal cord to the periphery [29]. Interneurons, the last key category of neurons, are responsible for integrating and processing information between sensory and motor neurons, playing critical roles in reflexes and more complex neural functions [26].

Relatively to epilepsy, the main nerve involved is the vagus nerve, one of the most complex cranial nerves that is represented in figure 1.3. It runs from the brainstem to the abdomen, passing through the neck and chest. It is a mixed

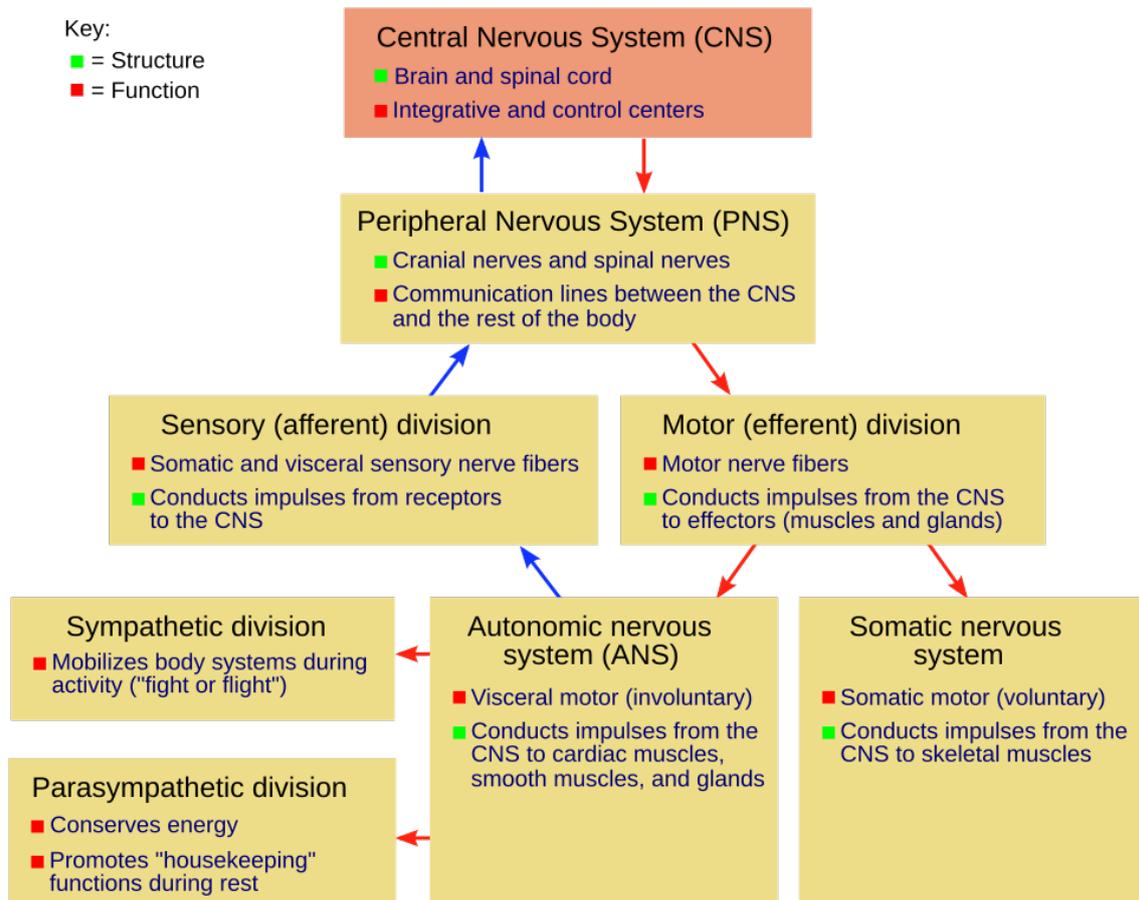
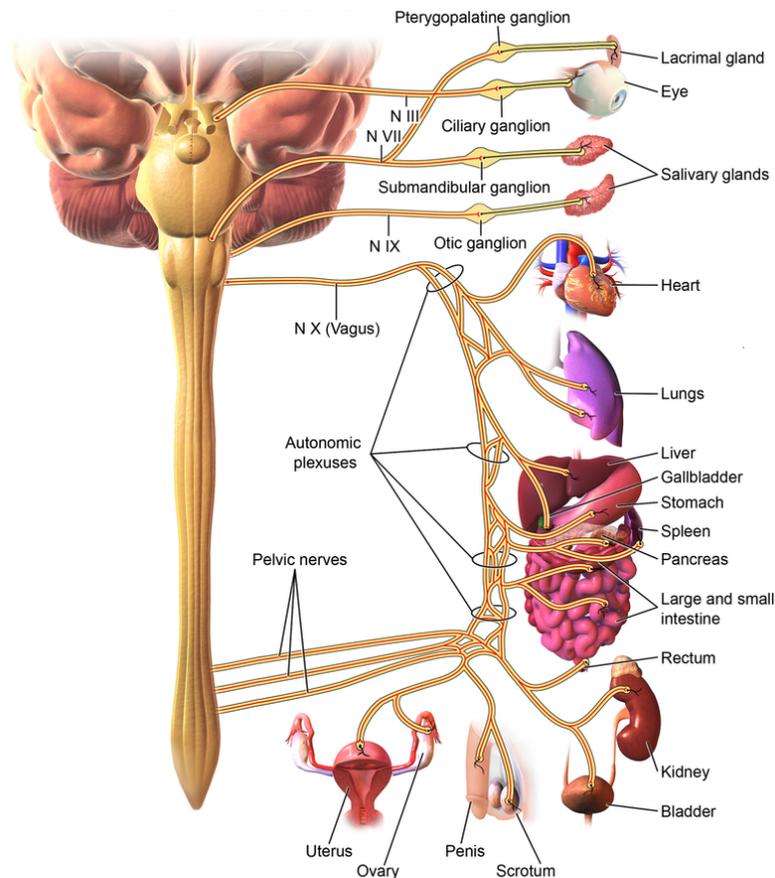


Figure 1.2: Nervous System diagram.

Source: <https://commons.wikimedia.org/wiki/File:NSdiagram.png>.

nerve, containing both sensory and motor fibers, with parasympathetic functions. The vagus nerve contains two sensory ganglia, the superior and inferior ganglia, from which many different branches originate, each controlling various functions. Among these branches is the cardiac branch, which plays a crucial role in regulating and slowing the heart rate. As will be covered in a later section, this nerve is fundamental for certain cardiac variations that can be used to predict an imminent epileptic seizure[9].



## Parasympathetic Innervation

Figure 1.3: Anatomy of the parasympathetic innervation.

Source: [https://me-pedia.org/wiki/Vagus\\_nerve](https://me-pedia.org/wiki/Vagus_nerve).

The CNS is composed of the brain and the spinal cord, as previously mentioned. While the spinal cord, which consists of a mix of gray and white matter, serves as a major signal pathway and plays a role in reflex responses (e.g withdrawal reflex) [8], most information processing, storage and the management of cognitive and emotional functions are controlled by the brain.

The brain is composed of the cerebrum, cerebellum, and brainstem, each with its own specific function.

- **Cerebrum:** makes up most of the brain volume and is divided into the right and left hemispheres, which are interconnected by the corpus callosum. It manages higher-level functions such as sensory interpretation, speech, motor control and reasoning.
- **Cerebellum:** is located below the cerebrum and is involved in learning and coordinating muscle activity and maintaining balance.
- **Brainstem:** acts as a mediator between the upper regions of the brain and the spinal cord and is responsible for regulating many automatic functions such as breathing, heart rate, body temperature and sleep-wake cycles.

The two sides of the cerebrum exhibit a contralateral approach to controlling the body, meaning the left hemisphere controls the right side of the body, and vice versa. This peculiarity provides important clues in cases of neurological issues, as the patient's symptoms reflect the location of the injury or lesion. To better describe its complex structure, the cerebrum is divided into four lobes: frontal, temporal, parietal and occipital. Each of these lobes is further divided into specific functional areas, where there is a coherence between the functions managed by that particular zone. However, no area or lobe operates independently; they are all deeply interconnected through physical, temporal, and functional connectivity.

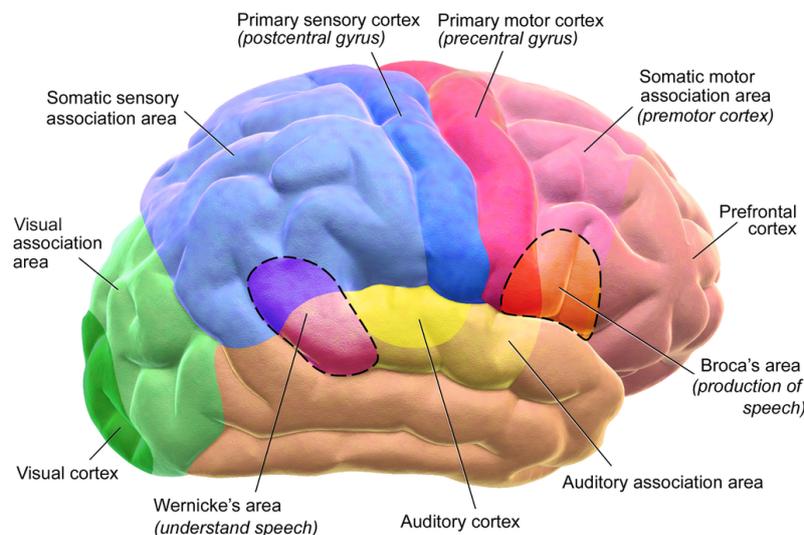


Figure 1.4: Brain functional map.

Source: [https://en.wikipedia.org/wiki/Human\\_brain](https://en.wikipedia.org/wiki/Human_brain).

Independently of the lobe and areas, the cerebrum is primarily composed of gray and white matter. Gray matter mainly consists of dendrites, somas, glial cells and capillaries, forming the outer part of the cerebrum known as the cortex. The gray matter is connected to the white matter, located deeper within the cerebrum via cortical columns, with each layer of the cortex having its own function. White matter is primarily composed of myelinated axons, glial cells and blood vessels.

## 1.2 Epilepsy

Epilepsy is an umbrella term that comprises a broad group of chronic, non-communicable disorders with diverse etiologies, presentations and outcomes [33]. They all share a common origin in surges of hypersynchronized electrical activity involving part or the entirety of the brain, causing a range of effects on the patient ranging from brief absences to tonic-clonic seizures involving the whole body. With 50 million people worldwide having epilepsy, it is one of the most common neurological diseases globally, with one-third of patients having drug and surgery-resistant seizures. Each year, 5 million new diagnoses are made, with an uneven distribution mainly affecting children and people over 65 years old [23, 16].

### 1.2.1 Diagnoses and consequences

It is important to note that having a seizure does not automatically translate into an epilepsy diagnosis, as epilepsy is diagnosed if a patient experiences more than one unprovoked seizure separated by at least 24 hours [34]. Various factors, including genetics, brain injuries and developmental disorders, can contribute to the onset of epilepsy.

Seizures can lead to various complications: the most immediate are related to trauma, drowning, suffocation and direct consequences of uncontrolled seizures in an hazardous environment. Additionally, there are consequences for sleep quality, memory and it can cause emotional issues that may lead to anxiety, depression and suicidal thoughts. Finally, less common complications include status epilepticus, a condition in which a seizure lasts more than 5 minutes or is characterized by multiple successive seizures without the patient regaining full consciousness. In this case, immediate medical assistance is required to avoid or reduce the risk of permanent brain damage and death. Another uncommon but serious condition is sudden unexpected death in epilepsy (SUDEP), a completely unexpected and still unexplained death that occurs in some cases, possibly due to cardiac or respiratory conditions. Because of this, the risk of premature death for a person with epilepsy is three times higher than for the general population.

The condition can significantly impact the quality of life of affected individuals, as seizures are usually unpredictable and leave the patient disoriented when they end. Due to the severity of some cases, the diagnosis can also lead to the revocation of driving privileges and other safety-related restrictions, further impacting the quality of life of patients with this diagnosis.

### 1.2.2 Classification

The standard reference guidelines for diagnosing epilepsy are provided by the International League Against Epilepsy and were reviewed in 2017. There are now three diagnostic levels to include in the diagnosis: seizure type, epilepsy type and epilepsy syndrome, with an emphasis on considering etiology and comorbidities at each level [33]. The complete diagram is shown in Figure 1.5, and each point is explained below [22]:

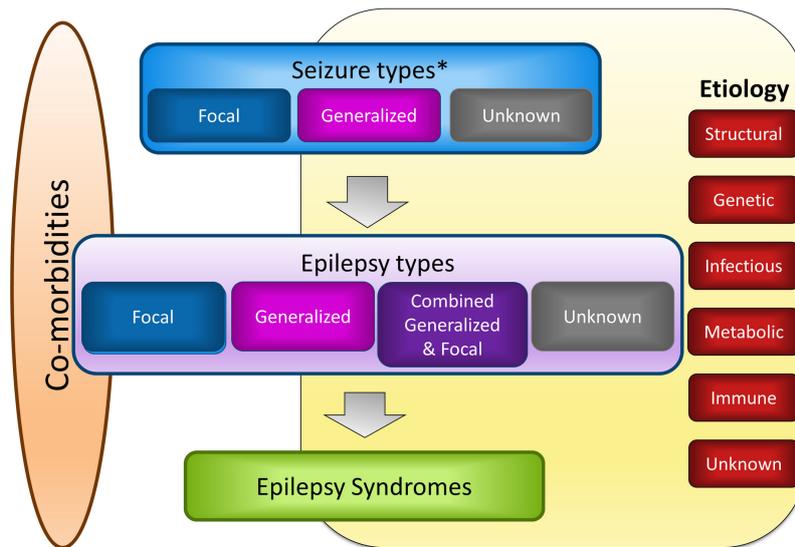


Figure 1.5: Epilepsy classification diagram.

