UNIVERSIDADE D
COIMBRA

Mariana Isabel da Silva Tavares

EEG Epilepsy Seizure Prediction: The Post-Processing Stage as a Chronology

Thesis submitted to the Faculty of Science and Technology of the University of Coimbra for the degree of Master in Biomedical Engineering with specialization in Clinical Informatics and Bioinformatics, supervised by Prof. Dr. César Teixeira and MSc Mauro Pinto.

February 2021

1 2 9 0

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE D COIMBRA

Mariana Isabel da Silva Tavares

EEG Epilepsy Seizure Prediction: the Post-Processing Stage as a Chronology

Thesis submitted to the University of Coimbra
for the degree of Master in Biomedical Engineering

Supervisors:
Prof. Dr. César Alexandre Domingues Teixeira (CISUC)
MSc Mauro Filipe da Silva Pinto (CISUC)

Coimbra, 2021

This work was developped in collaboration with:

Center for Informatics and Systems of the University of Coimbra

Agradecimentos

Quero começar por agradecer ao Professor Doutor César Teixeira e ao Mestre Mauro Pinto por me terem orientado na elaboração deste trabalho. O vosso apoio, a vossa disponibilidade e o vosso conhecimento foram fundamentais para todas as aprendizagens que adquiri ao longo deste percurso.

Um agradecimento também aos meus colegas de laboratório pela ajuda e por toda a partilha de ideias, críticas e sugestões.

Um agradecimento sincero a todos os que estiveram presentes ao longo deste meu percurso académico, pelos momentos especiais que partilhamos e que ficarão na memória para sempre. Àqueles que estiveram incessantemente ao meu lado, me deram força para continuar e me fizeram perceber que desistir não é opção, agradeço de coração.

Por fim, deixo o meu maior agradecimento à minha família pela paciência e pelas palavras de incentivo constantes desde o início do meu percurso escolar. A vocês ficarei eternamente agradecida.

Resumo

A possibilidade de prever crises é um aspeto encorajador para os 30% dos doentes epiléticos que não podem ser tratados nem com medicamentos anticonvulsivos nem através de procedimentos cirúrgicos. De forma a melhorar a qualidade de vida destes indivíduos podem ser criados algoritmos para incorporar dispositivos de aviso na eminência de uma crise ou para integrar sistemas que desarmam um ataque, por exemplo através de eletroestimulação cerebral. Tais algoritmos são baseados na identificação do período pré-ictal, caracterizado por ser um estado de transição entre a atividade cerebral normal e a crise.

Embora já existam muitos estudos de previsão com diferentes abordagens baseadas no Eletroencefalograma (EEG), nenhum atingiu a aplicabilidade clínica, sendo poucos aqueles que foram implementados em dispositivos médicos. Para além disso, os algoritmos que mais potencial apresentam levam a sistemas de caixa-negra, não estando os médicos dispostos a tomar decisões de alto risco a partir deles.

O presente trabalho propõe o desenvolvimento de um algoritmo de previsão de crises específico para cada doente, utilizando técnicas de pós-processamento para explorar, dentro do período pré-ictal, a existência de um conjunto de eventos cronológicos de atividade cerebral. Esta complexidade temporal pode ser estudada diretamente através de abordagens de Aprendizagem Profunda, penalizando a interpretabilidade. A metodologia proposta combina um conjunto de *features* lineares univariadas com um classificador baseado em Máquinas de Vetores de Suporte (SVM) e técnicas de pós-processamento, de forma a gerar alarmes antes dos ataques.

O estudo foi realizado em 36 doentes com Epilepsia do Lobo Temporal (TLE) da base de dados EPILEPSIAE e consistiu no desenvolvimento de duas abordagens diferentes de pós-processamento assentes no conceito de temporalidade. Foi testado um total de 106 crises num total de 1533.8 horas de dados. O desempenho de cada método foi apresentado para um conjunto de valores de Período de Ocorrência da Crise (SOP). A melhor performance foi alcançada para um SOP de 10 minutos.

Apesar da significância estatística alcançada ser de 69% (25 de 36 pacientes), estudos futuros devem investigar a Taxa de Falsos Positivos por Hora (FPR/h) máxima para cada estratégia de intervenção, a fim de avaliar a aplicabilidade desta metodologia.

Palavras-chave: Epilepsia, Previsão de crises, Eletroencefalograma, Métodos de Pós-processamento

Abstract

The possibility of predicting seizures is an encouraging aspect for the 30% of epileptic patients who cannot be treated either with anticonvulsant drugs or through surgical procedures. In order to improve the quality of life of these individuals, algorithms can be created to incorporate warning devices when a seizure is imminent or to integrate systems that disarm a seizure, for example through cerebral electrostimulation. Such algorithms are based on the identification of the pre-ictal period, characterized by being a state of transition between normal brain activity and the seizure.

Although there are already many predictive studies with different approaches based on the Electroencephalogram (EEG), none have achieved clinical applicability, with few being implemented in medical devices. In addition, the algorithms that have the most potential lead to black-box systems and physicians are not willing to make high-risk decisions from them.

The present work proposes the development of a patient-specific seizure prediction algorithm using post-processing techniques to explore, within the pre-ictal period, the existence of a set of chronological events of brain activity. This temporal complexity can be studied directly through Deep Learning approaches, penalizing interpretability. The proposed methodology combines a set of univariate linear features with a classifier based on Support Vector Machines (SVM) and post-processing techniques, to generate alarms before seizures.

The study was carried out in 36 patients with Temporal Lobe Epilepsy (TLE) from the EPILEPSIAE database and consisted of the development of two different post-processing approaches based on the concept of temporality. A total of 106 seizures were tested, comprising a total of 1533.8 hours. The performance of each method was presented for a set of Seizure Occurrence Period (SOP) values. The best performance was achieved for a 10-minute SOP. Despite having obtained a statistical significance of 69% (25 in 36 patients), future studies must investigate the maximum False Positive Rate per Hour (FPR/h) for each intervention strategy, to assess the

applicability of this methodology.

Keywords: Epilepsy, Seizure prediction, Electroencephalogram, Post-processing methods

Contents

List of Figures

List of Tables

List of Abbreviations

AR Autoregressive. 97

CNN Convolutional Neural Network. 33, 38, 41

DRE Drug-Resistant Epilepsy. 48
DSI Dynamic Similarity Index. 100

ECG Electrocardiography. 27
EEG Electroencephalogram. xiii, xiv, 2, 4, 5, 7, 8, 9, 10, 11, 12, 13, 18, 20, 21, 24, 25, 27, 28, 29, 30, 31, 33, 34, 35, 45, 47, 48, 49, 51, 59, 72, 73, 74, 95, 96, 97, 98, 100, 102
EMG Electromyography. 27

FFT Fast Fourier Transform. 97, 98
FIR Finite Impulse Response. 28
FOA Focal Onset Aware. 7, 74
FOIA Focal Onset Impaired Awareness. 7, 74
FPR/h False Positive Rate per Hour. xvii, 16, 17, 18, 43, 45, 58, 65, 66, 67, 70, 71, 72, 73, 74, 79

GA Genetic Algorithm. 37
GRU Gated Recurrent Unit. 41

iEEG Intracranial EEG. xiii, xvii, 12, 13, 27, 28
IIR Infinite Impulse Response. 28
ILAE International League Against Epilepsy. 5, 6

LSTM Long Short-Term Memory. 40, 41, 45

mDAD maximum Difference Amplitude Distribution of histogram. 36
ML Machine Learning. 3, 24
MLP Multi-Layer Perceptron. 41
MLTE Mesial temporal lobe epilepsy. 8

MPC Mean Phase Coherence. 101
mRMR minimum Redundance Maximum Relevance. 36, 37

PCA Principal Component Analysis. 36
PSD Power Spectral Density. 97

SEF Spectral Edge Frequency. 98
SEP Spectral Edge Power. 98
SOP Seizure Occurrence Period. xiv, xvii, 15, 16, 19, 28, 30, 43, 45, 47, 51, 55, 56, 57, 58, 61, 65, 68, 70, 71, 72, 73, 74, 75, 79
SPH Seizure Prediction Horizon. xiv, xvii, 15, 16, 20, 28, 30, 43, 51, 55, 65, 73
SS Sensitivity. xvii, 15, 17
SVM Support Vector Machines. 38, 40, 45, 52, 53, 61, 73

TLE Temporal Lobe Epilepsy. 8, 12, 72, 73, 79

1

Introduction

In this chapter, it is presented the motivation and the context of this project in section 1.1 and the expected goals and contributions in section 1.2. Section 1.3 states a brief description of the structure of this project

1.1 Motivation and Context

Motivation

Epilepsy is a chronic neurological disorder, affecting 65 million people worldwide [12], characterized by the prevalence of seizures. Seizures are provoked due to disturbances in the brain's electrical activity and characterized by an increase in neuronal synchrony. Without being controlled, epilepsy can lead to serious consequences, both at the physical and social levels, affecting the everyday life of the patient [5, 8, 12, 13].

both at the physical and social levels

Therefore, it is essential to treat patients in order to reduce their daily-life limitations. The majority can be controlled by anticonvulsant medication. When seizures are located in a part of the brain that can be removed, surgery may be also an option. Nonetheless, about 30% of patients are resistant to both treatments. It is for these people that seizure prediction looks promising [8, 13–15].

Main goal

A prediction model can allow minimizing seizure effects by identifying the pre-ictal period of a seizure and raising an alarm with sufficient time to take an action. The pre-ictal is a transitional period between a normal brain state and a seizure state (ictal period), on which the entire seizure prediction area rests [6, 8, 13].

Prediction mechanisms can pave the way for closed-loop systems investigation that manage to abort the seizure, by the automatic administration of antiepileptic-

drugs or by electrical brain stimulation, or to warn the patient to take the necessary precautions. This way, it would be possible to improve not only the quality of life of the patient himself, but also of those around him [13].

Difficulties

Seizure prediction started to be studied about 50 years ago, however the developments reached have been limited. So far, only two detection devices and one prediction study have been clinically approved. This is related to the brain's behavior, which is difficult to predict, but also with ethical issues [13, 14, 16]. Furthermore, there are other limitations, linked to databases or to the behavior of epilepsy, hindering the progress:

- Electroencephalogram (EEG) is a very complex signal and still not fully understood by the scientific community;
- No universal characteristic capable of predicting the pre-ictal period in epilepsy patients has so far been discovered. It is also not possible, at least to date, to clinically annotate the pre-ictal period [13, 14];
- Most databases are composed of discontinuous or short-term EEG records that do not represent well the reality. An EEG signal that is representative of real life, must contain significantly large periods of a normal brain state, i.e., inter-ictal data [13];
- Most databases contain records from pre-surgical monitoring patients in a hospital, away from their daily routine. Additionally, patients have more seizures than usual since the administration of medication is suppressed, which does not simulate a normal situation in their daily lives;
- Epileptic seizures also have an unpredictable nature that varies both from patient to patient and even seizure to seizure [13, 14].

Despite the existing efforts to create more robust algorithms, and the existence of increased computational power, the created models have to be able to offer an explanation of their decision, as stated in article 22 of the General Regulation of Data Protection of 2018 [17].