Source: <https://onlinelibrary.wiley.com/doi/10.1111/epi.13709>.

- **Seizure type:** This is the starting point, based on the assumption that the clinician has already formulated the epilepsy diagnosis and is now trying to classify it with more precision. Seizures can be classified as **focal**, where everything starts from a specific zone in the brain and the symptoms reflect the affected area; **generalized**, where the whole brain is involved from the start of the seizure; or **unknown**, when there are no clear clues about the origin of the seizures. Without further information from EEG, video, and other analyses, this might be the only achievable level of classification.
- **Epilepsy type:** This classification is based on the analysis of EEG, focusing on the presence of specific waveforms (such as spike-wave) and their physical extent in the brain. It is also based on patient symptoms, forming a clinical basis for the diagnosis.
- **Epilepsy syndrome:** This refers to a cluster of features that include seizure types, EEG and imaging characteristics, which usually appear together. It can also include seizure triggers, diurnal variation, prognosis, and comorbidities, such as functional impairments, commonly reflected in EEG and other study results.

As clearly shown in Figure 1.5, there are many possible causes behind epilepsy, reflecting the complexity of this disorder. It should be noted that while many cases may have a genetic cause or result from trauma that alters brain structure, many diagnoses still lack a clear origin (between 30-40%), leaving their etiology unknown.

### 1.2.3 Treatment

The majority of epilepsy patients can be treated to achieve a seizure-free life, and in some cases, complete remission is possible. The first approach involves personalized medicine. There are many different anti-seizure medications (e.g., Carbamazepine, Levetiracetam, Valproate), with the right combination that has to be tailored to each patient. Nonetheless, about one-third of patients will not be able to control the disease with drugs alone. Another viable solution is surgery, where the aim is to modify the part of the brain causing the problem and its connections (e.g. resective surgery, laser ablation). However, the risks associated with such operations, issues of accessibility and the possible lack of an epileptogenic focus mean that surgery may not be feasible for every patient. Finally, another opportunity could be the vagus nerve stimulation, VNS. This solution is suitable for patient in whom surgery is not a viable option and can lead to a reduction in the severity, length and frequency of seizures, although it cannot stop them completely. This method is based on a subcutaneous stimulator placed in the patient's chest which sends regular electrical stimulations through the vagus nerve. An example of pattern is 30 seconds of stimulation every 5 minutes but can also be manually activated using a magnetic switch if the patient feels an incoming seizure. [24]

For these reasons, many patients are still looking for a real solution that grants them a better quality of life: a seizure prevention system could be a major advancement, particularly if the system is portable, does not require clinician supervision and uses minimal hardware to acquire the signal.

### 1.3 EEG and ECG

Signal transmission within the nervous system relies on a biochemical process known as the action potential. In a resting state, a neuron is negatively charged relative to the extracellular fluid, with a typical membrane potential around  $-70$  mV. This is due to an ionic imbalance, characterized by a higher concentration of sodium  $Na^+$  along with a high concentration of chloride ions  $Cl^-$  outside the cell and a lower concentration of potassium  $K^+$  along with an higher concentration of organic anions inside the cell. According to the Nernst equation, this ionic distribution generates the negative membrane potential and this equilibrium is dynamically maintained by voltage-dependent ion channels, which adapt their permeability based on the membrane potential. [6]

However, when a signal is generated or propagated, sensory receptors partially open specific protein channels leading to a substantial sodium entrance that increases the membrane voltage. If this increase surpasses a critical threshold (approximately  $-55$  mV), the proper action potential is triggered. During the initial depolarization phase, additional sodium voltage-dependent channels open, leading to a rapid increase in membrane potential until it reaches approximately  $+35$  mV, simultaneously activating adjacent regions of the membrane. After this peak, slower potassium channels open, and along with the closing of sodium channels, this initiates the repolarization phase, where the membrane potential begins to return to its original negative value. Due to the inertia of these ion channels, after the membrane potential is restored, it briefly overshoots and becomes more negative during the hyperpolarization phase. This phase is terminated by the sodium-potassium pump, which finally restores the membrane potential to its resting level and allow for the possibility of generating a new action potential. This occurs after an absolute refractory period, during which the cell cannot respond to any stimulus, regardless of its intensity. [30].

This biochemical activity can be detected using specific hardware first developed in the early 20th century. Hans Berger invented the electroencephalogram (EEG) in the 1920s, while Willem Einthoven refined the electrocardiogram (ECG) at the beginning of the 20th century starting from earlier discoveries of bioelectric phenomena. These tools enabled the first detailed recordings of the electrical signals generated by neurons and cardiac muscle, introducing new possibilities in clinical routine. By capturing the rapid ion exchanges that drive action potentials, these early technologies laid the groundwork for modern neurophysiology and cardiology. In simpler forms, this is made possible by external, cutaneous electrodes that act as transducers between two worlds: ion-based physiology and electron-based electronic hardware. Ion fluxes cause interactions, and in some cases, exchanges, on the surface of the electrode, altering its potential. This potential is then processed and amplified using a specific hardware chain, finally being converted into a numerical value.

Even if the hardware acquiring signals share the same working principle, each signals has its own different characteristic, also because of the anatomical structures surrounding the zone of interest. In particular, amplitude and frequency content can change as shown below:

Technique	Amplitude range	Frequency range
ECG	0.5 - 4 mV	0.01 - 250 Hz
EEG	5 - 300 $\mu$ V	0.1 - 150 Hz
EMG	50 - 3000 $\mu$ V	0.1 - 300 Hz

Table 1.1: Comparison of amplitude and frequency ranges for ECG, EEG, and EMG techniques.

### 1.3.1 Electroencephalography (EEG)

This method focuses on the brain activity. It can be either invasive, with electrodes placed inside the skull or deeper within the brain tissue, or non-invasive, using small electrodes placed in a cap worn by the patient on the scalp. Clearly, the best signal quality is obtained with the invasive method, as there is no tissue interference with the readings and the distance from the source is smaller; however, this approach is used only in rare and specific research cases due to the problems related with it. Most of the time, the non-invasive method provides sufficient data to study brain activity, despite higher level of signal noise: this is also the case for the signals used in this project.

The signal acquired in EEG recordings is the postsynaptic potential, reflecting membrane potential changes at a chemical synapse caused by neurotransmitters released by the presynaptic neuron, which bind to specific receptors on the postsynaptic terminal. As a result, the signal generated lasts longer and can be temporally and spatially summed, providing a better overall reading. Action potentials, on the other hand, are too fast to be reliably captured in a stable manner.

These signals are usually acquired by 8-16 pairs of electrodes attached to a specific cap, following the 10-20 international standard shown in figure 1.6, where electrodes are positioned relative to two reference points (Nasion and Inion). The resulting signal is typically calculated as a single differential between pairs of electrodes or as a monopolar signal using an external reference point, usually placed on the ear or as the mean value of all recording channels. [7, 25]

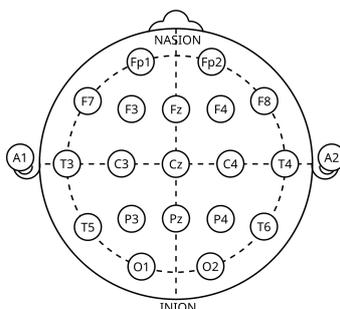


Figure 1.6: 10-20 international standard map.

Source: <https://en.wikipedia.org/wiki/Electroencephalography>.

The signal acquired can then be processed dividing its frequency content in different bands, each one with a different physiological meaning.

- **Delta ( $\delta$ ) waves:**

- Frequency range: 0.5 – 4 Hz
- Characteristics: Delta waves are the slowest EEG waves and are typically observed during deep sleep. They are also associated with unconscious processes and certain pathological conditions, such as brain injuries.

- **Theta ( $\theta$ ) waves:**

- Frequency range: 4 – 8 Hz
- Characteristics: Theta waves are commonly associated with light sleep, relaxation, and meditation. They are also related to memory processes, creativity and the early stages of sleep.

- **Alpha ( $\alpha$ ) waves:**

- Frequency range: 8 – 13 Hz
- Characteristics: Alpha waves are predominantly seen in relaxed, awake states, especially when the eyes are closed. They are linked to a state of relaxed alertness and are most prominent in the occipital and parietal regions of the brain.

- **Beta ( $\beta$ ) waves:**

- Frequency range: 13 – 30 Hz
- Characteristics: Beta waves are associated with active thinking, focus and concentration. They are prominent during wakefulness, especially when a person is alert and engaged in cognitive tasks.

- **Gamma ( $\gamma$ ) waves:**

- Frequency range: 30 – 100 Hz
- Characteristics: Gamma waves are associated with high-level cognitive functions, such as perception, consciousness and binding of sensory inputs. They are thought to play a role in the integration of information across different parts of the brain.

In an epileptic center, the neurons studied by the EEG display high membrane potential instability, with excessive and prolonged depolarization known as paroxysmal depolarizing shifts. Even during the interictal phase, the EEG shows some atypical waves with abnormal spikes. However, these become problematic and initiate a seizure only when a sufficient number of neurons are involved, allowing the pathological activity to propagate and spread to neighboring regions. In that case, the abnormal discharging activity started by abnormal neurons diffuses and entrain other neurons until a critical mass is reached, which leads to the start of a seizure. However, it's important to note that this critical is not highly localized but appears over different time and spatial scales and has a non-linear nature, that is the reason why predictions are difficult. The termination of the seizure is attributed to the activation of inhibitory mechanisms based on GABA, both cortical and extracortical.

Overall, some examples of epileptic EEG is shown below in figure 1.7 and figure 1.8

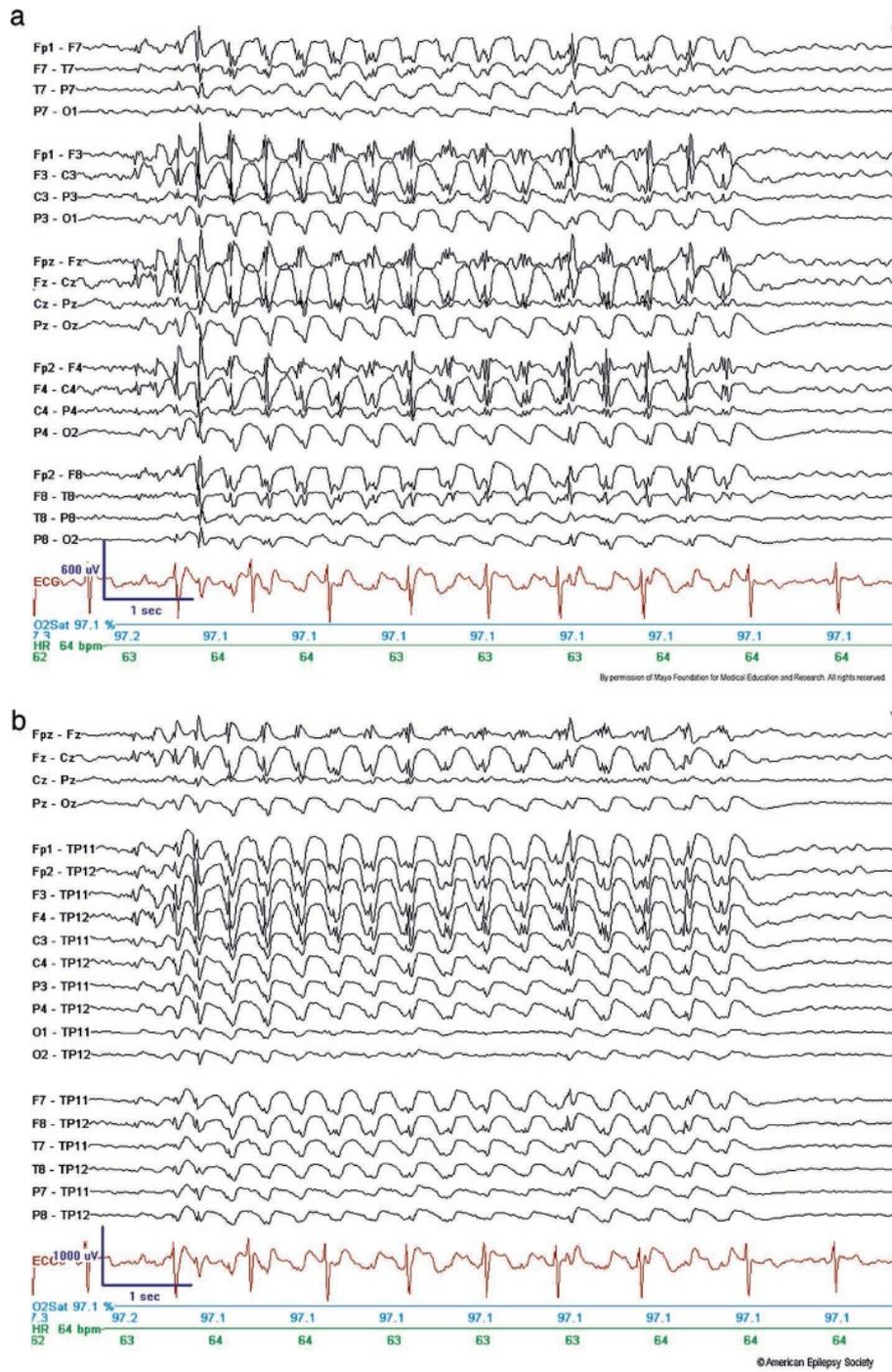


Figure 1.7: Example of 3-Hz (typical) generalized spike-wave IED.

Source: <https://www.ncbi.nlm.nih.gov/books/NBK390347/>

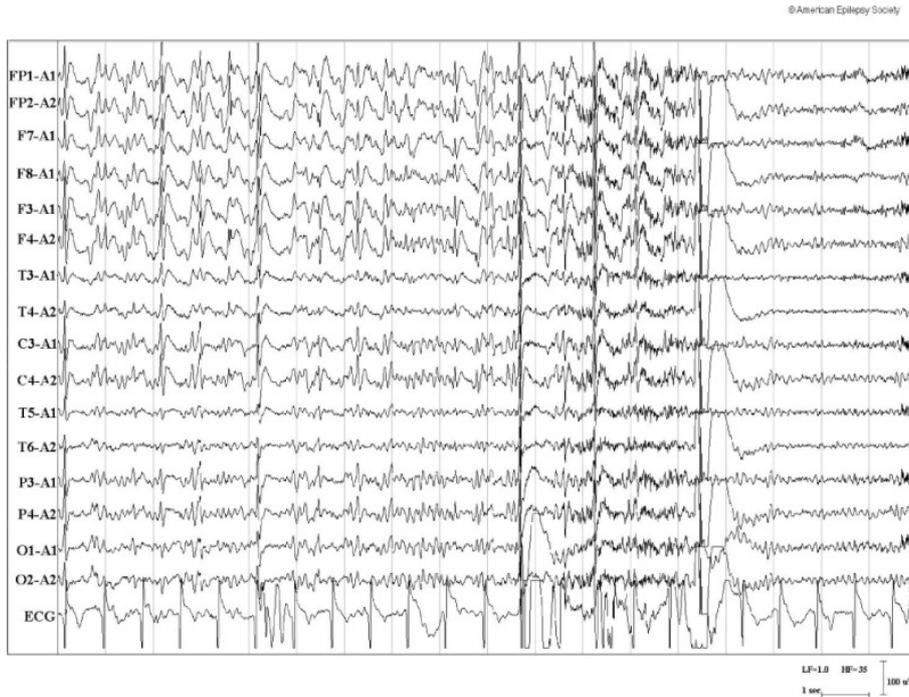


Figure 1.8: Example of a patient who had a generalized axial myoclonic jerk during second 14, which coincided with the generalized spike-wave discharge and electrodecremental pattern seen on EEG.

Source: <https://www.ncbi.nlm.nih.gov/books/NBK390347/>

### 1.3.2 Electrocardiography (ECG)

This section focuses on the electrical activity generated by heart tissues during their functioning. Due to the complex pattern of electrical signal propagation through the heart tissue, a comprehensive study of its path and characteristics usually requires 12 different leads divided into limb leads and precordial leads, as shown in figure 1.9.

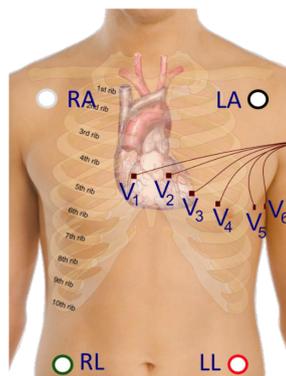


Figure 1.9: 12 leads electrodes placement.

Source: Leutheuser, Heike. (2019). *Wearable computing applications in eHealth*.

The first group is subdivided into standard bipolar and augmented unipolar leads, while the others are simply named V1 to V6. Limb leads provide information in the vertical plane, while the precordial leads offer insight into the horizontal plane.

The combination of all these data provides a precise picture of heart conditions, including not only the heart rate but also its morphology, potential conduction blocks, and other issues that may cause deviations from the typical ECG shape shown in figure 1.10, whether due to amplitude differences or delays.

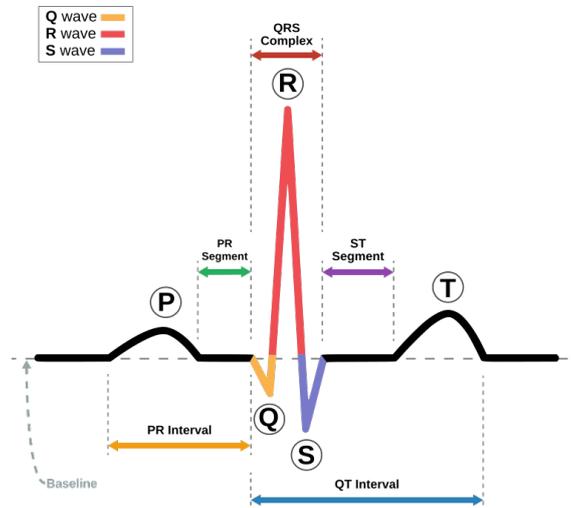


Figure 1.10: Physiological ECG shape.

Source: <https://en.wikipedia.org/wiki/Electrocardiography>

Some key information about the heart condition, however, can still be extracted with a lower number of leads. Even with the loss of different perspectives for analyzing the heart, information about heart rate will still be available. Among the various types of information available, Heart Rate Variability (HRV), also known as R-R variability, is one of the most important: this parameter measures the time difference between successive heartbeats, as shown in figure 1.11. High variability between heartbeats is a sign of good cardiovascular and nervous system condition, as it reflects the body readiness to respond to changes in condition, with a balanced interaction between the sympathetic and parasympathetic systems (through the vagus nerve). [4]

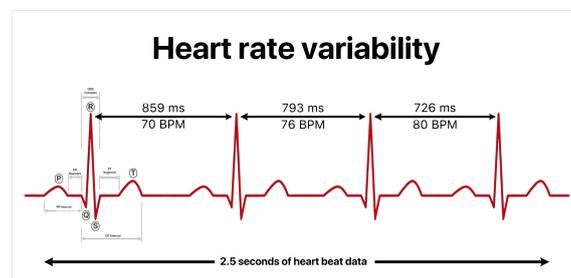


Figure 1.11: HRV calculated on an ECG signal.

Source: [https://en.wikipedia.org/wiki/Heart\\_rate\\_variability](https://en.wikipedia.org/wiki/Heart_rate_variability)

## 1.4 Machine Learning

Machine learning (ML) is a branch of artificial intelligence methods that enables systems to learn from data and make predictions without explicit programming using thresholds or other methods. In biomedical applications, ML is widely used for physiological signal analysis, classification and early event detection, making it a valuable tool for seizure prediction. By analyzing EEG and ECG data, ML models can identify patterns that precede seizure onset, allowing for real-time intervention strategies.

In this work, multiple supervised ML classifiers were explored to determine the most effective approach for seizure prediction. The following classifiers were implemented and tested:

- **Support Vector Machine (SVM)** – A supervised learning algorithm that finds an optimal hyperplane to separate different classes. Given a dataset  $\{(\mathbf{x}_i, y_i)\}_{i=1}^n$ , SVM solves:

$$\min_{\mathbf{w}, b} \frac{1}{2} \|\mathbf{w}\|^2 \quad \text{subject to} \quad y_i(\mathbf{w} \cdot \mathbf{x}_i + b) \geq 1, \forall i.$$

- **Linear Discriminant Analysis (LDA)** – A statistical method that projects data into a lower-dimensional space, maximizing class separation while assuming normally distributed features.
- **Quadratic Discriminant Analysis (QDA)** – An extension of LDA that allows for non-linear decision boundaries by considering class-specific covariance matrices.
- **Decision Tree** – A hierarchical structure that recursively splits data based on feature values to form decision rules.
- **Random Forest** – An ensemble of decision trees, where each tree is trained on a random subset of data. The final classification is determined by majority voting among the trees.
- **Gradient Boosting** – An iterative boosting technique that corrects errors by sequentially training weak classifiers on residual errors.
- **k-Nearest Neighbors (kNN)** – A non-parametric classifier that assigns a class label based on the majority class among the  $k$  nearest data points.
- **Logistic Regression** – A probabilistic model that predicts class membership using the sigmoid function:

$$P(y = 1|\mathbf{x}) = \frac{1}{1 + e^{-(\mathbf{w} \cdot \mathbf{x} + b)}}.$$

- **Naïve Bayes** – A probabilistic classifier based on Bayes' theorem, assuming feature independence.

## 1.5 Neural Networks

Neural Networks (NNs) are a class of machine learning models inspired by the biological structure and function of the human brain. They consist of multiple interconnected layers of artificial neurons that process information in a hierarchical manner. Each neuron performs a weighted sum of its inputs, applies an activation function, and passes the output to the next layer. This layered structure allows NNs to learn complex, nonlinear patterns in data, making them highly effective for biomedical applications such as EEG and ECG signal analysis.

The advantage of NNs over traditional machine learning methods lies in their ability to learn complex representations and patterns, also leveraging on improved feature extraction. The disadvantage, however, lies in the larger number of data that are needed to correctly train the system.

In this work, four different NN architectures were implemented to process EEG and ECG signals and classify different time windows as preictal (before seizure) or interictal (normal state).

### 1.5.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks are a special type of Recurrent Neural Network (RNN) designed to capture long-term dependencies in sequential data. Unlike traditional RNNs, which suffer from the vanishing gradient problem, LSTMs include memory cells that selectively store and retrieve information over extended periods. This ability makes them particularly useful for analyzing time-series data such as EEG and ECG, where past observations influence future predictions.

LSTMs achieve this by using three types of gates:

- **Forget gate**  $f_t$ : Decides what information should be discarded from the memory cell based on the current input  $x_t$  and previous hidden state  $h_{t-1}$ .
- **Input gate**  $i_t$ : Determines which new information should be stored in the memory cell.
- **Output gate**  $o_t$ : Regulates the amount of information from the memory cell that should be passed to the next time step.

These gates operate as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (\text{forget gate}) \quad (1.1)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (\text{input gate}) \quad (1.2)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \quad (\text{cell state update}) \quad (1.3)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (\text{output gate}) \quad (1.4)$$

$$h_t = o_t \cdot \tanh(C_t) \quad (\text{hidden state}) \quad (1.5)$$

The LSTM model used in this study first applies time-distributed dense layers to extract relevant features from the input signal before passing them to the LSTM layers. The final classification is performed by fully connected layers, with a softmax activation function at the output.

## 1.5.2 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) are a deep learning architecture originally designed for image processing but have proven effective in time-series classification as well. CNNs leverage spatial hierarchies by applying convolutional filters that capture local patterns in data.

The main components of a CNN are:

- **Convolutional layers:** Apply filters to the input, detecting features such as frequency patterns in EEG signals.
- **Batch Normalization:** Stabilizes training by normalizing activations within each batch.
- **Max-pooling layers:** Reduce dimensionality while preserving essential features by taking the maximum value from local regions.
- **Fully connected layers:** Aggregate extracted features and perform classification.

The convolution operation is defined as:

$$y_t = \sum_{i=0}^{k-1} x_{t+i} \cdot w_i + b,$$

where  $k$  is the kernel size,  $w_i$  are filter weights, and  $b$  is the bias term.

By applying multiple layers of convolutions, CNNs progressively learn higher-level abstractions of EEG and ECG data, making them useful for seizure detection.

## 1.5.3 Recurrent Neural Networks with Gated Recurrent Units (RNN-GRU)

Gated Recurrent Units (GRUs) are an alternative to LSTMs that simplify the recurrent structure while retaining the ability to capture long-term dependencies. GRUs remove the memory cell and instead use two gates:

- **Update gate  $z_t$ :** Controls how much past information should be carried forward to the next time step.
- **Reset gate  $r_t$ :** Determines how much past information should be forgotten.

These operations are defined as:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (\text{update gate}) \quad (1.6)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (\text{reset gate}) \quad (1.7)$$

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h) \quad (\text{candidate activation}) \quad (1.8)$$

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (\text{final activation}) \quad (1.9)$$

GRUs are computationally more efficient than LSTMs since they have fewer parameters. In this study, a GRU model was used to analyze EEG and ECG signals, taking advantage of its ability to model temporal dependencies while maintaining fast training times.

### 1.5.4 Temporal Convolutional Networks (TCN)

Temporal Convolutional Networks (TCNs) are an alternative to RNNs for processing sequential data. Instead of relying on recurrent connections, TCNs use 1D convolutional layers with dilation to capture long-term dependencies while enabling parallel processing.

The key features of TCNs include:

- **Causal convolutions:** Ensure that the model only uses past and present information, preventing data leakage from future time steps.
- **Dilation:** Expands the receptive field of the convolution, allowing the network to learn long-term dependencies.
- **Residual connections:** Improve gradient flow and make deeper networks easier to train.

Unlike LSTMs and GRUs, which process sequences step-by-step, TCNs apply filters across entire sequences simultaneously. This allows for more efficient computation, making TCNs well-suited for real-time applications.

## 1.6 Neural Network Intelligence (NNI)

A major challenge in deep learning is hyperparameter tuning, which traditionally requires extensive manual experimentation or inefficient grid search methods. Neural Network Intelligence (NNI) automates this process by intelligently exploring the hyperparameter space. Instead of performing exhaustive searches, NNI employs adaptive techniques such as Bayesian optimization and evolutionary algorithms to converge on optimal configurations, discarding combinations not leading to acceptable results. This significantly reduces training time while improving model accuracy and robustness. In this project, however, NNI has a larger adoption because it was used to tune not only the NN and ML hyperparameters, but also many different preprocessing parameters were treated as hyperparameters and so became object of optimisation. By leveraging NNI, the need for human intervention was minimized, ensuring that hyperparameter choices were driven purely by data and final results rather than subjective intuition or computational constraints.



# Chapter 2

## State of the art

Over the past decades, there has been growing interest in the prediction of seizures using physiological signals. This trend has been driven by advancements in informatics and the increased availability of powerful hardware capable of running more complex algorithms, including deep neural networks. Early solutions were based on simple machine learning algorithms, such as Support Vector Machines (SVM), or even simpler threshold methods that relied on specific features extracted from the signals and compared them with patient tuned thresholds. Neural networks have only started to gain attention in more recent studies, as advances in hardware and dedicated coding frameworks have made them more accessible for applied research.