1.2 Goals and Contributions

Predicting epileptic seizures implies knowing in advance when a seizure is about to happen, only in this way is it possible to minimize its consequences. To this end, in previous works, authors tracked a prior stage of the seizure, the pre-ictal

period which is considered a transitional period, between a normal brain state and a seizure. A fixed time period is adopted for the pre-ictal period. However, what has been seen is that this period varies depending on the seizure and from patient to patient, making it almost impossible to find a standard value.

For this reason, it is intended to embrace a different view assuming that the epileptic seizure is not the consequence of just a single event, but rather a sequence of events, as in many other biological processes. This vision is not particularly new as it is part of the network theory. This theory advocates an epileptic network with interactions over large regions of the brain, rather than just interactions close to a well-defined epileptic focus [13, 15].

There are more complex methods, such as Deep Recurrent Neural Networks, capable of dealing directly with this temporality. However, they lack interpretability as their decisions are difficult to explain. Our approach attempts to address the temporality in a more easily explainable way, using knowledge from network theory. Thus, it is intended to highlight the nature of the brain as a biological system so the pre-ictal period is seen as a process of generating seizures, that is, as a series of events in chronological order of the brain activity [15, 18, 19].

This project aims to improve the most common structure in seizure prediction studies, by using explicitly the notion of the network theory, when using Machine Learning (ML), and thus providing an intuitive explanation for the system decisions [18]. For this purpose, a classifier model for predicting seizures will be developed:

- exploring the European epilepsy database (EPILEPSIAE);
- developing patient-tailored algorithms;
- evaluating different post-processing methods that are based on the notion of a chronology of events that lead to a seizure;
- statistically validating the classifier;
- comparing the developed methodology with the standard framework.

Currently, the post-processing methods that are used to deal with a classifier's decisions are ad-hoc and have no biological connection over time. Post-processing techniques that explore the chronology of events in a pre-ictal period will be developed. This way, through the existing knowledge of the brain, temporality can be directly applied in an easy-understanding approach to imitate some potentials of Deep Learning and overcome its interpretability problems.

Therefore, the research question is to ascertain whether there is a performance improvement in prediction algorithms by introducing the concept of temporality and prioritizing interpretability through more robust post-processing techniques.

1.3 Structure

This thesis proposal is composed of five more chapters structured as follows.

Chapter 2 provides background information related to the main concepts of Epilepsy and Seizure prediction.

Chapter 3 is devoted to the state of art concerning EEG seizure prediction in the last decades.

Chapter 4 describes all the steps of the adopted methodology.

Chapter 5 reports the results obtained for both the training phase and the testing phase of the proposed methodology along with a discussion and comparison of those results.

Chapter 6 presents a conclusion.

2

Background Concepts

This chapter is dedicated to introducing the main concepts of Epilepsy, Electroencephalography and Seizure prediction. Firstly, in section 2.1 are presented some definitions related to Epilepsy. Section 2.2 includes an overview of Electroencephalogram (EEG) and section 2.3 provides theoretical aspects related to Seizure prediction.

2.1 Epilepsy and seizures concepts

In 2005, the International League Against Epilepsy (ILAE), which promotes and disseminates knowledge about epilepsy, proposed a conceptual definition for both "seizure" and "epilepsy": *"an epileptic seizure is a transient occurrence of signs and/or symptoms due to abnormal excessive or synchronous neuronal activity in the brain"* and *"epilepsy is a disorder of the brain characterized by an enduring predisposition to generate epileptic seizures, and by the neurobiologic, cognitive, psychological, and social consequences of this condition. The definition of epilepsy requires the occurrence of at least one epileptic seizure"* [20].

In order to have a more practical definition for clinical application and for being in accordance with how epileptologists interpret epilepsy, in 2014, ILAE formulated an operational clinical definition of epilepsy [21]. Epilepsy is a brain disease defined by any of the following conditions:

"1. At least two unprovoked (or reflex) seizures occurring >24 h apart

2. One unprovoked (or reflex) seizure and a probability of further seizures similar to the general recurrence risk (at least 60%) after two unprovoked seizures, occurring over the next 10 years

3. Diagnosis of an epilepsy syndrome

Epilepsy is considered to be resolved for individuals who had an age-dependent epilepsy syndrome but are now past the applicable age or those who have remained seizure-free for the last 10 years, with no seizure medicines for the last 5 years."

Epilepsy is defined conceptually as a disorder, however in the operational clinical definition it is presented as a disease. The main difference between them is the fact that the term disease gives the dysfunction a longer duration [21].

The term "unprovoked", used in the first two points, can be imprecise nevertheless, it involves the absence of a temporary or reversible factor that reduces the threshold and heads to a seizure at any given time [21].

The operational clinical definition of epilepsy is a more complicated definition but it brings clinical relevance to the diagnosis and allows a better clarification about the recurrence risk after an unprovoked seizure, thereby facilitating initiation of treatment [21].

2.1.1 Classification of seizure types

Over the years, the classification of seizure types has been changing. The most recent classification in force was proposed by ILAE, in 2017 [1], with the main purpose of providing a universal way of communication for clinical use. This classification is more interpretive, with greater flexibility and transparency, allowing the use of additional data to classify a seizure (see Figure 2.1).

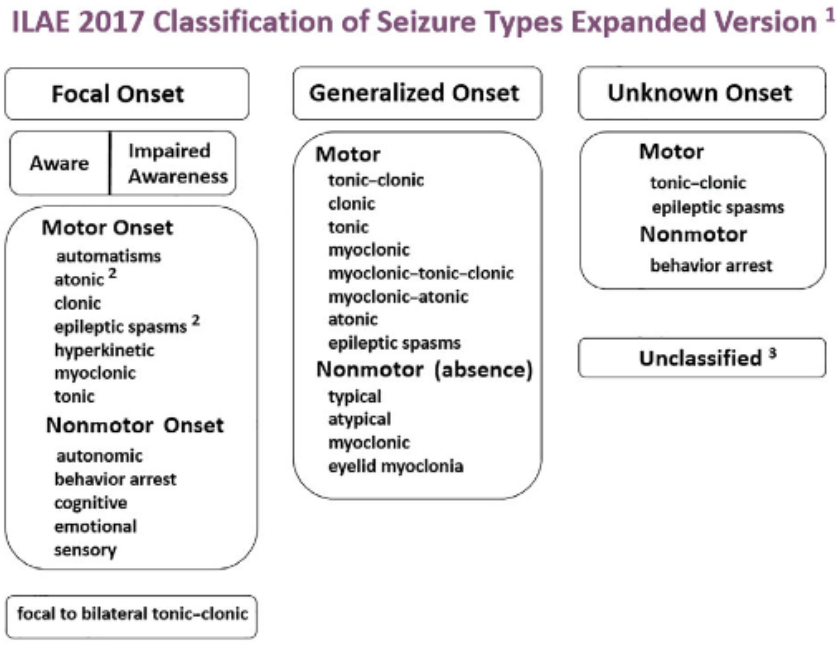

Figure 2.1: The expanded ILAE 2017 operational classification of seizure types. [1]Definitions, other seizure types and descriptors are listed in the accompanying paper and glossary of terms. [2] Due to inadequate information or inability to place in other categories. Source: Sheffer et al. 2017 [1].

The diagnosis process starts with the determination of the patient's seizure

type. Depending on where the initial manifestations begin in the brain, seizures are divided into three types: focal, generalized and unknown onset. Seizures can also be subcategorized into motor and non-motor. Those two classifications are optional, if there is no significant certainty, it is unknown or not applicable, must be omitted [1].

Focal seizures begin in neural networks of a single cerebral hemisphere. This seizure type is subgrouped in terms of awareness. During a Focal Onset Aware (FOA) seizure, a patient has knowledge of himself and environment. When this does not happen, the seizure is designated Focal Onset Impaired Awareness (FOIA). Focal to bilateral tonic-clonic is a singular type of seizure that has a focal onset but rapidly propagates to another brain location, reflecting a propagation pattern and not a single seizure type [1].

A generalized seizure originates from bilateral networks, that is, networks that involve both cerebral hemispheres. The level of awareness is not included in generalized seizures since most generalized seizures present partial or total disturbance of awareness [1].

The seizure onset may not be possible to determine being classified as unknown. When there is no great degree of confidence whether a seizure has a focal or a generalized onset, the seizure must be classified as "unclassified". An unknown seizure may be afterward classified as focal or generalized [1].

2.1.2 Classification of epilepsy types

In some cases, seizure type classification is the last level of classification achievable due to the nonexistence of EEG record, video, imaging studies, or to the insufficient information available. On the other hand, there are situations where it is possible to have a second level of classification according to the epilepsy type as seen in Figure 2.2. As per the figure, there are four types of epilepsy [2]:

- Focal epilepsies can be identified by showing focal epileptiform discharges on the interictal period and can encompass a variety of seizure types, such as FOA, FOIA, focal motor, focal non-motor and focal to bilateral tonic-clonic. This type of epilepsy comprises both seizures occurring in a single hemisphere and unifocal and multifocal disorders.
- Generalized epilepsies are identified by showing spike-wave activity on EEG and have the following seizure types: absence, myoclonic, atonic, tonic and tonic-clonic.
- Combined generalized and focal epilepsies show generalized spike-wave and

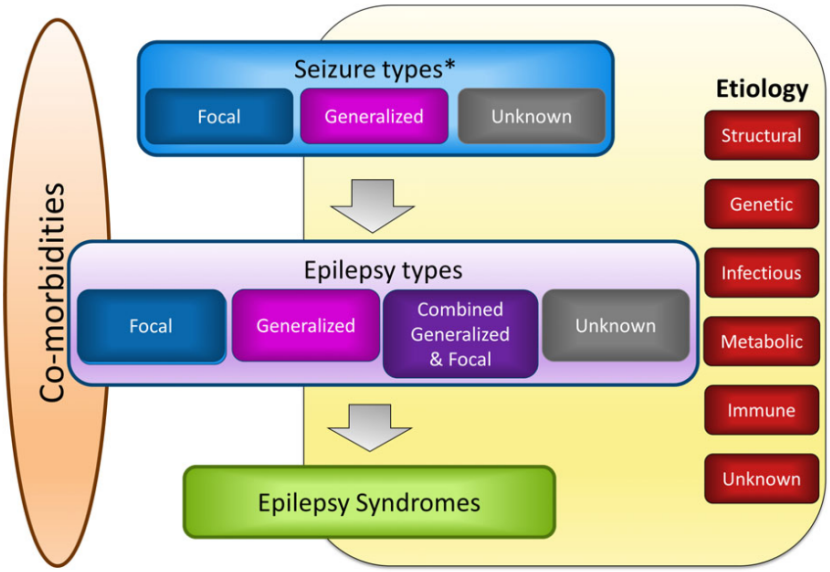

Figure 2.2: ILAE 2017 framework for classification of epilepsies. *Denotes onset of seizure. Source: Scheffer et al. 2017 [2].

focal epileptiform discharges on the interictal EEG once they comprise generalized and focal seizures.

- An epilepsy can be classified as unknown when a clinician cannot discern the type of seizure often because the information they hold is not sufficient or enlightening.