### 2.1 EEG based systems

This field of research started in the 1980s [15] thanks to major mathematical discoveries which led to advanced parameter calculations from EEG signals. This research domain saw a significant increase at the start of the new millennium. During this period, many interesting papers emerged claiming exceptional results by using various features calculated upon a moving time window, as summarized in [18]. The approach underlying these systems extends beyond the introduction of new metrics for signal characterization and is based on the hypothesis that epileptic seizures do not emerge abruptly without warning. Rather, a preictal state exists, representing a gradual transition from normal physiological conditions to seizure onset. However, a milestone review published in 2007 highlighted excessive optimism in these earlier results, as most of them were not replicable by other research groups and presented methodological flaws that falsely inflated the final metrics presented. Moreover, this paper suggests a common ground for future system metrics evaluation, presenting for example the time under false alarm as an important evaluation metric. [20] Aided by advances in both computational hardware and code, research resumed from these new foundations, combining updated strategies, parameters [1], and technologies. There was increasing reliance on machine learning and deep learning to improve results, achieving sensitivity rates ranging from 70% to 90% and time under false alarm up to 30%, with a prediction window typically ranging from 10 to 20 minutes. [19] [36] [35][31]. In these cases, the parameters used vary widely.

For univariate parameters, i.e. values calculated from single channels, these include:

- **Statistical moments of the signal amplitude:**
- **Power spectral parameters and power band ratios**
- **Correlation dimension and density**
- **Signal entropy**

As for bivariate features, i.e. parameters calculated across different channels, the most commonly used ones are:

- **Maximum linear cross-correlation:**
- **Autoregressive measure of synchrony**
- **Phase synchronization**

In other cases, convolutional networks were used to bypass manual feature selection and were directly used simply upon the filtered signal, and the final classification was performed either by the network's final layer or by an SVM utilizing the output from the features selected by the convolutional layers. [31]. The most significant recent advancement in this field is the ability to generate seizure forecasts with an alert span of several days, enabling predictions across different temporal scales. Currently, research is divided between the older deterministic approach, which offers shorter alert periods, and the newer probabilistic approach, which provides a longer prediction window but only offers a probability, not a definitive state. Moreover, this approach must be trained on continuous data spanning several months in order to follow physiological daily and multidaily trends on which these systems are based. Despite the longer alert window, these systems are more invasive as are usually based on intracortical electrodes and can be more confusing for patients as the output is a probability instead of a defined value. They may also have a psychological impact due to the extended prediction window, which still needs to be tested on subjects[5] [2] [3]. However, this long-term approach falls outside the scope of this thesis and is referenced solely to provide a comprehensive overview of the state of the art in this field. In both cases, still, the generalization of these systems remains problematic as there is no single parameter or value that can be simply universally applied to all patients and their variability. Therefore, each system typically requires some degree of personalization during the training phase, to varying extents. An example can involve a simpler fine-tuning to adapt a pre-trained model to a specific patient.[17]

While recent studies have demonstrated significant advancements in seizure prediction using artificial intelligence, a fundamental challenge remains in the preparation of datasets and the rigorous partitioning of training and validation data. A critical issue is the prevention of data leakage, which occurs when information from the validation set is inadvertently introduced into the training phase, leading to over-estimated model performance and a lack of generalizability in real-world scenarios. This leakage can be introduced through various mechanisms, including incorrect normalization strategies where statistics are computed over the entire dataset rather than restricted to the training set alone. A major concern highlighted in recent

research [32] is that many high-accuracy seizure prediction results may be driven by models learning spurious correlations or uninformative patterns rather than physiologically meaningful seizure precursors. In particular, the study emphasizes that EEG signals exhibit temporal correlations, meaning that samples close in time are more similar than those further apart. This phenomenon poses a risk when EEG windows from adjacent time periods are assigned to both training and validation sets, as the model may learn to classify subtle non-physiological features, such as session-specific noise, rather than genuine preictal biomarkers. Some studies [32] empirically demonstrated that noise within EEG recordings, including slow-varying background signals, can be exploited by machine learning models to achieve deceptively high classification performance. This confounding effect is particularly problematic in short-duration recordings and datasets with a limited number of samples, where deep learning algorithms may inadvertently learn noise patterns rather than universal seizure precursors. Furthermore, the study introduced a controlled experiment in which a neural network was trained on arbitrarily labeled EEG segments, completely decoupled from actual seizure events. The model still achieved high classification accuracy, demonstrating that non-seizure-related temporal variations could be sufficient to drive successful prediction. To mitigate these issues, the study proposed best practices for dataset preparation and model validation, including the use of patient-independent evaluation frameworks, rigorous cross-validation strategies and the adoption of a fully held-out test set. Importantly, they emphasized that validation procedures should prevent models from leveraging temporally adjacent EEG samples as a surrogate feature for seizure prediction. Additionally, another solution could employ adversarial validation techniques to assess whether models are learning physiologically meaningful patterns rather than noise-related artifacts. Without such rigorous methodological safeguards, seizure prediction models risk producing misleadingly high performance metrics that do not translate into clinically useful applications. Consequently, future research in this field must prioritize not only the development of accurate predictive algorithms but also the implementation of robust evaluation protocols to ensure that model performance is genuinely driven by seizure-related biomarkers. Addressing these challenges is essential for advancing AI-based seizure prediction toward reliable and clinically deployable solutions.

## 2.2 ECG based systems

To advance beyond classical EEG-based seizure prediction systems, recent studies have increasingly focused on the electrocardiogram (ECG) signal as a viable alternative for predicting seizures in epileptic patients. The primary advantage of this approach lies in the simpler hardware required for signal acquisition. Unlike EEG, which necessitates an array of electrodes distributed across the scalp or even worse intracranial electrodes, ECG monitoring can be performed using only a few electrodes placed on the patient chest.

It is important to note that in these studies, ECG data was acquired without employing a full 12-lead setup; instead, a minimal number of electrodes was used. This choice was driven by the primary objective of identifying individual heartbeats to derive features from the interbeat intervals rather than capturing the entire cardiac waveform. Specifically, the key metric of interest is heart rate variability (HRV), which quantifies the temporal fluctuations between consecutive heartbeats and includes additional parameters to enhance the characterization of cardiac dynamics. HRV analysis is typically conducted over a moving time window of approximately two minutes, enabling the extraction of time-domain, frequency-domain and non-linear features.

The most commonly computed time-domain indices include:

- **Mean NN:** Mean value of the interval between two successive heartbeats.
- **SD NN:** Standard deviation of the interval between two successive heartbeats.
- **RMSSD:** Root mean square of the differences between successive NN intervals.
- **NN50:** Number of adjacent NN interval pairs differing by more than 50 ms within the specified measurement period.

Similarly, the most frequently computed frequency-domain indices are:

- **Total Power:** Overall spectral power of HRV across all frequency bands.
- **Variance of NN Intervals:** Measure of the overall HRV magnitude.
- **Low-Frequency (LF) Power:** Spectral power within the 0.04–0.15 Hz range, associated with both sympathetic and parasympathetic modulation.
- **High-Frequency (HF) Power:** Spectral power within the 0.15–0.40 Hz range, primarily reflecting parasympathetic activity.
- **LF/HF Ratio:** Ratio of LF to HF power, often used as an index of autonomic balance.

Finally, non-linear HRV analysis provides additional insights into the complex dynamics of heart rate variability by capturing properties that are not evident in time or frequency-domain features. The most commonly employed non-linear metrics include:

- **SD1:** Standard deviation of Poincaré plot points perpendicular to the line of identity, reflecting short-term HRV.

- **SD2**: Standard deviation along the line of identity in the Poincaré plot, indicating long-term HRV.
- **CSI**: Cardiac Sympathetic Index, assessing sympathetic nervous system activity.
- **CVI**: Cardiac Vagal Index, related to parasympathetic modulation.
- **Katz FD**: Katz Fractal Dimension, a measure of the complexity of HRV.
- **Recurrence Plot Features**:
  - **Recurrence Rate**: Fraction of recurrent points in the phase space.
  - **Determinism**: Percentage of recurrence points forming diagonal structures, indicating predictability.
  - **Lmax**: Maximum diagonal line length, related to system stability.
  - **Laminarity**: Fraction of recurrence points forming vertical structures, linked to autonomic control.
  - **Trapping Time**: Average duration of laminar phases in the recurrence plot.
  - **Shannon Diagonal Entropy**: Entropy of the diagonal line distribution, representing signal complexity.

In some cases, additional parameters, such as entropy-based measures or higher-dimensional fractal metrics, may also be computed to further refine the characterization of HRV dynamics.

As previously noted, these HRV metrics are typically computed using a sliding window approach (that could be implemented using a First-In-First-Out (FIFO) buffer in real-time scenarios), which enhances measurement stability and reduces the influence of transient artifacts. To further improve the robustness of the extracted metrics, many studies also applied dedicated outlier removal techniques based on thresholding and interpolation strategies. For frequency-domain HRV parameters, signals are usually resampled to ensure uniform time intervals, and autoregressive models may be employed to extract spectral components not directly discernible from the raw data.



# Chapter 3

## Methods

The core objective of this project was to develop a comprehensive and adaptable toolbox capable of processing both EEG and ECG signals through a fully automated and generalizable pipeline. The system is designed to handle all stages of signal processing and machine learning/neural networks training with minimal user intervention, ensuring that key methodological choices are data-driven rather than dependent on subjective decisions or external constraints, such as excessive number of trials. This approach addresses one of the major limitations in traditional seizure prediction research, where many analytical pipelines require manual parameter tuning, often influenced by prior experience, computational restrictions, or practical time constraints, thereby limiting the true extent of the parameter search space and potentially neglecting key regions.

By minimizing the need for arbitrary a priori choices, the proposed system ensures that the selection of preprocessing parameters, feature extraction methods and classification algorithms is optimized systematically based on the final output results. This mitigates the risk of overlooking potentially significant parameter configurations that might otherwise be disregarded due to human bias or feasibility constraints. The automation of parameter selection also expands the explored search space, allowing for a more exhaustive evaluation of different signal processing and classification strategies, ultimately leading to improved performance and robustness of predictive models, despite the substantial inter-individual and intra-individual variability in physiological signals.

Beyond improving predictive accuracy, the automation of preprocessing and classification stages significantly reduces the workload required to implement and test different models. Traditionally, the design of seizure prediction systems requires extensive manual experimentation with various filtering techniques, feature engineering approaches and machine and deep learning architectures, often demanding extensive knowledge and substantial computational effort. By automating this process, the proposed framework accelerates the ideation, creation and testing of new methodologies and classifiers, enhancing both the quality of the achieved results and the efficiency of their development.

Additionally, the systematic nature of this pipeline reduces the risk of methodological inconsistencies that can arise from ad-hoc adjustments and manual fine-tuning. As mentioned earlier, many studies suffer from issues related to reproducibility and bias introduced by human intervention during model development. By enforcing a structured, automated workflow, the proposed toolbox enhances

reproducibility, ensuring that model performance is evaluated under standardized conditions and that results can be reliably compared across different datasets and studies.

### 3.1 Dataset

The whole system was trained and tested using the Siena Scalp EEG Database. This database consists of EEG recordings of 14 patients acquired at the Unit of Neurology and Neurophysiology of the University of Siena. Subjects include 9 males (ages 25-71) and 5 females (ages 20-58). Subjects were monitored with both cameras and EEG with a sampling rate of 512 Hz, electrodes were arranged on the basis of the international 10-20 System. Most of the recordings also contain 2 EKG signals. The diagnosis of epilepsy and the classification of seizures according to the criteria of the International League Against Epilepsy were performed by an expert clinician after a careful review of the clinical and electrophysiological data of each patient.[12][13][14].

The selection of this specific dataset was driven by the availability of both EEG and ECG signals, complemented by detailed metadata. This comprehensive data structure enables the utilization of both signal types and facilitates the implementation of fully automated functions, ensuring the level of automation required for the aim previously presented.

The preprocess and training pipelines were designed to be launched independently: this ensures a faster training section that can be run on different dataset versions that were precomputed using the different parameters, reducing redundant calculations and by so further improving the research speed.

To further optimize model performance, the preprocessing pipeline is designed to process each patient's data independently, ensuring that subject-specific characteristics are preserved. The decision to merge data from multiple patients or keep them separate is made during the training phase, maximizing dataset reusability while maintaining flexibility in cross-subject generalization and personalized model adaptation.

## 3.2 Preprocess

The preprocessing phase represents the first step of the system, during which raw data are automatically imported, verified and processed to generate a refined and structured dataset. This ensures that the data are optimized for classifier training, enhancing both signal quality and model performance. The preprocessing stage involves automated data and metadata import along with integrity checks to detect potential inconsistencies, followed by a series of transformations aimed at improving signal quality and features calculation.

As previously discussed, key preprocessing parameters are treated as tunable variables during each NNI experiments, reducing the number of fixed a priori assumptions and allowing for systematic optimization. This approach ensures that the final dataset configuration is not arbitrarily constrained, enabling the exploration of a broader search space and maximizing the likelihood of identifying optimal processing strategies.

### Import

The first function of the pipeline is responsible for correctly importing and structuring data, ensuring that signals and their corresponding metadata are properly organized for subsequent processing. Specifically, it reads EDF files containing EEG and ECG recordings while also extracting essential metadata from an external text file supplied in the dataset, since this information is not correctly embedded within the EDF files. The extracted metadata includes the recording start and stop times, the effective sampling rate and the seizure onset for each file.

During metadata extraction, minor inconsistencies were identified and manually corrected to maintain data integrity. For example, in some recordings, seizure onset times were erroneously reported as precedent to the start of the corresponding recording. These discrepancies were manually reviewed and adjusted based on plausible corrections, such as shifting the timestamp by a reasonable value in cases likely due to classical typographical errors.

At the end of this step, the system automatically flags any remaining inconsistencies, allowing the user to review potential errors and ensure data quality before proceeding. Given the importance of this stage, randomized manual verifications were also performed, confirming that the imported data accurately reflects the original values.

Once the signals are successfully imported and stored in the appropriate data structure, a channel-wise mean subtraction is applied to each signal. This step, performed before any downstream processing, removes DC offset and standardizes baseline levels, improving the robustness of subsequent analyses.

### 3.2.1 Filtering

Given the distinct nature of EEG and ECG signals this filtering procedure is applied separately to each modality, ensuring that each signal type is processed with optimal parameters. In both cases, the primary objective is to attenuate low-frequency noise caused by artifacts and baseline drift, suppress high-frequency noise originating from acquisition electronics and mitigate power line interference, which was observed to be particularly strong in these recordings.

To achieve this, a Chebyshev band-pass zero-phase filter is applied, implemented as a cascade of high-pass and low-pass filters. This configuration effectively isolates the relevant frequency bands while preserving phase characteristics, ensuring minimal waveform distortion. Following band-pass filtering, a notch filter is applied to suppress power line interference at the fundamental frequency and its harmonics. The specific filtering parameters for both EEG and ECG signals are reported in Table 3.1.

To further enhance filtering effectiveness, the system automatically adjusts the filter attenuation, starting from predefined parameters and incrementally increasing suppression levels while preserving signal stability. This approach led to an attenuation of 65 dB for both EEG and ECG signals, balancing effective noise suppression with signal integrity.

Finally, the filtering function includes a visualization module, allowing users to compare raw and filtered signals alongside their Power Spectral Density (PSD) representations shown in 3.1, 3.2.

<b>Parameter</b>	<b>EEG</b>	<b>ECG</b>
Cut-off frequency (Hz)	0.5 - 100	0.5 - 100
Notch (Hz)	50	50
Attenuation (dB)	65	65

Table 3.1: Filtering parameters

### 3.2.2 Signal labeling

The preprocessing pipeline proceeds by merging EEG and ECG data into a unified matrix, where ECG values are appended as additional rows. This structure enables the joint analysis of both signal types, leveraging their complementary information for seizure prediction. The resulting multimodal matrix is then passed to the labeling function, a critical component that directly influences the quality and interpretability of the final model predictions.

A key limitation in previous seizure prediction studies has been the lack of a well-defined, medically validated threshold for differentiating normal and preictal states. Traditionally, the alert horizon — the time window before a seizure in which the model should detect preictal patterns — was chosen arbitrarily, without a consensus-driven guideline. This rigid approach restricted the discovery of novel patterns, as informative signal segments might still be considered "normal" under an arbitrarily chosen alert threshold.

To address this, the presented system redefines the alert horizon as a tunable hyperparameter, rather than a fixed a priori value. This means that the optimal alert window is determined dynamically, based on the hyperparameters combination that yields the best classification performance. This expands the parameter search space, allowing the system to potentially identify previously overlooked but clinically relevant preictal patterns.

Another key enhancement introduced in this function is the transition from a binary labeling system to a multilabel approach. Instead of simply classifying each

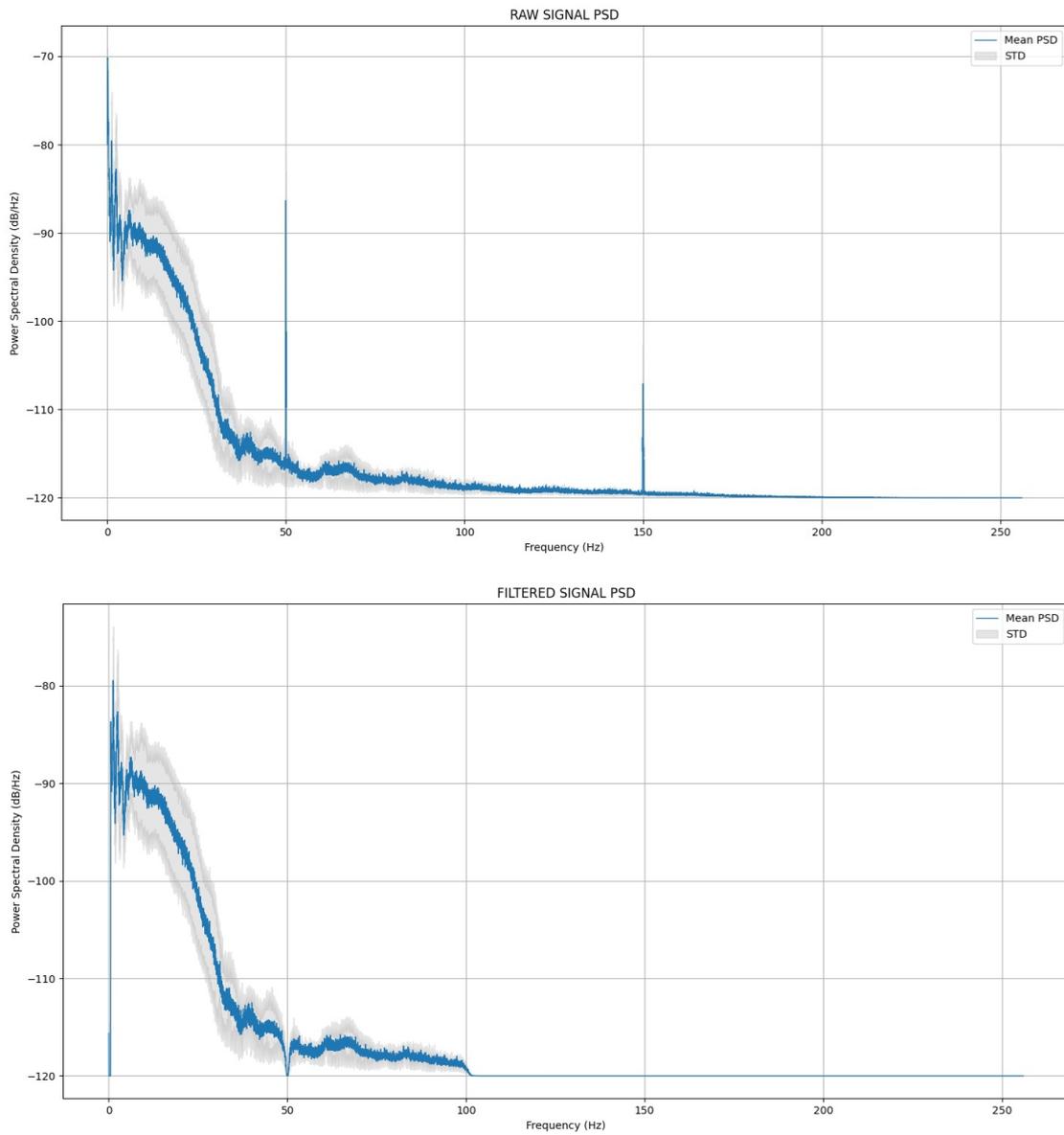


Figure 3.1: PSD comparison for the raw and filtered ECG signals. The raw acquisition clearly presents important power-line induced interference, successfully removed by the notch filter without causing an excessive information loss

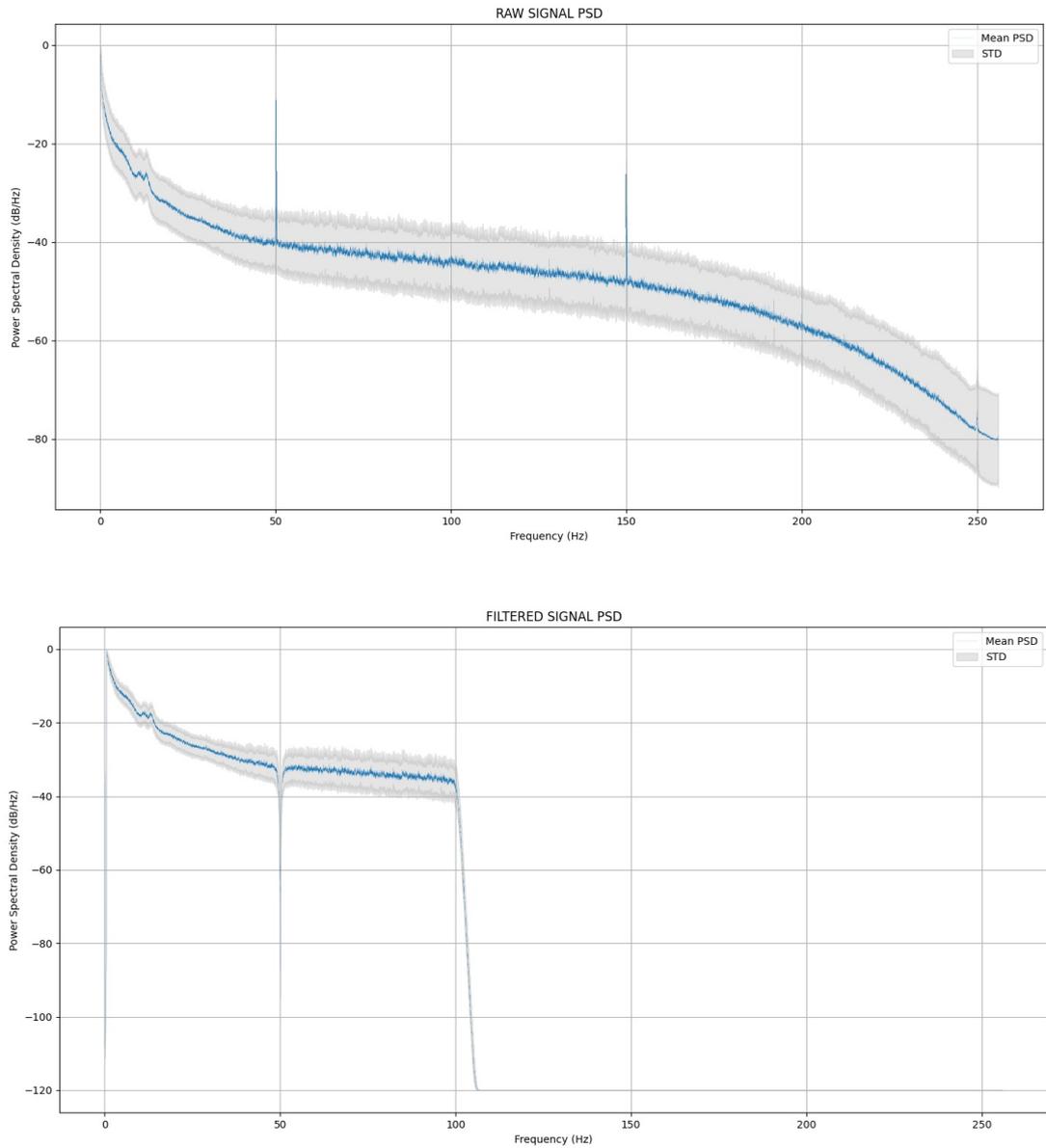


Figure 3.2: PSD comparison for the raw and filtered EEG signals. The raw acquisition still show power-line induce noise, but with a lower amplitude than the ECG case

time segment as "normal" or "preictal," the multilabeling strategy enables the detection of gradual, time-dependent preictal changes, if present. This approach could lead to improved classification performance by capturing progressive variations in seizure onset dynamics. The specific labels used in this system are outlined in Table 3.2.

Label	Description
0	Normality (windows distant more than 3 times the alert horizon)
1	Windows distant from 3 times the alert horizon to 2 times the alert horizon
2	Windows distant from 2 times the alert horizon to 1 time the alert horizon
3	Windows distant from 1 time the alert horizon to the actual start of the event

Table 3.2: Labeling system for event prediction based on alert horizon

It is important to note that while the system is theoretically capable of labeling seizure and post-seizure instants, these phases were excluded from the current implementation to maintain a strict focus on seizure prediction. To ensure that only normal and pre-seizure windows are retained for analysis, the recordings are truncated precisely at seizure onset, preventing seizure and post-seizure activity from influencing downstream processing, as they could become confounding factors.

In future developments, the system could be changed to allow predictions of temporally close seizures, allowing it to predict multiple seizure events within a short time frame. However, this functionality was not included in the present implementation due to dataset limitations, particularly the scarcity of recordings containing multiple closely spaced seizures.

### 3.2.3 Region mean calculation

This dimensionality reduction step is applied exclusively to EEG signals, as their multichannel nature results in high computational costs, particularly when bivariate features such as coherence or phase synchronization are considered. EEG caps typically contain a large number of electrodes, with higher electrode density improving spatial resolution. However, this increased precision comes at the cost of greater computational demands, making it essential to balance signal detail and processing requirements.

Despite the standardization provided by the 10-20 electrode placement system, minor variations in electrode positioning can still occur across different recordings due to inter-subject variability, cap fitting adjustments, or hardware-specific differences. To mitigate these issues while preserving spatial information, this pre-processing step reduces the number of channels while maintaining key topographic features.

The dimensionality reduction process consists of two main stages. First, electrodes are grouped into eight distinct scalp regions, defined based on their relative position along the coronal and sagittal planes, ensuring that the most relevant spatial information is retained. Next, a single representative channel is extracted from each region, computed as the median value of all electrodes within that zone. The median was chosen over the mean to reduce the impact of outlier noise and maintain a more robust regional representation.

The final output of this step is a reduced EEG matrix with eight channels, representing a trade-off between spatial resolution and computational efficiency.

### 3.2.4 Signal windowing

The primary objective of this step is to apply windowing to the original signal, segmenting it into fixed-length time windows for subsequent feature extraction and classification and making this approach real-time executable with the help of a FIFO buffer. However, in contrast to previous approaches where window length was manually tested and selected, this system treats window length as a tunable hyperparameter, optimizing it alongside other processing parameters. This automation eliminates the need for manual iterations, allowing the system to explore potential correlations between window duration and other hyperparameters, ultimately improving model performance.

The range of possible window lengths was determined by both practical constraints and feature calculation requirements. The lower limit was imposed primarily by the computational requirements of feature extraction algorithms. Since AR-based spectral estimation methods used in this work rely on sufficiently long segments to provide stable low-frequency power estimates, excessively short windows would compromise feature reliability.

In particular, among all extracted features, Heart Rate Variability (HRV) low-frequency power (0.04-0.15 Hz) posed the strictest requirement on window length.