Temporal Lobe Epilepsy (TLE)

Temporal lobe epilepsy is the commonest form of focal epilepsy and is characterized by seizures beginning or involving the temporal lobes. TLE can be divided into two types: Mesial temporal lobe epilepsy (MLTE) and Neocortical or lateral temporal lobe epilepsy. While the first type involves the internal structures of the temporal lobe, the second occurs when seizures affect more outer structures. MLTE is the most prevalent type accounting for almost 80% of all temporal lobe seizures. In some cases, seizures often do not respond to antiepileptic medications, so surgery can be a good solution to improve risks related to lack of memory and emotional impairment. Most seizure prediction studies focus on TLE patients [8, 22, 23].

2.1.3 Epilepsy syndrome

A cluster of features occurring together characterizes an epilepsy syndrome. These features may include signs and symptoms such as epileptic seizure type, patterns on the EEG and brain images, age of onset and remission, diurnal variation,

seizure triggers, among others. A syndrome doesn't need to be defined by all existing features [2, 24, 25].

Besides being the third level of classification, the diagnosis of epileptic syndromes is very useful, since it helps the clinician to identify the seizure causes and to choose which is the best way to treat a patient. It can also help the prediction of new seizure remissions [2, 24].

2.2 EEG

The neuron is the fundamental unit of the brain. In the nervous system, there are millions of neurons, and they work together in a complex network to process and transmit information through electrical impulses [3, 26].

EEG is a registration of the electrical activity generated by the brain. More specifically, the EEG records represent voltage fluctuations, in space and time, due to all excitatory and inhibitory postsynaptic potentials produced by cortical pyramidal neurons [26–28]. It is needed thousands of synchronized neurons for the activity to be detected in EEG [3, 27, 29].

This medical measurement tool has excellent temporal resolution allowing the distinction of rapid changes in voltage. In contrast, it has poor spatial resolution depending on the number of used electrodes and how they are placed [3].

EEG enables analysis and interpretation of brain disorders, as the case under study [26]. It is noteworthy that the diagnosis is made by the physician's visual analysis of the EEG. Most of epileptic patients have EEG abnormalities [30]. For this reason, it is relevant to distinguish the different types of EEG waveforms (normal and abnormal) beyond identifying properly the characteristics linked to epilepsy [26].

Frequency, amplitude, shape and region of occurrence are characteristics analysed from the EEG, and depend on patient's age, level of awareness, mode of behaviour, etc [3, 26].

Generally, the EEG is characterized by its oscillatory behaviour, however, it also has other patterns of shorter duration not necessarily rhythmic. Consequently, two different types of phenomena can be registered in an EEG (see Fig. 2.3): oscillations and transients. Oscillations are commonly characterized by the waveform frequencies, so they are divided into different sub-bands: delta (2-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30), gamma (greater than 30Hz) [26, 27].

EEG sharp transients are subdivided into normal and abnormal. Normal transients can be related to sleep episodes or to artifacts [27]. Most of sleep potentials occur when a subject change its state of alertness, mainly in stage 2 of NREM sleep

[26]. Artifacts are electrical potentials from sources other than the brain. They are considered noise and can be physiological such as the eye blinks, cardiac impulses, breathing, chewing, muscle movements, or external as electromagnetic interference from the surrounding environment and improper electrode placement. Both the identification and the removal of these signs from EEG records present major challenges to the experts [3].

Abnormal transients are crucial for epilepsy diagnosis, so they are divided into epileptical and non-epileptical. Non-epileptiform EEG transients point to other brain disorders [26, 27].

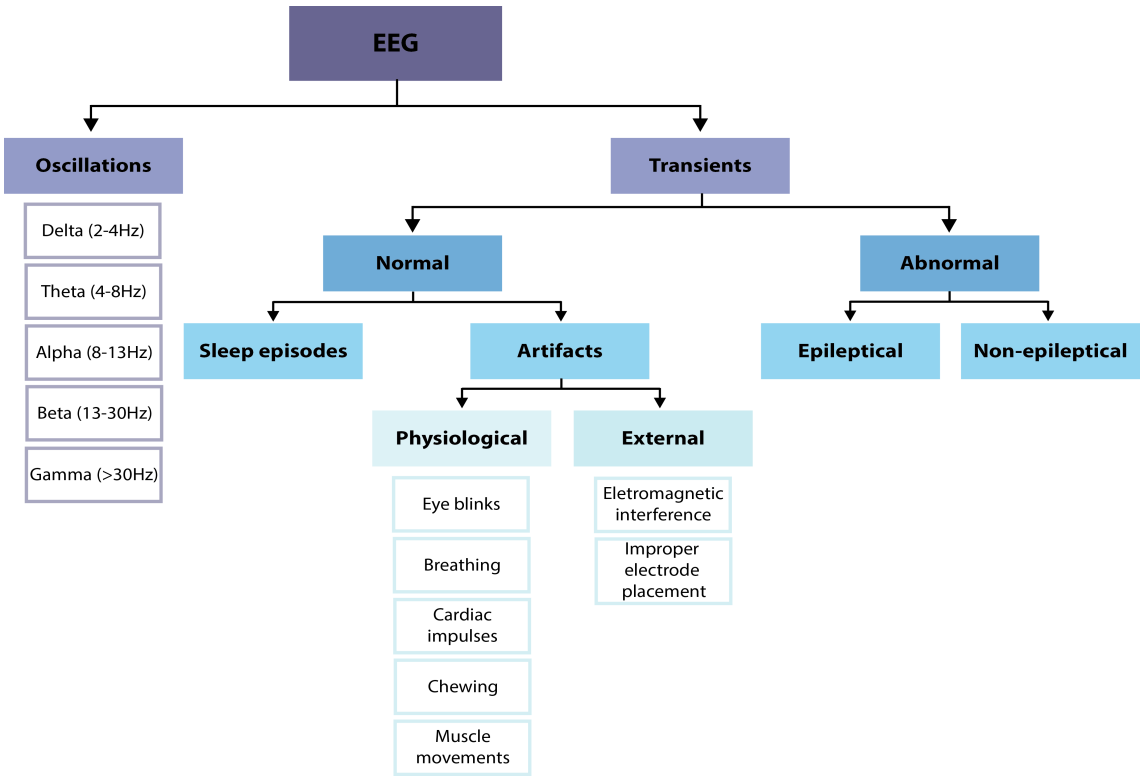

Figure 2.3: Categorization of EEG activity.

The electrodes placement and the chosen reference system have a major impact on the acquisition of the EEG. The positioning of the electrodes can be in the scalp or directly in the exposed surface of the brain [3].

2.2.1 Scalp EEG

10-20 System

For recording the EEG, electrode placement normally follows the international 10-20 system (see Fig.2.4 a)). The locations of the electrodes are defined according to

the proportions of 10% and 20% of distances between some anatomical landmarks. The front-back distance of the head is calculated between the nasion and inion points, and the pre-auricular points allow the determination of right-left distance [3, 29, 31].

The electrodes on the left hemisphere are named with odd numbers, on the right hemisphere with even numbers and midline electrodes are referred with a "Z". According to the brain's region where the electrode is located, it is denoted with different prefixes: "P" for parietal, "F" for frontal, "Fp" for frontopolar, "T" for temporal, "C" for central and "O" for occipital [27, 32]. This standard convention takes into consideration the heterogeneity of heads size allowing proportional head coverage, thus the records are more easily compared between patients [3, 27, 29, 31].

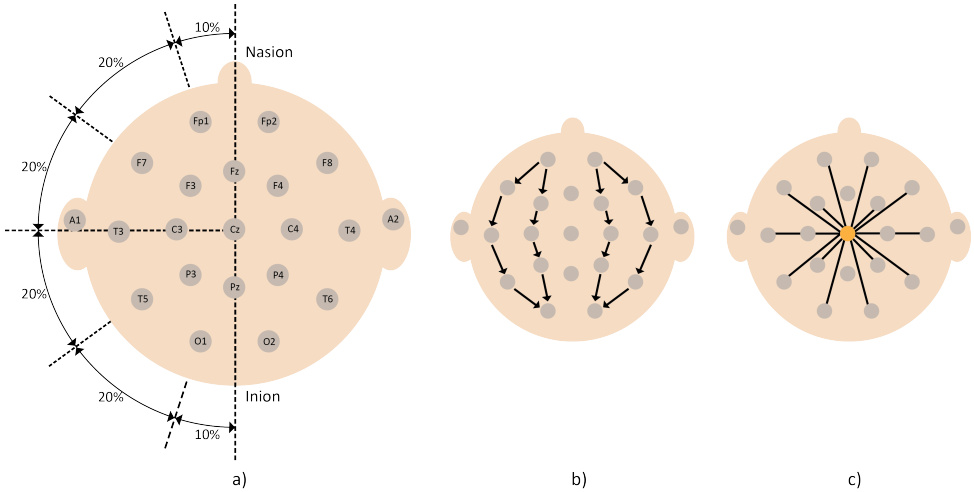

Figure 2.4: EEG electrodes placement maps. In (a) is shown the positions and standard names of the electrodes. In (b) is presented a bipolar montage and in (c) a referential montage. Adapted from: [3].

Montages

EEG signal morphology depends on the chosen montage, i.e., how a particular EEG electrode is referenced. Normally it is preferable to choose a small set of montages to the familiarization process to become faster and allowing a more efficient EEG interpretation [32]. Nowadays, the digital recording of EEG enables changing the montages during the reading process [29].

Bipolar and referential montages are two different strategies to record EEG signals. In an approach, EEG channels are disposed of in a chain of electrodes, where the second input is the next electrode in line. Each channel displays the voltage difference between two adjacent electrodes [27, 29, 32]. Figure 2.4 b) shows an anterior-posterior longitudinal montage, better known as "double banana", one

of the most commonly used [3, 27]. In c) is represented a referential montage, where the second input of each channel is a reference electrode placed either on the scalp or in other body parts (nose, ear lobe, chin...) [27, 29, 32].

2.2.2 Intracranial EEG (iEEG)

As an alternative to scalp EEG where electrodes monitor electrical activity from outside the brain, iEEG can be performed to record the activity from cortex or deeper structures (such as subcortical systems), thereby the electrodes are placed directly on the brain.

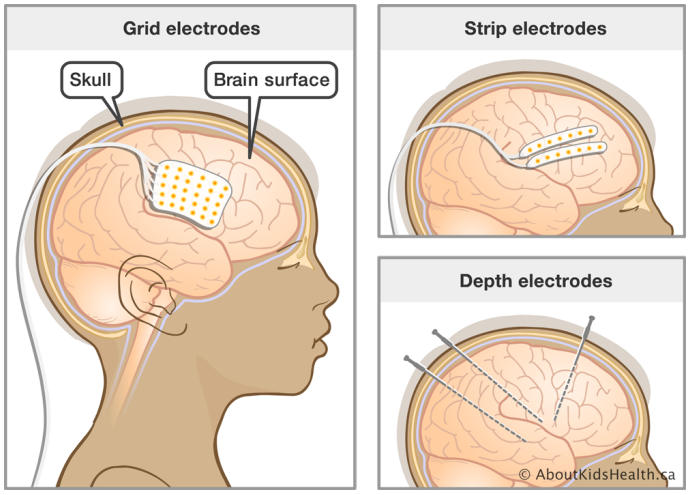

Figure 2.5: Different types of iEEG. Extracted from: [4]

Subdural electrodes are placed on the surface of the brain (cortex) and can be strips or grids. These electrodes can record electrical activity from many points but grids can cover larger areas of the brain (see Fig. 2.5). On the other hand, depth electrodes are thin and flexible wires capable of recording in several places along it and targeted toward deep lesions inside the cortex. Surgery is needed to place these invasive electrodes. This method is only used for pre-surgical evaluation when non-invasive electrodes cannot localize the seizure onset zone. iEEG is very effective for pre-surgical assessment of temporal lobe epilepsy (TLE) [3, 23, 33–35].