### 3.2.5 Feature extraction

Feature extraction represents the final and most computationally demanding step of the preprocessing pipeline. The objective of this phase is to extract and compute a diverse set of features that have been widely utilized in the literature for seizure prediction, leveraging both EEG and ECG signals. Due to the substantial differences between these two modalities, they are processed independently in separate sections. However, despite this separation, the output structure remains consistent across both modalities:

1. A **matrix-based representation** intended for neural network training. In this format, each row corresponds to a specific extracted feature, each column represents a subwindow over which the feature was computed, and the different physical zones from which the signals originate are preserved as additional matrix dimensions.
2. A **vector-based representation** for machine learning classifiers. Each element in this vector corresponds to a feature extracted over an entire window, providing a more compact representation for traditional classification algorithms.

## EEG features

- **Cross-correlation:** Measures the similarity between two signals as a function of their relative time shift. It is useful for detecting synchronization and dependencies between different signals.
- **Phase Locking:** Quantifies the consistency of the phase difference between two signals across multiple time points. It is commonly used in neuroscience to assess phase synchronization in brain activity.
- **Dynamical Entrainment:** Describes the process where two oscillatory systems interact and synchronize their rhythms. It is relevant in physiological studies where external or internal factors influence neural and cardiac dynamics.
- **Relative Power:** Represents the proportion of power contained in a specific frequency band relative to the total power of the signal. This is crucial in EEG analysis to assess brain activity in different frequency ranges.
- **Power Band Ratio:** Computes the ratio between the power of two frequency bands, such as the theta/beta ratio, which is widely used in cognitive and neurological assessments.
- **Spectral Entropy:** Measures the randomness or complexity of a signal in the frequency domain. A higher value indicates a more irregular and unpredictable signal, which is relevant in distinguishing normal and abnormal brain activity.
- **Hjorth Parameters:** A set of statistical measures including activity (signal variance), mobility (frequency variability), and complexity (higher-order variations), used for characterizing EEG signals.
- **Spectral Edge Frequency:** The frequency below which a given percentage (typically 90%) of the total power of the signal spectrum is concentrated. It is a valuable parameter in EEG and ECG signal analysis for assessing dominant frequency components.
- **Approximated Entropy:** A statistical measure that quantifies the regularity and predictability of time-series data. Lower values indicate more predictable patterns, while higher values suggest increased complexity, often used in biomedical signal analysis.

## ECG features

- **Mean R-R Interval:** The average time interval between successive R-wave peaks in the ECG signal, measured in seconds. It reflects the mean heart rate over the analyzed period.
- **Standard Deviation of R-R Intervals :** Measures the overall variability of the heart rate by computing the standard deviation of R-R intervals. Higher values indicate greater heart rate variability (HRV), which is typically associated with better autonomic control.
- **Root Mean Square of Successive Differences:** Quantifies short-term HRV by calculating the square root of the mean squared differences between successive R-R intervals. It is particularly sensitive to parasympathetic (vagal) activity.
- **Percentage of Successive R-R Differences Greater than 50 ms:** Represents the proportion of successive R-R interval differences that exceed 50 milliseconds. It is commonly used to assess vagal tone and short-term HRV.
- **Low-Frequency Power:** The power spectral density in the low-frequency band (typically 0.04–0.15 Hz), which reflects both sympathetic and parasympathetic influences on heart rate modulation.
- **High-Frequency Power:** The power spectral density in the high-frequency band (typically 0.15–0.4 Hz), which is primarily linked to parasympathetic activity and respiratory-driven heart rate modulation.
- **LF/HF Ratio:** The ratio of low-frequency to high-frequency power, often used as an indicator of the balance between sympathetic and parasympathetic nervous system activity.
- **Variance of R-R Intervals:** Measures the dispersion of R-R intervals, providing insights into overall HRV. Higher values generally indicate better autonomic flexibility.
- **Poincaré SD1:** A nonlinear HRV metric derived from the Poincaré plot that reflects short-term variability, primarily influenced by parasympathetic modulation.
- **Poincaré SD2:** Another Poincaré plot-derived measure representing long-term HRV. It captures both sympathetic and parasympathetic contributions to heart rate dynamics.
- **Cardiac Sympathetic Index:** A measure derived from Poincaré plot analysis that provides an estimate of sympathetic nervous system dominance.
- **Cardiac Vagal Index:** An index that quantifies the influence of the vagus nerve (parasympathetic activity) on heart rate regulation.
- **Katz Fractal Dimension:** A nonlinear metric that assesses the complexity of heart rate time series, providing insight into fractal-like properties of HRV.

- **Recurrence Rate:** A measure derived from Recurrence Plot Analysis that quantifies how often similar patterns in the heart rate time series reappear, reflecting system stability.
- **Determinism:** Indicates the percentage of recurrence points forming diagonal structures in a Recurrence Plot, which is related to the predictability and regularity of heart rate dynamics.
- **Maximum Diagonal Line Length:** Represents the longest uninterrupted diagonal line in the Recurrence Plot, linked to the degree of deterministic structure in heart rate variability.
- **Laminarity:** The proportion of recurrence points forming vertical structures in the Recurrence Plot, associated with sustained physiological states and slow heart rate fluctuations.
- **Trapping Time:** The average length of laminar (vertically structured) segments in the Recurrence Plot, providing insights into the persistence of heart rate states.
- **Shannon Diagonal Entropy:** A complexity measure derived from the distribution of diagonal line lengths in a Recurrence Plot, indicating the degree of randomness in heart rate dynamics.

After feature extraction, the resulting matrices can be considerably large, particularly due to the preservation of data from the eight different physical scalp zones as additional dimensions. While this enhances spatial resolution, it also increases computational complexity and required resources and may introduce risks of overfitting, particularly in neural network-based models.

To mitigate these issues, an optional dimensionality reduction step can be applied to transform the data into a more compact 2D matrix. For each extracted feature (i.e., for each row of the original matrix), four statistical descriptors are computed across all channels: mean, median, standard deviation, and maximum absolute deviation. These values are then stored as separate rows in the final representation, ensuring that each feature is described by four distinct values while maintaining the temporal structure of the dataset with the matrix columns.

### 3.3 Dataset handling

This stage of the pipeline is executed separately from preprocessing to increase computational efficiency. By allowing independent run, different dataset versions - each corresponding to a unique combination of preprocessing hyperparameters — can be computed once and stored. This significantly reduces redundant computations during training, as the dataset can be directly loaded from storage, making the execution time dependent solely on the model training process. This efficiency gain enables a higher number of trials to be conducted within the same execution time. The first step in the training preparation is the splitting of the dataset into training, validation and test sets. Instead of keeping entire recordings within a single subset, individual windows are randomly assigned across the three sets. This merging of different recordings ensures that the model does not overfit to session-specific artifacts or learn non-physiological trends, such as baseline drift or recording-specific noise patterns. By constructing a mixed dataset, the model is encouraged to focus on generalizable physiological features rather than stochastic recording variations.

It is important to note that during this phase two different approaches were followed: with the first one, patient recording were not mixed, leading to a dataset keeping different patients separated. Following the complementary approach, a general dataset was created where both windows and patients were mixed. This choice follows the disputed idea that a generalizable system is feasible and achievable. Ideally, once the training is terminated, this system would be able to classify even new, unseen patient. However, even if this approach leads to decreased performance, a general system could still be useful as it could require simple fine tuning to adapt it to a specific patient, leading to increased performances with a lower number of data required when compared to a complete training starting from zero.

A key advantage of this dataset split approach is its impact on algorithms training robustness. In this implementation, each independent training iteration is conducted on a unique dataset partition, reducing the likelihood that high accuracy is merely the result of a favorable dataset assignment. If multiple training runs with different dataset splits converge to similar results, this provides stronger evidence of genuine model performance, rather than an artifact of dataset partitioning.

Following dataset partitioning, a balancing phase is applied to address the class imbalance inherent to seizure prediction. Since seizures are rare events, the dataset naturally contains more normal (interictal) samples than preictal samples, which can bias model training. To mitigate this, a hybrid balancing strategy is implemented: Oversampling of the minority class (preictal windows): Additional preictal samples are generated by applying Gaussian noise augmentation, where the noise amplitude is calculated relative to the original signal amplitude to preserve realistic variability. Undersampling of the majority class (normal windows): The interictal class is downsampled to match the number of preictal samples, ensuring class balance.

Before model training, a Z-score normalization step is applied to standardize the dataset. To prevent data leakage, the normalization parameters (mean and standard deviation) are computed exclusively on the training set. These same parameters are then used to normalize the validation and test sets each channel independently, ensuring a consistent feature distribution without introducing biases from future data.

## 3.4 Classification training

This represents the last section of the code: due to the scarcity of available data, it was not feasible to implement a dedicated postprocessing phase, as this would have required reusing training data, thereby introducing the risk of data leakage and overfitting. The execution flow diverges at this stage depending on the chosen classification approach, whether using neural networks or traditional machine learning algorithms. An important consideration in single-patient training is the need to avoid model overfitting to session-specific artifacts. To address this, only 3 out of the 14 original patients were selected for training, as these were the only subjects with at least 5 different recordings. This criterion was necessary because a limited number of recordings per patient could lead the model to learn classification based on recording-specific noise, baseline wandering, or acquisition artifacts, rather than true physiological patterns. By ensuring that each selected patient has at least five independent recordings, the dataset achieves a more diverse representation of intra-subject variability. This helps the model generalize beyond recording-specific biases and instead focus on medically relevant seizure-related patterns, ultimately leading to more reliable and robust classification results. In particular, the final dataset after balancing contained recordings with the following duration:

- **Patient 00:** 1 hour and 15 minutes, mean preictal period of approximately 15 minutes
- **Patient 06:** 7 hours and 30 minutes, mean preictal period of approximately 1 hour and 30 minutes
- **Patient 10:** 3 hours, mean preictal period of approximately 40 minutes

### 3.4.1 ML training

Machine learning algorithms serve as statistical-based classifiers that typically require fewer training samples compared to neural networks. This characteristic makes them particularly well-suited for this system, given the limited size of the dataset. Unlike NNs, which require large amounts of data to learn hierarchical feature representations, ML models are highly effective with structured feature vectors, making them an ideal choice in this context.

To further enhance model reliability and robustness, a stratified k-fold cross-validation approach was implemented, ensuring balanced class distribution across folds. This process was performed on the merged training and validation set, maximizing training data availability while maintaining unbiased performance estimation. Importantly, the test set was kept entirely separate to preserve an independent final evaluation. Given the diversity of ML algorithms, multiple classifiers were tested in each fold, and their performance was ranked based on Out-Of-Fold (OOF) metrics.

To further mitigate the risk of overfitting, each classifier was evaluated using two different feature selection strategies:

1. **Full Feature Set** – The model is trained using all available features in the vector.
2. **Forward Feature Selection (FFS)** – The model starts with a single feature and additional features are sequentially added based on their impact on model performance. At each iteration, the feature that improves classification metrics the most is retained, and the process continues until performance plateaus or drops below a predefined tolerance threshold.

Following these optimization steps, the best-performing classifier — determined based on cross-validation performance — was selected and applied to the independent out-of-fold samples. This final evaluation provided unbiased classification metrics, ensuring that the model’s generalization ability was accurately assessed.

### 3.4.2 NN training

Differently from the ML-based approach, the dataset is not further modified in this section; instead, it is processed according to the original assignments into training, validation and test sets. The choices available in this phase relate to the selection of different neural network architectures and their respective training parameters.

A key aspect of this approach is that neural networks operate on matrices rather than feature vectors, making them more sensitive to time-dependent variations. Each column of the input matrix corresponds to a distinct subwindow over which the features were extracted, so the architectures are explicitly designed to exploit this temporal structure for classification. However, it is important to note that due to the data-intensive nature of neural networks, the performance obtained on this dataset may not fully reflect the system potential in scenarios where larger datasets are available.

Another critical consideration is explainability. While ML algorithms retain a degree of interpretability — since the contribution of individual features to the final decision can be analyzed — neural networks lack this level of transparency. In medical applications, this lack of explainability poses a significant challenge, as both physicians and patients may be skeptical about the classification outcome without an interpretable rationale.

To optimize the neural network training process, Neural Network Intelligence (NNI) played a pivotal role, aligning well with its intended purpose. Instead of relying on manual trial-and-error or suboptimal grid search methods, the tuning algorithm explored the hyperparameter search space adaptively. This approach allowed the system to focus on the most promising configurations while systematically discarding regions of the search space that consistently led to suboptimal performance. In particular, each trial was conducted using a Tree-structured Parzen Estimator (TPE) optimizer to select the different hyperparameters since it offers efficient exploration-exploitation even in complex search spaces. An example of search space can be found in the table 3.3 reported here below.

Table 3.3: Hyperparameter Search Space

Hyperparameter	Type	Values
risk_window_length	Choice	{5, 10, 15, 20}
window_length	Choice	{70, 90, 110}
use_dynamic_cost_matrix	Choice	{false, true}
dense_units	Choice	{16, 32, 64}
kernel_size	Choice	{3, 5, 7}
dilation_rate	Choice	{1, 2, 4}
use_second_conv1d_block	Choice	{false, true}
activation	Choice	{tanh, sigmoid, relu}
learning_rate	LogUniform	$[10^{-7}, 10^{-4}]$
regularization_l2	LogUniform	$[5 \times 10^{-4}, 5 \times 10^{-2}]$
loss	Choice	{sparse_categorical_crossentropy, kullback_leibler_divergence}
batch_size	Choice	{2, 4, 6}
dropout	Choice	{0, 0.15, 0.3}

### 3.4.3 Final Classification

As previously discussed, the training phase employs a multilabel approach to enhance sensitivity to potential variations observed in the pre-ictal phase. However, this classification scheme may be less intuitive for the end user. To improve usability, the final output can be binarized by merging different classes. This simplification not only enhances user experience but may also reduce the cognitive load and stress associated with interpreting a multilabel system, despite its potential for higher classification granularity. For this reason, the performance metrics presented in the next chapter will include both multilabel and binary classification approaches.