Scalp EEG and iEEG comparison

Scalp EEG is an accessible technique in terms of both procedure and cost. It is used as a non-invasive diagnostic tool, sometimes providing enough information to the physician [3]. Despite all the advantages, it has some limitations. The signal is attenuated since its propagation occurs by multiple layers each affecting it differ-

ently. This type of recording has more noise including several extra-cranial artifacts mentioned above, cannot be used to record deep structures, and also does not enable an accurate record, in the beta and gamma sub-bands, of faster frequencies [3, 27]. Besides, for continuous recording over time, it brings some discomfort to the patient and makes it difficult to maintain the skin-electrode interface [27].

Scalp EEG is usually performed for surgical monitoring. This acquisition method may not always be able to detect the seizure focus and it can also lead to inconclusive results. In these cases, iEEG can be used [35]. In spite of being an invasive procedure, iEEG has a higher signal-to-noise ratio (SNR) since the signal is less contaminated by artifacts, records smaller spatial scales due to the greater proximity to seizure onsets and has fewer issues with reference electrodes. Furthermore, although the scalp EEG covers the entire head and the iEEG only a certain region, in the area where the iEEG is being acquired, the spatial resolution is good. The problem is that there is only knowledge of that specific area and not of other surrounding areas. However, this technique involves a difficult procedure and can lead to infections and intracranial hematomas. It also lacks a standard methodology for the electrode placement since that decision depends on case by case [3, 27, 35].

2.2.3 EEG division into periods

Besides allowing to differentiate an epileptic patient from a normal one, EEG records also make it possible to distinguish the different events in time associated with a seizure, due to its own characteristics and patterns. It can be characterized in different periods: the pre-ictal is the period prior to seizure onset, the ictal period corresponds to the seizure, the post-ictal is the period after a seizure and the inter-ictal is the stage between consecutive seizures (see Figure 2.6) [5, 36].

Seizure onset is the moment when a seizure begins. There are two types of onsets, the clinical that corresponds to the physical symptoms and the EEG onset related to changes in brain dynamics. The EEG onset happens before the clinical onset. The determination of clinical signs can be ambiguous since not all types of seizures have identifiable symptoms. Consequently, the onset we are interested in is the one based on the EEG[15].

Although there are several studies that report the existence of a transition phase prior to the seizure onset, there is still a great lack of information about the pre-ictal period, as it is associated with very complex brain dynamical mechanisms manifesting itself in different ways. The pre-ictal period varies from patient to patient and from seizure to seizure, so there is no pattern associated with it hampering its detection in seizure prediction [13, 14, 36].

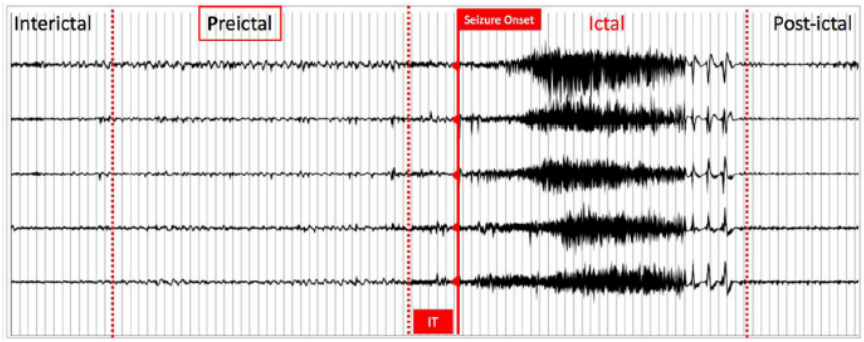

Figure 2.6: Periods of a seizure episode represented on an EEG signal. Source: Iasemidis 2003 [5].

The existence of pre-ictal is an assumption of the seizure prediction area and is seen as the point from which there is no return, that is, the brain can no longer defuse a seizure. This view is very consensual, although limited. Some authors assume a period of susceptibility to seizures, and not a period that necessarily leads to a seizure [14, 16].

2.3 Seizure Prediction

The seizure prediction field has as main goal the construction of a method capable of anticipating a seizure, referring in which time window it will occur later. Thus, an alarm is issued with an associated occurrence window and a guaranteed minimum time for the patient to intervene.

Seizure prediction and seizure detection

It is important to understand the distinction between seizure prediction and seizure detection, since they have different applications. Seizure prediction intends to recognize the presence of seizures in advance in order to allow seizure control. The seizures can be foreseen by the tracking of the pre-ictal state well before the onset time. Seizure detection is important both to discover the seizure onset and to find the epileptic focus. It does not provide enough time for intervention, just a few seconds. Although, detecting a seizure is simpler than prediction for this reason the performance for detection is better [15, 37].

2.3.1 Seizure prediction characterization

M. winterhalder et al.[6] suggested, in 2003, the seizure prediction characteristic, based on clinical and statistical considerations, to facilitate comparison and

evaluation of methods. Seizure prediction approaches aim to develop algorithms for clinical application, in on-line conditions, to be able to integrate warning systems and immediate intervention systems [5, 6, 37]. Since anticipating models cannot determine the exact time when a seizure will occur, they propose two concepts to handle this uncertainty: Seizure Occurrence Period (SOP) and Seizure Prediction Horizon (SPH) [6].

The models can allow minimizing seizure effects by identifying the pre-ictal period of a seizure and raising an alarm. The alarm is associated with a certain SPH, which is the time for intervention, SOP, the period in which the seizure is expected to occur (see Fig. 2.7) [6].

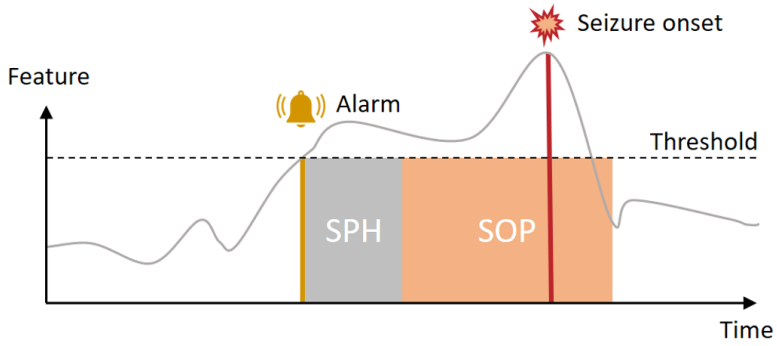

Figure 2.7: Practical definition of SPH and SOP. Adapted from [6].

In an accurate prediction, the seizure onset must occur during the SOP and not during the SPH [6].

Optimum values for SOP and SPH have not yet been found since the pre-ictal time varies from patient to patient and from seizure to seizure but also because each patient may have different preferences regarding these periods. The SPH must have a minimum time to allow intervention (SPH_{min}). Likewise, there must also be a maximum limit for the SOP (SOP_{max}). Authors commonly fix its value from minutes to hours, depending on the studies. When the goal is to alert patients to seizures, the high SOP values lead to greater emotional distress and anxiety. Long SOPs also make continuous intervention difficult, as for example in the case of electrostimulation systems. The values adopted for both parameters have a significant impact on the performance of the classifiers [5, 6, 8, 27].

2.3.2 Performance

Performance must be evaluated with data that was never seen by the trained models. The performance of predictive methods is evaluated by the Sensitivity (SS)

and False Positive Rate per Hour (FPR/h) [6]. To calculate them, it is necessary to distinguish a correctly generated alarm from an incorrect one [8].

A false alarm occurs when it is triggered during any non-pre-ictal period, meaning it occurs in SPH or in the inter-ictal period. A true alarm occurs when an alarm is raised during pre-ictal period (see Fig. 2.8) [27]. It is common for authors to choose for SOP the duration of the pre-ictal period, so if the seizure onset does not occur during a SOP, the alarm has to be classified as a false prediction [6, 8, 38].

Sensitivity (2.3) is a metric adopted by most studies and consists of the proportion of seizures correctly predicted among all seizures.

$$Sensitivity = \frac{Predicted\ seizures}{All\ seizures} \tag{2.1}$$

To evaluate specificity, in the context under study, the FPR/h is a more appropriate measure. This measure quantify the number of false predictions in a certain time [6, 8]. Nevertheless, there are several definitions for FPR/h according to the studies assumptions [27]. Generally, FPR/h is the number of false alarms divided by the inter-ictal period. When considering a refractory period, there is a certain time after an alarm (SPH+SOP) when it is not possible to raise a new alarm. For this reason, that time is removed in order to account for the period of time during which it is possible to raise an alarm. FPR/h can be calculated with (2.2):

$$FPR/h = \frac{FP}{Inter\text{-}ictal\ duration - FP \times (SPH + SOP)} \tag{2.2}$$

The two parameters must be evaluated together, due to their interdependent relationship. If we adjust the algorithm in order to increase the sensitivity, the FPR/h will also be influenced [6]. This is exemplified in Fig. 2.8.

In the literature, when the objective is to simply warn the patient of possible seizures, it is assumed for pre-surgical patients an $FPR/h_{max} = 0.15$ (3.6 seizure per day). When there is anticonvulsant drugs administration, patients with pharmacorefractory focal epilepsy have about 3 seizures per month. In this case, an FPR/h of 0.15 is unsuitable, as 3 seizures per month means 0.0042 seizure per hour, thus no less than 50% of the alarms would be false if all seizures were predicted [6].

No reference FPR/h value has yet been adopted however a maximum value should be set to minimize the negative impact on the patient's life. For a patient warning application regarding seizures, a high FPR/h can lead to two situations: disbelief in relation to the alarms generated and the lack of preparation when a seizure actually occurs, or an increase in psychological stress if all alarms are seri-

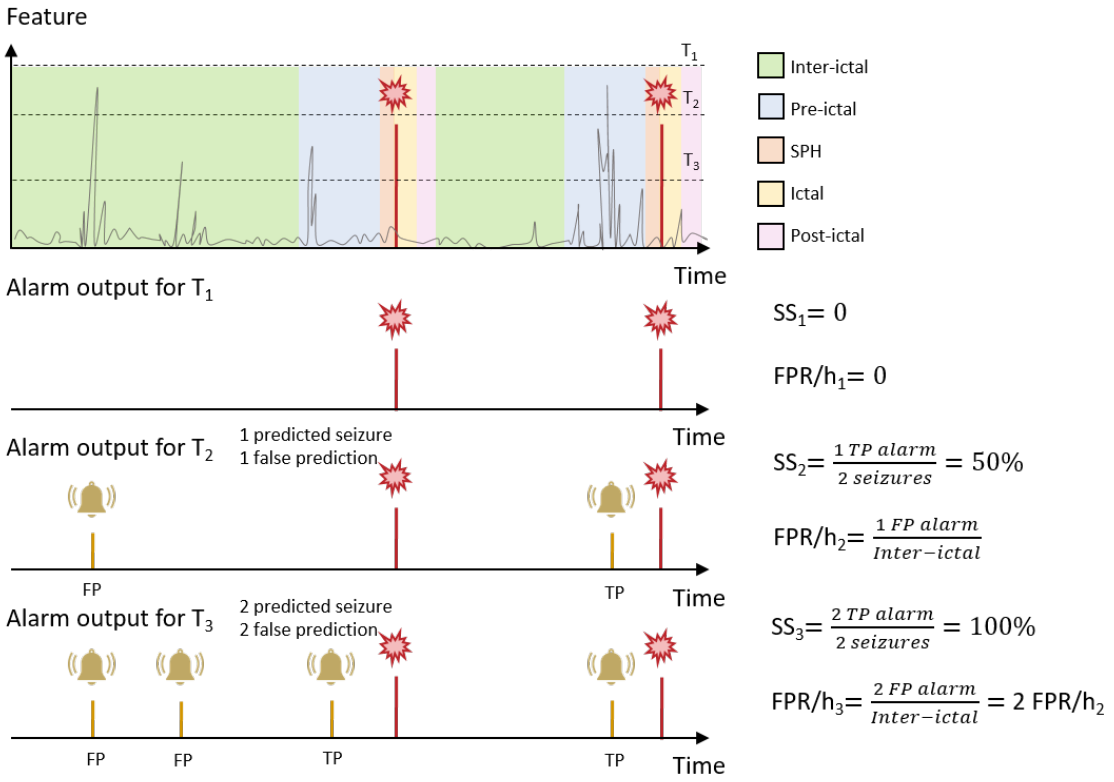

Figure 2.8: Example of an extracted feature used by a seizure prediction method. There are three thresholds (dashed lines) to trigger alarms and they illustrate the dependency between SS and FPR/h. Seizures onsets are marked by the burgundy vertical lines. For the first threshold (T_1) no alarm was raised. T_2 raises two alarms but only have a correct prediction for the second seizure. For the lowest threshold T_3, there is a production of two false alarms but the two seizures are predicted. Decreasing the threshold, despite allowing the correct prediction of the two seizures (SS = 100%), the number of false alarms and consequently the value of FPR/h increases. In a prediction method it is necessary to evaluate both metrics and find the best balance between them. Adapted from: [6].

ously taken into account. For an intervention system, such as electrostimulation or medication, a high number of false alarms can cause side effects due to the excess of electrical stimulation of the brain or over-medication [6].

The performance presented above is the performance of a seizure forecasting system. This is achieved by working on the decisions of a binary Machine Learning model over time. For this, the training samples are classified into two classes, inter-ictal (0) and pre-ictal (1). After the algorithm is trained, the model obtained is applied to test data. The following confusion matrix (Fig. 2.9) shows the relationship between the actual data and the model output (predicted data).

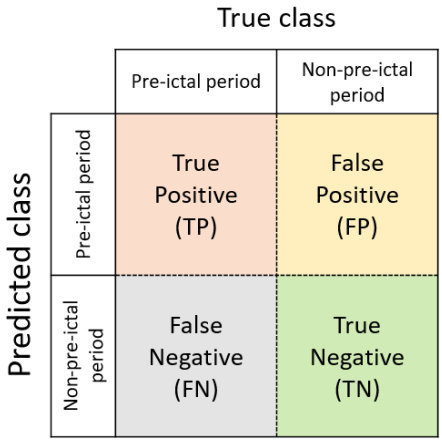

Figure 2.9: Confusion matrix for evaluating Machine Learning performance.

The confusion matrix is a performance measurement for Machine Learning classification. Nevertheless, it should be noted that this classification is carried out sample by sample, which is not the case with the measures previously presented for the classification of alarms. For the evaluation of Machine Learning algorithms the following measures (equations 2.3 and 2.4) can be used.

$$Sample\ Sensitivity = \frac{TP}{TP + FN} \tag{2.3}$$

$$Sample\ Specificity = \frac{TN}{TN + FP} \tag{2.4}$$

2.3.3 Statistical Evaluation

Given the complexity of the seizure prediction, the impossibility to have really long-term recordings, and the fact that there is a trade-off between sensitivity and FPR/h, in addition to the calculation of metrics for performance evaluation, statistical validation should also be implemented to understand if the proposed methods perform above chance level. For this reason, a comparison method is used. It is required that the algorithm performance outperforms a random one. Otherwise, the developed method will not be predicting seizures [6, 8, 15]. Although there are several techniques, the most used are the random predictor [7, 38] and the surrogate time series analysis [7, 39, 40].

2.3.3.1 Random prediction method

A random predictor triggers alarms randomly without resorting to EEG information [6, 7]. Schelter et al. [7, 38] proposed an analytic random predictor based

on a homogeneous Poisson process for false predictions. Therefore, the probability of raising an alarm at each sample point of the time series of N samples is given by (2.5):

$$P_{Poisson} = \frac{FP}{N}. \tag{2.5}$$

Considering a time interval equal to SOP, the probability of raising at least an alarm in that interval is given by (2.6):

$$P \approx 1 - e^{(-FPR/h_{max}) \times SOP} \approx (FPR/h_{max}) \times SOP \tag{2.6}$$

When taking into account more than one seizures, the probability of randomly predict k of K seizures follows a binomial distribution with probability P. Considering also several electrodes and multiple features, the probability is described by (2.7):

$$P_{binom,d}(k; K; P) = 1 - \left[\sum_{j<k} \binom{K}{j} P^j (1-P)^{K-j} \right]^d \tag{2.7}$$

where the degrees of freedom is $d = 1$ if only a unique predictor is used with the decision of being an alarm or not [41].

For a given significance level α, it is possible to determine the critical value of sensitivity for the random predictor (2.8):

$$\sigma_{rand} = \frac{max_k\{P_{binom,d}(k; K; P) > \alpha\}}{K} \times 100\% \tag{2.8}$$

This value provide information about the minimum number of seizures that the algorithm have to predict to ensure that performance above chance level can be shown. However, its concept is based on the homogeneous Poisson distribution for false predictions, which may not be correct [38, 41].

2.3.3.2 Surrogates time series

Schelter et al. [7] suggested an alternative approach to analytic random predictors, the bootstrapping techniques. This method uses the concept of Monte Carlo simulations and perform constrained randomizations of the original seizure predictor to construct surrogate seizure predictors. It allows testing several null hypotheses depending on the set of assumptions and restrictions. Only if the performance of the original seizure predictor is noticeably better than the one of the surrogate predictor, the null hypothesis can be rejected.

Andrzejak et al [39] introduced the seizure-times surrogates. This method consists of randomly shuffle the inter-seizure intervals maintaining however the measures profiles. Thus, it is generated artificial seizures onset times under the constraints of preserving the distribution of intervals between consecutive seizures, the total number of seizures and the clustering of the seizures. The number of seizures and the presence of gaps in the EEG records can make it difficult to generate a sufficient number of independent surrogates needed to obtain significance [7, 15, 39, 40]

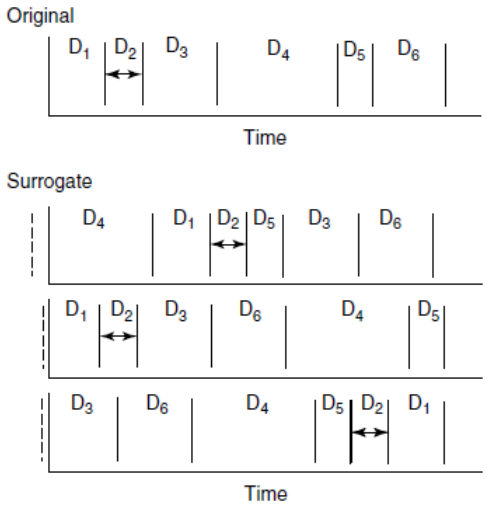

Figure 2.10: Original seizure times and the surrogate times bootstrapped from the interseizure intervals. The arbitrary onset times for the surrogates are obtained from a uniform distribution and are indicated by the dashed vertical lines. Source: Schelter et al. 2008 [7].

Comparison between the two approaches

The surrogate predictor has the advantage of being more flexible, adaptable and taking into account several crucial parameters to the study. However, this method is slower and require greater care in implementing to avoid biasing the null hypotheses. While the analytic random predictor allows for almost immediate calculation of critical values, it does not take into account the SPH value in its definition. Therefore, in this case, having an SPH of 10 seconds or 1 minute does not change its value, but a SPH of 1 minute changes the difficulty of the problem. Thus, some authors consider the surrogate predictor to be more robust. In addition, both techniques are subject to assumptions, however those of the analytic random predictor are more risky. When the homogeneous Poisson distribution is not applicable to the case under study, bootstrapping techniques can be a solution [7, 13].

2.3.4 Concept drift and data imbalance

In real-world problems, as is the case of epilepsy, the data often has class imbalance and a non-stationary distribution that results in the concept drift. For an adequate forecast, the algorithm must have a general structure capable of adapting and handle such scenarios [42].

The distribution of data may depend on a particular context. Changes in this context can induce changes in the data, the so-called concept drifts. The state of the brain varies depending on external and internal factors, such as circadian rhythm, psychological states (such as stress and anxiety), medication, daily tasks, vigilance states, among others. These changes, in turn, can influence the predictive features of seizures [13, 14, 43].

In the EPILEPSIAE database and in almost all of the other databases, the EEG recordings belong to patients under pre-surgical monitoring where the anti-epileptic medication was suppressed. However, the daily life situation is quite different from pre-surgical context, influencing the neurological state differently. This is observed even by the average number of seizures in both conditions, which varies from several seizures per day to several in a month. Another problem is the fact that although the medication is withdrawn, the effect of it does not disappear immediately, with a visible increase in seizure frequency over time. In addition, people in pre-surgical condition are mostly seated or lying down, activities very different from those of normal daily experience. Because of this, most of the tested methodologies with this type of databases constitutes only a proof-of-concept that will have to be tested, when possible, with data collected under realistic conditions [6, 14, 44].

There is also the problem of data imbalance, since periods under seizures (ictal period) are severely underrepresented in the dataset. Seizures are rare and therefore, the pre-ictal period is significantly shorter than the inter-ictal period. Summarily, handling concept drift and data imbalance is challenging but important for a robust clinical application [42].

3

State of the art

This chapter presents the current state of the art in seizure prediction. The first section (3.1) presents an introduction to the seizure prediction area by providing a summary of the historical context and an overview of the most important framework steps of studies. Sections 3.2 to 3.8 provide a more detailed description of the most common techniques.

3.1 Overview

The apparently unpredictable nature of seizures makes it difficult to find a good prediction method. However, several efforts have been made in the last 50 years for a correct detection of the pre-ictal period and consequently an accurate forecast of seizures onset [8, 13].

Historical context

The first studies on seizure prediction date back to the 1970s. Early works used linear approaches, such as autoregressive models [8, 15]. Immediately afterward, some researchers start introducing the concept of pre-ictal phenomenon. Then there were studies based on non-linear time series analysis [15]. Later, proof-of-principle works appeared introducing inter-ictal periods into their focus of interest, enabling comparisons with pre-ictal periods [8, 15]. Around the year 2003, the lack of reproducibility of studies and their statistical validation started to be discussed [8, 13]. With the development of technology and consequently of the data storage capacity and computational power, it was possible to record long-term data [13, 14]. With the appearance of such records, it was found that the problem was still far from being solved, that is, the performance of the studies was much lower than those reported [8, 15]. Therefore, in recent years, studies based on Deep Learning techniques have emerged [14].