# Chapter 4

## Results

The experiments were conducted with two distinct objectives. The first aimed to develop a system capable of correctly predicting seizures using a patient-specific approach, where models were trained on data from multiple recordings of the same patient. This approach ensures that the system learns individual-specific preictal patterns, optimizing its performance for each subject.

The second objective focused on generalization, where recordings from different patients were combined into a single dataset. This strategy allows the model to learn more general seizure-related features that are not limited to a specific individual, making the system compatible with new, unseen patients.

In both scenarios, EEG and ECG features were treated separately, as initial test trials showed that different values in the hyperparameters choice are required to achieve optimal performance. While a unified system integrating both signals is feasible, it would require an additional classification layer, such as a Multi-Layer Perceptron (MLP) or another machine learning model to merge the outputs from EEG- and ECG-based classifiers into a final decision, leading to an increased amount of training data required.

Another solution could be the direct training of a single system using both EEG and ECG features: despite being theoretically possible, it poses a significant risk of overfitting due to the limited dataset size. In such a scenario, the model might learn to classify based on non-physiological patterns, relying on session- or patient-specific artifacts rather than genuine seizure-related features. This highlights the importance of further data collection and careful feature integration when designing multimodal seizure prediction models.

A final consideration regarding introducing a post-processing step has to be made: many different post-processing methods could be applied to reduce sporadic erroneous classifications and highlight the general trend, but to properly tune this approach an increased amount of data is required, since using the same training set already employed to train the system would significantly increase the risk of overfitting. A correct approach would instead require the use of entire recordings specifically and exclusively dedicated to this step, but the limited number of recordings originally available made this impossible, and therefore no post-processing was applied, despite its potential advantages.

Before presenting the results, it is important to clarify that although the systems were trained using the multilabel approach previously described to enhance sensitivity to the temporal evolution of the signals, the final evaluation metrics were

binarised. This binarisation involved merging the normal class and class 1 samples into one category and similarly merging classes 2 and 3 into another. Despite its simplicity, this strategy provided a practical starting point, avoiding the complexity of a full multilabel system and resulting instead in a straightforward binary alarm system.

## 4.1 Patient specific system

This section introduces the results obtained in the experiments focusing on a system-specific approach, where each patient is treated independently. As previously mentioned, a reduced number of patients presented at least 5 recordings and therefore were used while following this approach. EEG and ECG were treated separately, therefore results are divided in two different subsections.

### 4.1.1 ECG signal

#### Machine Learning

Given the limited dataset size, a stratified k-fold cross-validation strategy was preferred to obtain more robust performance metrics, with  $k = 3$ . This approach ensured that all reported values in Table 4.1 were computed using out-of-fold samples, thereby providing a more reliable estimate of the model generalization capability. It is important to remember that each sample represents a single signal window which was taken from the 5 different recordings previously mixed. The duration of those windows, tuned for each patient, was:

- **Patient 00:** 90 seconds
- **Patient 06:** 70 seconds
- **Patient 10:** 80 seconds

The alert horizon represents the time threshold used to separate normal and preictal windows while labeling. The lower horizon in patient 00 is due to the seizure starting close to the recording start. Finally, the model selected after the training phase from the previously presented methods was the random forest algorithm.

Patient	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
00	82.61	81.82	83.33	8.70	6
06	79.25	82.61	77.62	5.66	20
10	66.38	68.29	65.33	11.21	14
<b>Mean <math>\pm</math> Std</b>	<b>76.08 <math>\pm</math> 6.99</b>	<b>77.57 <math>\pm</math> 6.57</b>	<b>75.43 <math>\pm</math> 7.51</b>	<b>8.52 <math>\pm</math> 2.27</b>	<b>13.33 <math>\pm</math> 5.73</b>

Table 4.1: Final metrics reported on the selected patients using ML on ECG signal. Note: FA = False Alarm.

## Neural Networks

Unlike the ML models, NN results were obtained without k-fold cross-validation due to the higher computational request demanded by the NN during the training. Table 4.2 presents the performance metrics reported by the system on the test set, but the reduced size of this set results in fewer samples, leading to increased differences among the different patients. Up to 5000 independent training processes were conducted for each patient, each one with a distinct partitioning of training, validation and test sets. The consistency observed across different training reinforced the reliability of the reported results, mitigating concerns that performance variations were merely caused by specific data partitioning rather than genuine indicators of model effectiveness. It is important to note that due to the limited size in the ECG matrices, only the LSTM architecture was tested since it can better cope with the smaller matrices when compared to the other architectures selected in this work. As explained in the ML section, each sample represents a single signal window which was taken from the 5 different recordings previously mixed. The duration of those windows, tuned for each patient, was:

- **Patient 00:** 90 seconds
- **Patient 06:** 90 seconds
- **Patient 10:** 100 seconds

This length was then divided in 3 subwindows with equal length to create the final matrix used in the EEG case.

Patient	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
00	90.00	100.00	90.00	0.00	8
06	86.79	66.67	89.36	3.77	14
10	65.62	53.33	87.50	30.43	14
<b>Mean <math>\pm</math> Std</b>	<b>80.80 <math>\pm</math> 10.82</b>	<b>73.33 <math>\pm</math> 19.63</b>	<b>88.95 <math>\pm</math> 1.06</b>	<b>11.40 <math>\pm</math> 13.54</b>	<b>12.00 <math>\pm</math> 2.83</b>

Table 4.2: Final metrics reported on the selected patients using an LSTM neural network on ECG signal. Note: FA = False Alarm.

Before analyzing these results it is important to highlight a limitation in the patient 00 data, where each seizure occurs approximately 15 minutes after the start of each recording. This significantly reduces the number of available preictal windows for training, consequently shortening the prediction horizon for this patient and reducing the number of samples available for training and testing the system. As later discussed and frequently mentioned, the size of the available dataset represents a critical point in the development of these systems.

## Analysis and Discussion of Results

Overall, the obtained results are promising, demonstrating the feasibility of seizure prediction and in particular of ECG-based systems. While a direct comparison with previous research is challenging due to the different methodologies reported and different metrics usage, the results indicate an improvement in classification performance and enhanced system reliability. Notably, the preliminary findings suggest an

increased sensitivity, with a 7% gain compared to [c7] and lower percentages of time under false alarm, using the metric suggested in [c2]. This translates into a system that could gain a higher level of trust from the final user, leading to an increased number of predicted seizures while still achieving a reduced number of false alarms.

## Comparison Between Machine Learning and Neural Networks

A comparison of ML and NN outcomes suggests that, given the limitations of the current data set, the ML models demonstrate superior stability and have greater intrinsic transparency in the criteria used to classify the signal, leading to better explainability. This aligns with expectations, as traditional ML methods generally require fewer training samples and exhibit less sensitivity to small dataset variations. In contrast, NNs have greater learning potential potentially leading to improved outcome, but require substantially larger datasets to fully exploit their advantages. Given this, ML approaches may remain preferable for small datasets, while NN techniques could achieve better performance once larger datasets become available.

### 4.1.2 EEG signal

#### Machine learning

The results reported in this subsection are obtained under the same constraints previously explained for the ECG case, but use the EEG signal. The duration of those windows, tuned for each patient and in the case of the best performing network, was:

- **Patient 00:** 90 seconds
- **Patient 06:** 80 seconds
- **Patient 10:** 80 seconds

Finally, the model selected after the training phase from the previously presented methods was the random forest algorithm.

Patient	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
00	79.25	66.67	87.50	7.55	8
06	83.46	81.48	85.29	7.63	30
10	84.00	62.86	92.22	5.60	30
<b>Mean <math>\pm</math> Std</b>	<b>82.24 <math>\pm</math> 2.12</b>	<b>70.34 <math>\pm</math> 8.03</b>	<b>88.34 <math>\pm</math> 2.89</b>	<b>6.93 <math>\pm</math> 0.94</b>	<b>22.67 <math>\pm</math> 10.37</b>

Table 4.3: Final metrics reported on the selected patients using ML on EEG signal. Note: FA = False Alarm.

## Neural networks

This subsection presents the final results achieved by the different architectures tested, with a table for each of the selected patients. The duration of those windows, tuned for each patient and in the case of the best performing network, was:

- **Patient 00:** 90 seconds
- **Patient 06:** 90 seconds
- **Patient 10:** 100 seconds

This length was then divided in 8 subwindows with equal length to create the final matrix used in the EEG case. The alert horizon represents the time threshold used to separate normal and preictal windows while labeling. The lower horizon in patient 00 is due to the seizure starting close to the recording start. Differently from the ECG case, EEG trials tested all the architectures since the final matrix had higher dimension.

The table 4.4 below presents the results achieved on the patient 00:

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)
LSTM	65.49	57.86	69.23	20.66
CNN	68.88	53.77	76.00	16.31
TCN	69.48	62.14	73.08	18.08
RNN	52.40	44.33	56.90	27.68

Table 4.4: Final metrics reported on patient 00 using NN on EEG signal.  
Note: FA = False Alarm.

The table 4.5 below presents the results achieved on the patient 06:

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)
LSTM	64.71	53.57	72.50	16.18
CNN	66.18	50.00	77.50	13.24
TCN	70.59	53.57	82.50	10.29
RNN	73.58	61.11	80.00	13.21

Table 4.5: Final metrics reported on patient 06 using NN on EEG signal.  
Note: FA = False Alarm.

The table 4.6 below presents the results achieved on the patient 10:

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)
LSTM	57.14	61.11	54.84	28.57
CNN	56.72	51.74	53.34	22.11
TCN	62.50	46.67	71.05	17.86
RNN	54.10	31.82	66.67	21.31

Table 4.6: Final metrics reported on patient 10 using NN on EEG signal.  
Note: FA = False Alarm.

For coherence with the data representation previously used, the table 4.7 represents the performance of the best system for each patient.

Patient	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)	
00	TCN	69.48	62.14	73.08	18.08	
06	RNN	73.58	61.11	80.00	13.21	
10	TCN	62.50	46.67	71.05	17.86	
<b>Mean <math>\pm</math> Std</b>		68.52 $\pm$ 4.57	56.64 $\pm$ 7.06	74.71 $\pm$ 3.83	16.38 $\pm$ 2.25	-

Table 4.7: Summary of the final metrics reported on each patient with the best NNs on EEG signal. Note: FA = False Alarm.

## Analysis and Discussion of Results

Overall, the obtained results are promising but still distant from the requirements of clinical implementation, reflecting the complexity and variability of the EEG signal. A direct comparison with previous research is challenging due to the different methodologies reported, different metrics usage and the large number and variability of EEG seizure prediction publications, however a general improvement in the alert horizon and other performance metrics can be noted.

## Comparison Between Machine Learning and Neural Networks

ML was able to achieve higher performance metrics with reduced time under false alarm, despite the lower complexity of the system. This difference could find its root in different key aspects: first of all, the higher complexity of the EEG signal would require a higher number of samples to truly leverage the advantages offered by the neural networks. Another explanation is in the features extraction step: ML relies on features calculated over the whole window, while NN uses a matrix where each column is the value of the feature calculated upon a subwindow. While a deeper focus would be required to confirm those assumptions, these tests are outside of the general, introductory overview aimed by this work, therefore they were not conducted.

## 4.2 Generalised system

Seizures generally present a high level of variability and this is one of the main causes that put in doubt the feasibility of a patient a-specific prediction system, but no definitive position is generally accepted. As a consequence of this incertitude, this system was tested on a general dataset mixing the different patients recordings, as previously mentioned. This forces the system to look for general, universal patterns that may be found in those recording and as a natural consequence, further reduces to the minimum the risk of using non-physiological pattern to classify. As explained in the patient-specific approach, EEG and ECG signals were still treated independently, avoiding a multimodal approach.

### 4.2.1 ECG signal

#### Machine learning

Even if the creation of a mixed dataset led to an increased number of available samples, the same k-fold strategies was implemented to grant metrics homogeneity and the possibility of a direct comparison with the patient specific case. In this case a single window duration was selected with a length of 70 seconds. Finally, the model selected after the training phase from the previously presented methods was the random forest algorithm.

Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
69.17	32.69	89.24	6.94	40

Table 4.8: Final metrics reported on general dataset using ML on ECG signal. Note: FA = False Alarm.

#### Neural networks

As explained in the ECG section, the same criterion used for the patient-specific approach was used also in this section. In particular, the matrix dimension restricted the test to the LSTM architecture only, since the window used - lasting 80 seconds - was divided in 3 subwindows with equal length to create the final matrix used in the EEG case.

Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
64.71	44.00	71.96	20.76	40

Table 4.9: Final metrics reported on general dataset using NN on ECG signal. Note: FA = False Alarm.

## Analysis and Discussion of Results

The results obtained are clearly reflecting the complexity of generalizing these systems, with reduced metrics when compared to single patient solutions. However, the level of performance reached is still highlighting the presence of some pattern, even if not strong and stable enough to obtain acceptable results. Another confounding factor that may have impacted the results is the use of the ECG signals: the metrics extracted are influenced by many factors, first of all the activity of the patient. The dataset contains data of both sleeping and awake patient, their mix can therefore introduce noise that further improve the difficulty of the problem.

## Comparison Between Machine Learning and Neural Networks

The creation of a general system has greatly improved the number of available training samples and this has clearly impacted the final results, with an emerging trend reflecting what was expected in this case: the use of a general system has increased the complexity of the problem, but the larger number of samples have led the neural networks to improve specificity, leveraging on their higher complexity level better coping with the differences in the present pattern.

### 4.2.2 EEG signal

#### Machine learning

Even if the creation of a mixed dataset led to an increased number of available samples, the same k-fold strategies was implemented to grant metrics homogeneity and the possibility of a direct comparison with the patient specific case. In this case a single window duration was selected with a length of 80 seconds.

Finally, the model selected after the training phase from the previously presented methods was the random forest algorithm.

Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
82.20	76.30	86.98	7.20	40

Table 4.10: Final metrics reported on general dataset using ML on EEG signal. Note: FA = False Alarm.

## Neural networks

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	Time Under FA (%)	Alert Horizon (min)
LSTM	64.35	59.43	66.67	22.66	40
CNN	64.65	60.38	66.67	22.66	40
TCN	57.31	38.33	64.10	24.91	30
RNN	55.16	40.71	62.24	25.35	40

Table 4.11: Final metrics reported on general dataset using NN on EEG signal. Note: FA = False Alarm.

## Analysis and Discussion of Results

This experiments lead to interesting results, in particular when considering the machine learning solution: the training metrics have reached the highest level when compared with all other solutions, with a quality that is near what could be started to be considered interesting in clinical routine. This is particularly true once a larger dataset becomes available, allowing for improved training and the creation of a post-processing technique.

## Comparison Between Machine Learning and Neural Networks

Differently from the results obtained in the ECG case, NN have reached significantly lower results than the best ML algorithm trained, as already observed in the EEG, patient specific systems. There may be different reasons behind this outcome: firstly, ML could better cope with the higher complexity level reached in this case combined with the number of data, with NN suffering the lack of more data needed to reach this higher level result. Secondly, there is a difference in the feature extraction section: ML receives a vector where each element is the value of the feature calculated upon the complete window, while NN receives a matrix where each column represent the feature value upon a subwindow. In this case, with all the variabilities coming from the different patients, the smaller dimension could increase the noise present in the matrix and therefore impact the classification quality.

### 4.3 Final discussion and conclusion

All the results presented above highlight several important aspects and open interesting perspectives for future testing and exploration. The first observation is that in all different scenarios, this project actually produced promising results despite some variability depending on specific conditions. This demonstrates the reusability of the overall pipeline across the various tested situations, regardless of differences in signals and other parameters used. Additionally, the general performances also show the feasibility and potential advantages of a data-driven approach, allowing specifically tuned parameters for each scenario and thus increasing the quality of the final outcomes. Another implicit advantage was that manual interventions and subjective decisions were almost completely removed during the training phase.

A deeper focus is now shifted toward the obtained results themselves: as recent studies indicate, ECG has confirmed its potential in seizure prediction, paving the way for seizure prediction systems with significantly enhanced portability and comfort for the patient. The previously presented tables for patient-specific systems show that ECG-based systems generally achieved better metrics compared to their EEG counterparts. This could be explained by the physiological changes preceding a seizure, such as imbalances in the parasympathetic nervous system, but it might also be due to technical reasons like higher noise and acquisition challenges present in EEG signals. Beyond the comfort advantages of ECG over EEG acquisition, another important implication arises: since the features were extracted exclusively from heart rate variability, without requiring information from the actual ECG waveform itself, a simple PPG-based smartwatch with appropriate preprocessing could potentially be used as a simplified acquisition interface. This approach would significantly enhance ease of use and patient comfort, providing greater freedom while still reliably alerting them to incoming seizures. However, the situation reverses in the general system scenario: here, EEG systems achieved significantly better performance, likely due to differences in wakefulness and other physiological factors across patients, indicating a need for some personalization to enhance classification accuracy.

Another significant finding concerns the feasibility of a generalized system: although metrics were lower when compared to patient-specific experiments, performance still reached a level that suggests further research could obtain valuable results. In particular, having a larger number of recordings could help achieve two separate objectives. First, the creation of a truly generalized system capable of accurately predicting seizures in new, unseen patients without individual customization. Alternatively, a hybrid approach could leverage the expanded dataset to create a robust baseline model, subsequently fine-tuned using a limited number of patient-specific recordings to incorporate individual-specific seizure patterns along with generalized ones. These results also highlight an important difference between machine learning techniques and neural networks. As expected due to the limited dataset size, machine learning generally showed better performance, except in the general system using ECG signals, where neural networks demonstrated higher sensitivity. This could be explained by the increased complexity arising from mixing patient recordings in varying states of wakefulness, which might be better addressed by the more complex neural network models. Machine learning techniques remain preferable for developing medical applications due to their greater interpretability com-

pared to the more "black-box" neural network models. Although the performance difference observed is likely influenced by dataset size limitations, the current results are still acceptable. Nevertheless, the dataset size represents the primary limitation of this study; therefore, future work should prioritize acquiring larger and longer datasets, including a higher number of recordings per patient accompanied by detailed metadata. Addressing this limitation is expected to enhance performance and significantly increase the robustness of the system, finally leading the long and winding road of seizure prediction to an end.



## Acknowledgements

Arrivati alla conclusione di questo lavoro desidero esprimere la mia profonda gratitudine a tutti coloro che mi hanno accompagnato e sostenuto durante questo percorso.

Innanzitutto, ringrazio il mio relatore, Professor Luca Mesin, per i preziosi consigli e la grande fiducia che mi ha fornito. La sua disponibilità e competenza sono state per me esemplari.

Un ringraziamento sentito va alla mia famiglia, che da sempre mi ha sostenuto con affetto e pazienza, incoraggiandomi nei momenti di difficoltà e garantendomi un supporto su cui contare.

Profonda riconoscenza va poi ai miei amici e colleghi che mi hanno accompagnato in questa strada, chi partendo fin da subito e chi aggiungendosi più avanti. Le discussioni e le esperienze fatte, oltre che essere un prezioso ricordo, mi hanno profondamente arricchito e hanno contribuito al raggiungimento di questo obiettivo.

Un ringraziamento speciale, infine, va alla persona che più di tutte è riuscita ad accompagnarmi in questo percorso, fornendo sempre sostegno ed esempio ammirevoli.

# Bibliography

- [1] A. Aarabi, R. Fazel-Rezai, and Y. Aghakhani. “EEG seizure prediction: Measures and challenges”. In: *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. Accessed: 2024-08-26. Minneapolis, MN, USA, 2009, pp. 1864–1867. DOI: 10.1109/IEMBS.2009.5332620.
- [2] R. G. Andrzejak et al. “Seizure forecasting: Where do we stand?” In: *Epilepsia* 64 (2023). Accessed: 2024-09-2, S62–S71.
- [3] M. O. Baud et al. “Seizure forecasting: bifurcations in the long and winding road”. In: *Epilepsia* 64 (2023). Accessed: 2024-09-2, S78–S98.
- [4] Deepak L. Bhatt. *What is heart rate variability?* Accessed: 2024-08-18. Harvard Medical School. URL: <https://www.health.harvard.edu/heart-health/what-is-heart-rate-variability>.
- [5] Benjamin H. Brinkmann et al. “Seizure Diaries and Forecasting With Wearables: Epilepsy Monitoring Outside the Clinic”. In: *Frontiers in Neurology* 12 (2021). Accessed: 2024-09-2. DOI: 10.3389/fneur.2021.690404. URL: <https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2021.690404>.
- [6] T. Editors of Encyclopaedia Britannica. *action potential*. Accessed: 2024-08-17. Encyclopedia Britannica. 2024. URL: <https://www.britannica.com/science/action-potential>.
- [7] T. Editors of Encyclopaedia Britannica. *electroencephalography*. Accessed: 2024-08-17. Encyclopedia Britannica. URL: <https://www.britannica.com/science/electroencephalography>.
- [8] T. Editors of Encyclopaedia Britannica. *Spinal cord anatomy*. Accessed: 2024-08-10. Encyclopedia Britannica. 2023. URL: [https://www.physio-pedia.com/Spinal\\_Cord\\_Anatomy](https://www.physio-pedia.com/Spinal_Cord_Anatomy).
- [9] T. Editors of Encyclopaedia Britannica. *Vagus Nerve*. Accessed: 2024-08-09. Encyclopedia Britannica. 2024. URL: <https://www.britannica.com/science/vagus-nerve>.
- [10] Kendra Cherry. *How Does the Nervous System Work With the Endocrine System?* Accessed: 2024-08-09. 2023. URL: <https://www.verywellmind.com/the-nervous-and-endocrine-systems-2794894>.
- [11] Kendra Cherry. *How the Peripheral Nervous System Works*. Accessed: 2024-08-09. 2022. URL: <https://www.verywellmind.com/what-is-the-peripheral-nervous-system-2795465>.
- [12] Paolo Detti. *Siena Scalp EEG Database (version 1.0.0)*. PhysioNet. Accessed: 2024-08-26. 2020.

- [13] Paolo Detti, Giulia Vatti, and Gorka Zabalo Manrique de Lara. “EEG Synchronization Analysis for Seizure Prediction: A Study on Data of Noninvasive Recordings”. In: *Processes* 8.7 (2020). Accessed: 2024-08-26, p. 846.
- [14] A. Goldberger et al. “PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals”. In: *Circulation [Online]* 101.23 (2000). Accessed: 2024-08-26, e215–e220.
- [15] Jean Gotman. “Automatic recognition of epileptic seizures in the EEG”. In: *Electroencephalography and Clinical Neurophysiology* 54.5 (1982). Accessed: 2024-09-5, pp. 530–540.
- [16] Patty Osborne Shafer Joseph I. Sirven. *Facts and Statistics About Epilepsy*. Accessed: 2024-08-16. Epilepsy foundation. 2013. URL: <https://www.epilepsy.com/what-is-epilepsy/statistics>.
- [17] M. G. Leguia et al. “Learning to generalize seizure forecasts”. In: *Epilepsia* 64 (2023), S99–S113.
- [18] Klaus Lehnertz et al. “Seizure prediction by nonlinear EEG analysis”. In: *IEEE Engineering in Medicine and Biology Magazine* 22.1 (2003). Accessed: 2024-08-26, pp. 57–63. DOI: 10.1109/MEMB.2003.1191451.
- [19] Piotr Mirowski et al. “Classification of patterns of EEG synchronization for seizure prediction”. In: *Clinical Neurophysiology* 120.11 (2009). Accessed: 2024-08-26, pp. 1927–1940. ISSN: 1388-2457. DOI: 10.1016/j.clinph.2009.09.002. URL: <https://doi.org/10.1016/j.clinph.2009.09.002>.
- [20] Florian Mormann et al. “Seizure prediction: the long and winding road”. In: *Brain* 130.2 (Feb. 2007). Accessed: 2024-08-26, pp. 314–333. DOI: 10.1093/brain/awl241. URL: <https://doi.org/10.1093/brain/awl241>.
- [21] C. R. Noback et al. *Human Nervous System*. Accessed: 2024-08-09. Encyclopedia Britannica. 2024. URL: <https://www.britannica.com/science/human-nervous-system>.
- [22] I.E. Scheffer et al. “ILAE classification of the epilepsies: Position paper of the ILAE Commission for Classification and Terminology”. In: *Epilepsia* 58 (2017). Accessed: 2024-08-16, pp. 512–521. DOI: 10.1111/epi.13709. URL: <https://doi.org/10.1111/epi.13709>.
- [23] Unknown. *Epilepsy*. Accessed: 2024-08-16. WHO. 2024. URL: <https://www.who.int/news-room/fact-sheets/detail/epilepsy>.
- [24] Unknown. *Vagus Nerve Stimulation (VNS)*. Accessed: 2024-08-19. Epilepsy foundation. URL: <https://epilepsyfoundation.org.au/understanding-epilepsy/treatments-and-management/vagus-nerve-stimulation-vns/>.
- [25] Unkown. *Electroencephalography*. Accessed: 2024-08-17. Wikipedia foundation. URL: <https://en.wikipedia.org/wiki/Electroencephalography>.
- [26] Unkown. *Interneurons*. Accessed: 2024-08-09. 2024. URL: <https://www.physio-pedia.com/Interneurons>.
- [27] Unkown. *Nervoso, sistema*. Accessed: 2024-08-08. 2024. URL: <https://www.treccani.it/enciclopedia/sistema-nervoso/>.
- [28] Unkown. *Nervous system*. Accessed: 2024-08-08. 2024. URL: [https://en.wikipedia.org/wiki/Nervous\\_system](https://en.wikipedia.org/wiki/Nervous_system).

- [29] Unkown. *Nervous system*. Accessed: 2024-08-08. 2024. URL: <https://my.clevelandclinic.org/health/body/21202-nervous-system>.
- [30] Unkown. *The membrane potential*. Accessed: 2024-08-17. Khan Academy. URL: <https://www.khanacademy.org/science/biology/human-biology/neuron-nervous-system/a/the-membrane-potential>.
- [31] S. Muhammad Usman, S. Khalid, and M. H. Aslam. “Epileptic Seizures Prediction Using Deep Learning Techniques”. In: *IEEE Access* 8 (2020). Accessed: 2024-09-2, pp. 39998–40007. DOI: 10.1109/ACCESS.2020.2976866.
- [32] J. West et al. “Machine learning seizure prediction: one problematic but accepted practice”. In: *Journal of Neural Engineering* 20.1 (2023). Accessed: 2024-09-5, p. 016008.
- [33] E. Wirrell et al. “Introduction to the epilepsy syndrome papers”. In: *Epilepsia* 63.6 (2022). Accessed: 2024-08-16, pp. 1330–1332. DOI: 10.1111/epi.17231.
- [34] Lily Wong-Kisiel. *Epilepsy*. Accessed: 2024-08-16. Mayo clinic. 2023. URL: <https://www.mayoclinic.org/diseases-conditions/epilepsy/symptoms-causes/syc-20350093>.
- [35] X. Yang et al. “An Effective Dual Self-Attention Residual Network for Seizure Prediction”. In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021). Accessed: 2024-09-2, pp. 1604–1613. DOI: 10.1109/TNSRE.2021.3103210.
- [36] Y. Zhang et al. “Epilepsy Seizure Prediction on EEG Using Common Spatial Pattern and Convolutional Neural Network”. In: *IEEE Journal of Biomedical and Health Informatics* 24.2 (Feb. 2020). Accessed: 2024-08-26, pp. 465–474. DOI: 10.1109/JBHI.2019.2933046.