Framework

Most studies follow a common framework consisting of several stages (Fig. 3.1), although there may be some variations.

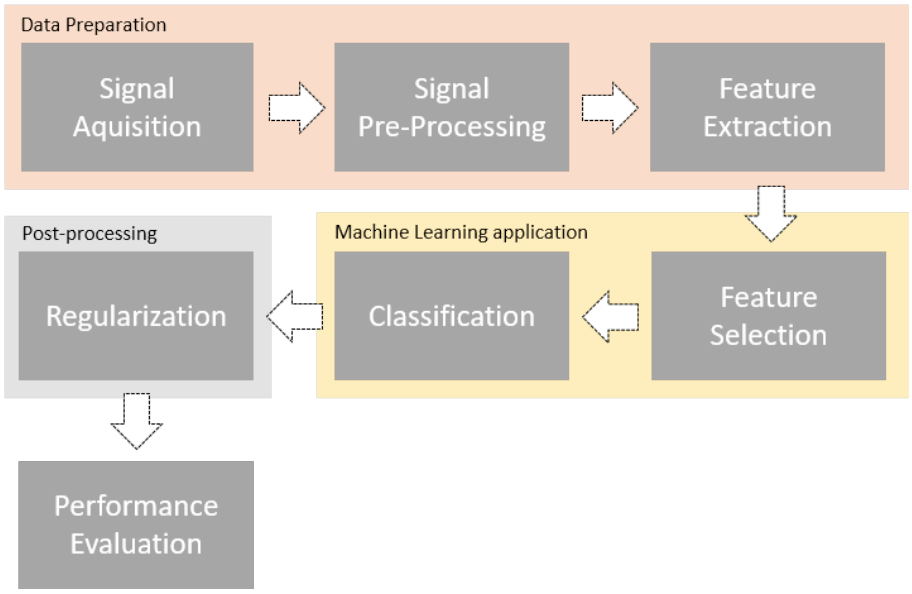

Figure 3.1: Common framework of prediction algorithms. Adapted from [8].

- Signal Acquisition: gather of Electroencephalogram (EEG) information to be used as a study tool.
- Signal Pre-processing: preparation of the collected information for feature extraction. It aims to improve the quality of the EEG signals in order to highlight the components of interest.
- Feature Extraction: collection of characteristics of the records
- Feature Selection: identification the most useful features to characterize the transition for pre-ictal.
- Classification: detection of the pre-ictal changes from the selected features using classifiers. In the last decades employing Machine Learning (ML) models.
- Post-Processing: usually consists of a regularization step.
- Performance: evaluation and testing of the data not used for training the model [3, 8, 45].

Latest studies employ Deep Learning techniques [10, 19, 46–48]. Such approaches can present some changes in the presented framework since they allow feature extraction from the data without explicit knowledge of its structure. Therefore, it is not necessary to have independent pre-processing and feature extraction steps. Furthermore, these algorithms can handle raw data, missing data, and time

implicitly. However, due to overfitting problems, it requires very large amount of data to have a better performance than other techniques. As a result of their great complexity, those approaches create black box systems, which are difficult to interpret clinically [14, 49].

The adoption of this structure does not imply that all studies are similar since the parameters and methods can significantly vary. In fact, there is a large variety of studies with different approaches, thus making comparison difficult [8].

For the reasons mentioned above, the framework used in this thesis will follow the pipeline presented in Figure 3.1.

3.2 Signal acquisition

The first step involves the collection and selection of data to be used in the study. Both the type of EEG signal and the database influence the results and conclusions of the different studies [8]. Table 3.1 contains the signal acquisition parameters for some of the studies in recent years. Both the total number of hours of EEG and the number of seizures are higher in more recent studies than in older ones. For this reason, in the first studies it was often difficult, if not impossible, to provide good estimates of specificity for inter-ictal periods [8, 44]. There are also studies performed with EEG records of healthy and epileptic patients [50, 51] and further studies using recordings from animal models [52, 53].

3.2.1 Databases

As mentioned before, first studies in seizure prediction were based on limited data in terms of periods before seizure, number of patients and seizures [8, 13]. Over time, scientists realized the need for long-term records, with a greater number of patients and seizures to face the problems of reproducibility and comparison between studies [44]. Figure 3.2 shows a comparison between some databases used in seizure prediction.

New databases were created, such as EPILEPSIAE and NeuroVista. NeuroVista database is the largest in terms of recordings duration per patient. Despite only having data from 10 patients, they were followed for a period of up to two years. However, this type of long-term tracking is very complicated to perform, due to ethical and logistic issues. EPILEPSIAE comprises a larger number of patients even though it only has a few weeks of data for each. This database allowed studies to be carried out with a larger number of patients, as is the cases of Direito et al.

Table 3.1: Signal acquisition parameters for some of the studies in recent years.

Study	Database	Patients	Total EEG (h)	Seizures	Type of EEG	Electrodes
Agarwal et al.[48] (2018)	Kaggle (U. Pensylvania and Mayo Clinic)	8	N.A.	N.A.	Invasive	All
Chamseddine et al.[19] (2018)	Kaggle (American Epilepsy Society)	1 dog	85	N.A.	Invasive	16
Sun et al.[47] (2018)	Kaggle (American Epilepsy Society)	5 dogs 2 humans	N.A.	N.A.	Scalp + Invasive	16
Khan et al.[46] (2018)	CHB-MIT	28	N.A.	N.A.	Scalp	22
Kiral-Kornek et al.[10] (2017)	NeuroVista	10	142700	2817	Invasive	16
Aggarwal et al.[50] (2017)	Personal	10 Epileptic 2 Healthy	N.A.	N.A.	Scalp	18 (9 in each hemisphere)
Direito et al.[23] (2017)	EPILEPSIAE	216	16729.8	1206	Scalp + Invasive	3 methods: 6 random F7, Fz, F8, T5, Pz, T6 6 in focal area
Bandarabadi et al.[54] (2015)	EPILEPSIAE	24	3565.4	183	Scalp + Invasive	3 in focal area + 3 far from focus
Bandarabadi et al.[55] (2015)	EPILEPSIAE	18	470	94	Scalp + Invasive	2 in focal area
Bou Assi et al.[52] (2015)	Kaggle (American Epilepsy Society)	5 dogs	N.A.	44	Invasive	16
Rasekhi et al.[56] (2015)	EPILEPSIAE	10	1388	86	Scalp + Invasive	3 in focal area + 3 far from focus
Mporas et al.[57] (2015)	ARMOR project	3	N.A.	N.A.	Scalp + Invasive	21
Teixeira et al.[58] (2014)	EPILEPSIAE	278	>40 000	>2662	Scalp + Invasive	3 methods: 6 random F7, Fz, F8, T5, Pz, T6 6 in focal area
Moghim et al.[37] (2014)	Freiburg	21	N.A.	N.A.	Scalp + Invasive	6
Alvarado Rojas et al.[41] (2014)	EPILEPSIAE	53	12744	558	Invasive	10-124
Rabbi et al.[59] (2013)	EPILEPSIAE	1	36	7	Invasive	N.A.
Rasekhi et al.[60] (2013)	EPILEPSIAE	10	1388	86	Scalp + Invasive	3 in focal area + 3 far from focus
Bandarabadi et al.[61] (2012)	EPILEPSIAE	12	2093	82	Scalp + Invasive	3 in focal area + 3 far from focus
Teixeira et al.[62] (2012)	EPILEPSIAE	10	1916	58	Scalp	2 methods: F7, Fz, F8, T5, Pz, T6 6 in focal area
Direito et al.[63] (2012)	EPILEPSIAE	10	497.3	30	Scalp	All
Valderrama et al.[64] (2012)	EPILEPSIAE	10	3178	108	Scalp + Invasive	All
Acharya et al.[51] (2012)	Boon	5 Epileptic 5 Healthy	N.A.	N.A.	Invasive	128
Direito et al.[65] (2011)	EPILEPSIAE	3	468.6	32	Scalp	Selected based on the power of the feature
Park et al.[66] (2011)	Freiburg	18	N.A.	80	Invasive	3 in focal area + 3 far from focus
Chisci et al.[67] (2010)	Freiburg	9	N.A.	N.A.	Invasive	3 in focal area + 3 far from focus
Mirowksi et al.[68] (2009)	Freiburg	21	N.A.	N.A.	Invasive	3 in focal area + 3 far from focus
Mormann et al.[69] (2005)	Personal	5	311	46	Invasive	All
Kreuz et al.[40] (2004)	Bonn	1	N.A.	10	Invasive	2 depth electrodes (each with 10 contact surfaces)
D'Alessandro et al.[70] (2003)	Personal	4	160	46	Invasive	All
Mormann et al.[71] (2003)	Bonn	18	117	32	Invasive	2 depth electrodes (each with 10 contact surfaces)
Mormann et al.[72] (2003)	Bonn	10	31	14	Invasive	2 depth electrodes (each with 10 contact surfaces)
Le Van Quyen et al.[73] (1999)	Personal	13	15	23	Invasive	32
Geva et al.[53] (1998)	Personal	25 rats	N.A.	N.A.	Scalp	2

[23] and Teixeira et al. [58] studies, with more than 200 patients.

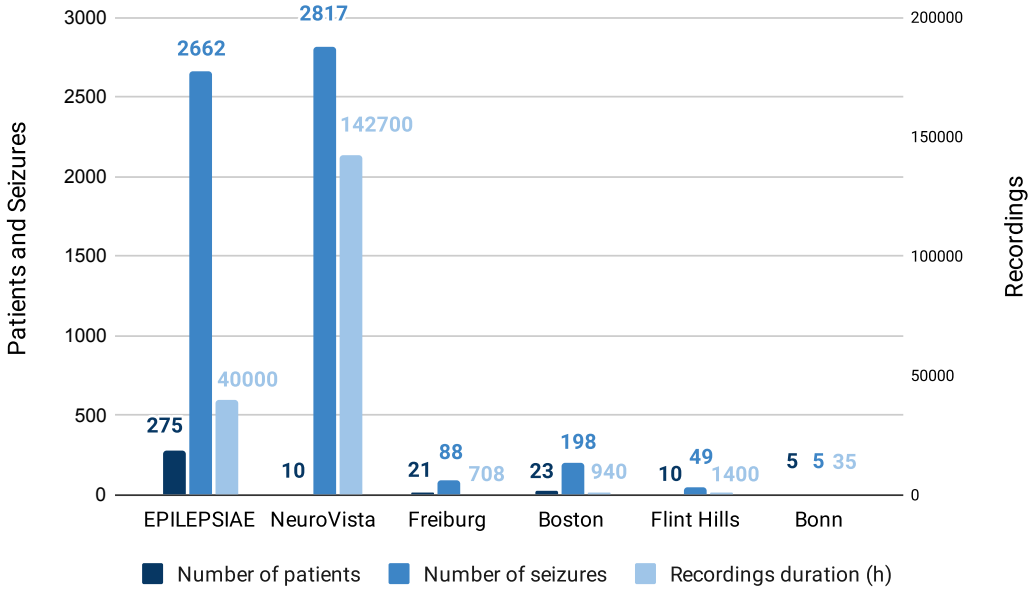

Figure 3.2: Databases comparison in relation to the number of patients, number of seizures and duration of records (in hours). Sources:[8–10].

Furthermore, EPILEPSIAE database contains metadata information and standard annotations. The stored metadata are about technical and clinical aspects, such as patient's age, epilepsy characteristics, medication, type of electrodes. Examples of standardized annotations are seizure dominant pattern and type, state of vigilance, clinical onset and offset.

The database is comprised of 275 pre-surgical monitoring patients registered in University Hospital of Coimbra, Portugal, University Hospital Freiburg, Germany and Hospital de la Pitier-Salpetriere in Paris, France. Although being possible to find Intracranial EEG (iEEG) and Scalp EEG recordings, the latter are in majority. On average, the recordings last 165h and the number of seizures per patient is 9.68. Sampling rates ranging from 250 to 2.5kHz are available. Furthermore, other types of signals such as Electrocardiography (ECG) and Electromyography (EMG) were also included [44].

3.2.2 EEG type and electrodes

Regarding the type of signal, both the scalp EEG and intracranial EEG have been used. Several studies reveal that there are no significant differences in performance between the two types of EEG [54, 58, 60]. However, for application in

clinical prediction devices iEEG is apparently more suitable [8].

The electrodes chosen vary both in number and location. There are studies that choose both electrodes locations, near the focal area and far from focus, in order to find relationships between places where there are stronger pre-ictal traces and places where there is more information about the general brain state [54, 56, 60, 61, 66–68]. There are also studies that focus on electrodes in the focal zone [23, 55, 58, 62], others on random electrodes assuming that the seizure generation process occurs in all parts of the brain [23, 58]. In addition, there are approaches that try to maximize scalp coverage with a smaller number of electrodes, selecting them (e.g. F7, Fz, F8, P7, Pz, P8) [23, 58, 62]. Certain authors choose to use all the information available, using all electrodes [63, 64, 69, 70]. However, it should be noted that a high number of channels may not be suitable for a portable system due to the increased complexity of analysis and computation [60].

3.3 Signal Pre-processing

This stage includes the preparation of the raw EEG signals, to highlight the components of interest, for feature extraction [3]. For a better analysis of the EEG, the signal must be treated in order to minimize the impact of unwanted artifacts and noise. In addition, it is necessary to decide how the signal will be segmented. A sliding window analysis is usually performed. Then, one of the most important steps for the prediction of seizures is the definition of a duration (or several) for the pre-ictal period and Seizure Occurrence Period (SOP) and also for Seizure Prediction Horizon (SPH). Table 3.2 summarizes the pre-processing decisions made in the studies in recent years.

3.3.1 Filters and artifact removal

This step involves the removal of unwanted aspects, such as noise, artifacts and high frequency content [3]. The interference from the power supply is usually removed with a band-rejection filter (notch) of 50Hz [23, 50, 52, 54, 56, 58, 60, 62, 63, 66–68] or 60Hz [19, 53, 60, 70].

To eliminate high and low frequencies, a band-pass is usually applied. For this, a low-pass with the highest cut-off frequency can be applied followed by a high-pass with the low cut-off frequency. The cut-off frequencies vary depending on the bands of interest, but generally the lowest frequency is 0.5Hz [50, 52, 59, 60, 66, 68, 71–73]. Digital filters, such as Infinite Impulse Response (IIR) and Finite Impulse Response

Table 3.2: Signal pre-processing parameters for some of the studies in recent years.

Study	Filters	Sliding window	Pre-ictal (min)	SPH (min)
Agarwal et al.[48] (2018)	1 - 47Hz bandpass	1s (without overlap)	N.A.	N.A.
Chamseddine et al.[19] (2018)	0 - 190Hz bandpass 60Hz notch	5s (without overlap)	60	N.A.
Sun et al.[47] (2018)	0.1 - 180Hz bandpass	30s (without overlap)	60	5
Khan et al.[46] (2018)	0 - 128Hz bandpass	1s (without overlap)	10	N.A.
Kiral-Kornek et al.[10] (2017)	N.A.	N.A.	15	N.A.
Aggarwal et al.[50] (2017)	0.5 Hz - 40 Hz bandpass 50Hz notch	5s (without overlap)	60	4-10
Direito et al.[23] (2017)	50Hz notch	5s (without overlap)	10, 20, 30, 40	1/6
Bandarabadi et al.[54] (2015)	50Hz notch	5s (without overlap)	10, 20, 30, 40	N.A.
Bandarabadi et al.[55] (2015)	Outlier removal and normalization	8s (50% overlap)	5 - 180	N.A.
Bou Assi et al.[52] (2015)	0.5 - 180Hz bandpass 50Hz notch	5s (without overlap)	60	5
Rasekhi et al.[56] (2015)	50Hz notch	5s (without overlap)	10, 20, 30, 40	N.A.
Mporas et al.[57] (2015)	N.A.	1s (without overlap)	N.A.	N.A.
Teixeira et al.[58] (2014)	50Hz notch	5s (without overlap)	10, 20, 30, 40	1/6
Moghim et al.[37] (2014)	Artifact removal	5, 9 and 180s (without overlap)	5	N.A.
Alvarado Rojas et al.[41] (2014)	filtered in the bands of interest (0.5 - 140 Hz)	1min (without overlap)	60	N.A.
Rabbi et al.[59] (2013)	0.5 - 100Hz bandpass 60Hz notch	10s (50% overlap)	15, 30, 45	N.A.
Rasekhi et al.[60] (2013)	50Hz notch	5s (without overlap)	10, 20, 30, 40	N.A.
Bandarabadi et al.[61] (2012)	N.A.	5s (without overlap)	10, 20, 30, 40	N.A.
Teixeira et al.[62] (2012)	50Hz notch	5s (without overlap)	10, 20, 30, 40	N.A.
Direito et al.[63] (2012)	50Hz notch	5s (without overlap)	2-60	N.A.
Valderrama et al.[64] (2012)	N.A.	5s (without overlap)	5, 10, 20, 30, 45, 60	N.A.
Acharya et al.[51] (2012)	N.A.	23.6s (20% overlap)	N.A.	N.A.
Direito et al.[65] (2011)	N.A.	5s (without overlap)	30, 40	N.A.
Park et al.[66] (2011)	Artifact removal 50Hz and 100Hz notch	20s (50% overlap)	30	N.A.
Chisci et al.[67] (2010)	50Hz notch filter	2s (without overlap)	50	N.A.
Mirowksi et al.[68] (2009)	0.5 - 120Hz bandpass 50Hz notch	5s (overlap N.A.)	50	N.A.
Mormann et al.[69] (2005)	N.A.	17 - 20.5s (without overlap)	5, 30, 120, 240	N.A.
Kreuz et al.[40] (2004)	0.5 - 85Hz bandpass	20.48s (without overlap)	240	N.A.
D'Alessandro et al.[70] (2003)	0.1 - 100Hz bandpass 60 Hz notch	10s (75% overlap) and 60s (96% overlap)	10	N.A.
Mormann et al.[71] (2003)	0.5 - 85Hz bandpass	23.6s (20%overlap)	10 - 240	N.A.
Mormann et al.[72] (2003)	0.5 - 85Hz bandpass	23.6s (20%overlap)	10 - 105	N.A.
Le Van Quyen et al.[73] (1999)	0.5 - 99Hz bandpass	25s (without overlap)	20	N.A.
Geva et al.[53] (1998)	1 - 30Hz bandpass 60Hz notch	1s (50% overlap)	5	N.A.

(FIR), are commonly used [8].

A visual analysis of the signal can also be made to remove bad segments, such as motion artifacts, so as not to affect the next phase [66].

However, it should be noted that due to the complexity of the EEG, it becomes difficult to understand what are the best measures to be taken regarding filtering and artifact removal. For this reason, there is a great heterogeneity of choices in the studies.

3.3.2 Sliding window analysis

The segmentation of the EEG is taken through a sliding window analysis where the windows have a certain length and percentage of overlap. These windows must

be of an adequate size to extract relevant EEG characteristics. It is important to maintain a balance between the ability to capture specific patterns and assumptions of stationarity, and also take into account its application in online systems [3, 8].

There is no ideal standard length to be applied in studies, so its value varies depending on the study. The same applies to the overlap values. However, windows with a duration of less than 1 min are usually applied, being the 5-s window the most adopted. Concerning the overlap value, it is often discarded. Moghim et al. [37] compared two windows of different lengths, the short-term windows with 9s (STE) and the long-term windows with 180s (LTE), and concluded that the STE features had lower predictive value than LTE features. D'Alessandro et al. [70] also used two windows with different lengths and overlapping values to extract the features of different levels.

3.3.3 Definition of pre-ictal duration, SOP and SPH

Furthermore, in the pre-processing phase, it is necessary to define the length of the pre-ictal period. Some studies define a fixed pre-ictal time of 5 [53], 10 [46, 70], 30 [66], or even of 60 minutes [19, 41, 47, 50, 52], others consider several values for the pre-ictal time [23, 54, 56, 58–62, 64, 69]. The definition of the pre-ictal is important when supervised learning methods are used.

No optimal value has yet been defined for the duration of the pre-ictal period, hence the values used in the literature are diverse. Some studies have tried to find an optimal pre-ictal value as the case of Bandarabadi et al.[55]. In this study, scientists varied the pre-ictal time between 5-180min and concluded that the optimal pre-ictal value varied between seizures and from patient to patient.

There is also no standard value for SPH. This value may vary depending on the final application. In the case of a warning system, an SPH of a few minutes is adequate to allow patient intervention. In the case of an intervention system, a smaller duration may be sufficient. In the literature, SPH varies from a few seconds [23, 58] to a few minutes [47, 50, 52]. In many studies, the SPH duration is omitted.

The algorithm must be tested for a range of SOP and SPH values in order to achieve general performance, but this rarely happens. It is common for authors to choose for SOP the duration of the pre-ictal period.

3.4 Feature Extraction

Feature extraction is a key factor to the algorithm performance. It is needed a deep understanding of the subject to select the most appropriate discriminating measures [3].

There is a significant variety of features grouped into four main categories according to the linearity and the number of used channels (see Figure 3.3). No optimal feature type has been determined, although some have a lack of clinical interpretability and non-linear features are computational heavy, difficult to parameterize and apply in real-time [8, 15].

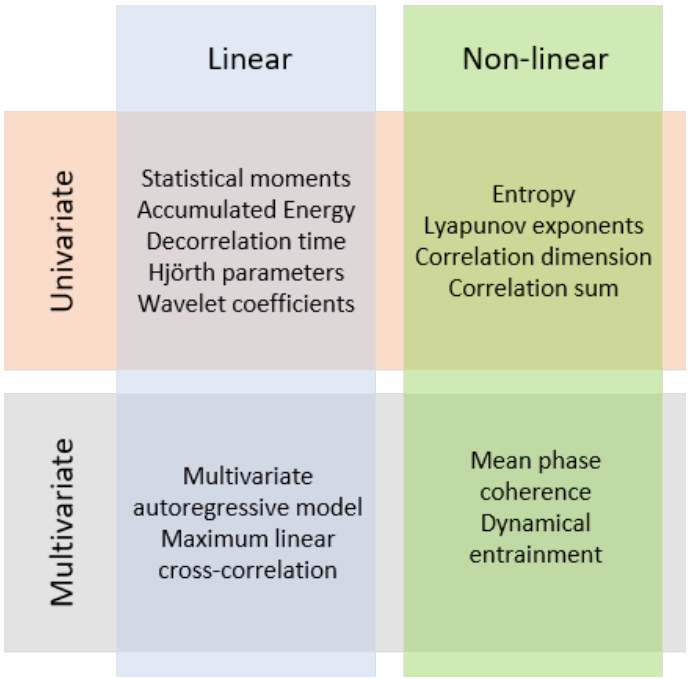

Figure 3.3: Feature categorization according to the linearity and number of channels.

According to the number of used channels, features are denominated univariate (one single EEG channel) or multivariate (several EEG channels). The term bivariate, which is included in the multivariate, is commonly used in the literature and refers to features performed in two channels. Multivariate measures have more information and characterize the relation between the channels. Regarding to linearity, features are classified as linear or non-linear.

Concerning univariate features, linear predominate in the set of studies shown in Table 3.3. As for multivariate, the scenario is already the opposite. Univariate linear measures are better understood and computed quickly since they are lighter. There is also a predominance in the choice of univariate features, probably since

multivariate features need greater computational power.

Table 3.3: Type of features adopted in some of the studies in recent years.

Study	Univariate		Multivariate	
	Linear	**Non-linear**	**Linear**	**Non-linear**
Agarwal et al.[48] (2018)				
Chamseddine et al.[19] (2018)	x			
Sun et al.[47] (2018)	x			
Khan et al.[46] (2018)	x			
Kiral-Kornek et al.[10] (2017)	x			
Aggarwal et al.[50] (2017)				x
Direito et al.[23] (2017)	x			
Bandarabadi et al.[54] (2015)	x			
Bandarabadi et al.[55] (2015)	x			
Bou Assi et al.[52] (2015)	x			
Rasekhi et al.[56] (2015)	x		x	
Mporas et al.[57] (2015)	x	x		x
Teixeira et al.[58] (2014)	x			
Moghim et al.[37] (2014)	x	x		
Alvarado Rojas et al.[41] (2014)		x		
Rabbi et al.[59] (2013)		x		x
Rasekhi et al.[60] (2013)	x			
Bandarabadi et al.[61] (2012)	x			
Teixeira et al.[62] (2012)	x			
Direito et al.[63] (2012)	x			
Valderrama et al.[64] (2012)	x			
Acharya et al.[51] (2012)		x		
Direito et al.[65] (2011)	x			
Park et al.[66] (2011)	x			
Chisci et al.[67] (2010)	x			
Mirowksi et al.[68] (2009)			x	x
Mormann et al.[69] (2005)	x	x	x	x
Kreuz et al.[40] (2004)				x
D'Alessandro et al.[70] (2003)	x	x		
Mormann et al.[71] (2003)				x
Mormann et al.[72] (2003)			x	x
Le Van Quyen et al.[73] (1999)		x		
Geva et al.[53] (1998)	x		x	

Some studies extract different types of features for prediction purposes, but not all make a comparison between them. However, over the years some studies reported some noteworthy points. Mormann et al. [72] compared two multivariate features, one linear and other non-linear, and concluded that both have similar performance. In another study, Mormann et al. [69] evaluated the performance of 30 different features both univariate and multivariate as well as linear and non-linear. They concluded linear features performed similarly or better than non-linear ones and multivariate measures were capable of being sensitive to pre-ictal changes on a longer time scale before seizure onset than univariate measures. Rasekhi et al. [56] noted that multivariate features generate a smaller number of false predictions

comparing with univariate ones.

Deep Learning methods do not require previous extraction of features, since they are capable of employing automatic feature engineering methods from a time series data [14, 49]. However, some studies performed an traditional feature extraction followed by its application in deep classification models [10, 19, 46, 47]. Agarwal et al. [48] opted for extracting features regardless of classification, using Deep Learning (Convolutional Neural Network (CNN)).

A general description of the most adopted features in seizure prediction studies is presented in the following sections. A more detailed characterization is presented in Appendix A.

3.4.1 Univariate Linear features

Univariate linear measures are computed based on the amplitude and frequency information of EEG time series that satisfy the linearity property [15]. However, it is known that the EEG signal is non-stationary so it must be segmented into shorter semi-stationary parts. Table 3.4 presents a feature extraction overview of univariate linear measures in recent years.

Table 3.4: Univariate Linear Features adopted for some of the studies in recent years

Study	Statistical Moments	Energy related	Hjörth parameters	Decorrelation time	Autoregressive modelling	Power related	Wavelet coefficients
Chamseddine et al.[19] (2018)						x	
Sun et al.[47] (2018)						x	
Khan et al.[46] (2018)							x
Kiral-Kornek et al.[10] (2017)						x	
Direito et al.[23] (2017)	x	x	x	x	x	x	x
Bandarabadi et al.[55] (2015)						x	
Bou Assi et al.[52] (2015)			x	x		x	
Rasekhi et al.[56] (2015)	x	x	x	x	x	x	x
Mporas et al.[57] (2015)	x	x			x	x	x
Teixeira et al.[58] (2014)	x	x	x	x	x	x	x
Moghim et al.[37] (2014)	x	x				x	x
Rasekhi et al.[60] (2013)	x	x	x	x	x	x	x
Bandarabadi et al.[61] (2012)						x	
Teixeira et al.[62] (2012)	x	x	x	x	x	x	x
Direito et al.[63] (2012)						x	
Valderrama et al.[64] (2012)	x	x	x	x	x	x	x
Direito et al.[65] (2011)	x	x	x	x	x	x	x
Park et al.[66] (2011)						x	
Chisci et al.[67] (2010)					x		
Mormann et al.[69] (2005)	x	x	x	x		x	x
D'Alessandro et al.[70] (2003)	x	x				x	x
Geva et al.[53] (1998)	x						x

To analyze the dynamics of the EEG signal in the time domain, studies have been adopted several measures. Statistical moments (mean, variance, skewness, kurtosis), which characterize the amplitude distribution of a signal, are the most frequently used features due to their simplicity [23, 37, 56–58, 60, 62, 64, 65, 69].

Energy-related features, such as accumulated energy which assumes that seizure-generation mechanisms induce an increase in brain activity, are also reported in seizure prediction studies [37, 56, 60, 64, 70]. Some studies [23, 52, 56, 58, 60, 62, 64, 65, 69] employ the Hjörth parameters (activity, mobility and complexity) to describe the EEG quantitatively. An increase in mobility and complexity has been reported prior to seizures [8, 52]. In order to analyse regularity on the EEG signal, the autocorrelation function, which attempts to find repeating patterns, can be computed. Decorrelation time, defined as the first zero crossing of a given EEG signal's autocorrelation function, has been reported to decrease prior to seizures [8, 23, 56, 58, 60, 62, 65, 69]. Furthermore, several studies [23, 56–58, 60, 62, 64, 65, 67] mention measures computed from autoregressive models, for example the prediction error, which has been found to decrease as seizure onset approaches.

Signals can also be analyzed concerning the spectral dynamics. Several studies computed the relative spectral power in different frequency bands (delta, theta, alpha, beta and gamma) to explore if the seizure-generation processes change the patterns of brain activity [8, 66, 69]. Mormann et al. [69] found that the spectral power of low frequencies bands, as delta band, decreases in the pre-ictal period while the spectral power increases in the remaining bands. Many of the studies that choose to calculate the relative spectral power of bands also choose the spectral edge frequency and power as one of the features to be evaluated [23, 37, 52, 56, 58, 60, 62, 64, 65, 69]. Wavelet coefficients can also be used for both spectral and temporal analysis of the brain signals [23, 56, 57, 60, 65, 70]. Deep Learning approaches, which opted for univariate linear features, extracted power related features [10, 19, 47] or performed wavelet-transformation on EEG [46].

3.4.2 Univariate Non-linear features

The mechanisms that lead to the epileptic process are still not well understood. So univariate linear measures may not be suitable. Univariate non-linear features consider brain mechanisms are associated with several dynamic states and try to explain the non-stationary nature of EEG signals, noise and intermittency of seizures [15, 74]. Table 3.5 shows the feature extraction overview univariate non-linear measures in recent years.

Non-linear measures, such as correlation sum and correlation dimension [37, 69], Lyapunov exponents [37, 68, 69], dynamical similarity index [59, 73] and entropy measures [51, 57], are based on the theory of dynamic systems also known as chaos theory. Dynamical systems describe the evolution of a system over time through rules which are functions that determine what future states follow from the cur-

rent state. That is, for an epileptic patient, the behavior of brain activity can be increasingly predictable with the approach of a seizure [8, 37].

Table 3.5: Univariate Non-Linear Features adopted for some of the studies in recent years.

Study	Correlation dimension and correlation sum	Lyapunov exponent	Dynamical similarity index	Entropy
Mporas et al.[57] (2015)				x
Moghim et al.[37] (2014)	x	x		
Rabbi et al.[59] (2013)			x	
Acharya et al.[51] (2012)				x
Mirowksi et al.[68] (2009)		x		
Mormann et al.[69] (2005)	x	x		
Le Van Quyen et al.[73] (1999)			x	

These features are complex, computationally heavier, more sensible to parameters and lack reproducibility [8, 68]. Some authors presented different conclusions regarding the predictive power of the Lyapunov exponents [69, 75] and the dynamic similarity index [6, 73].

3.4.3 Multivariate Linear features

Measures applied to multiple EEG channels can be very useful for studying the synchronization between different brain regions and understanding how their interaction modulates the activity of epileptic seizures [8]. In Table 3.6 is presented the multivariate measures, both linear and non-linear, used in studies in recent years.

Table 3.6: Multivariate features adopted for some of the studies in recent years.

Study	Linear	Non-linear		
	Max. Linear cross-correlation	Mean phase coherence	Dynamical entrainment	Nonlinear interdependence
Aggarwal et al.[50] (2017)		x		
Rabbi et al.[59] (2013)		x		x
Mirowksi et al.[68] (2009)	x		x	x
Mormann et al.[69] (2005)	x	x		x
Kreuz et al.[40] (2004)		x		
Mormann et al.[71] (2003)		x		
Mormann et al.[72] (2003)	x	x		

Maximum linear cross-correlation is the most prominent linear multivariate feature [68, 69, 72]. This measure quantifies the similarity between two-time series, measuring the degree of lag synchronization. In other words, it assesses the similarity of two signals shifted in time. Therefore, it allows the distinction between the

period prior to seizure onset and the seizure period itself, since there are drops in synchronization in the pre-ictal period [8, 15, 72].

Recently, Bandarabadi et al.[54] introduced a new spectral power approach extended for two channels. The bivariate spectral band power provides the cross-power information between two different sub-bands across all possible channel pairs. This bivariate method was utilized for tracking changes before seizures and in 90% of the cases under study was selected as the best predicting feature.

3.4.4 Multivariate Non-linear features

Non-linear concepts can be also applied to multiple channels. These features (see Table 3.6) are based on mutual information and similarity between channels from different regions of the brain [8].

The most used features are mean phase coherence [40, 50, 59, 69, 71, 72], dynamical entrainment [68] and non-linear interdependence [59, 68, 69]. The first one evaluates the phase synchronization of two time series, the second one is a multi-channel version of the Lyapunov exponent and the last one is measure of generalized synchronization [71, 76, 77].

3.5 Feature selection

Feature selection allows, from a high dimensional feature set, to identify the most relevant subset of features to characterize the transition from the inter-ictal to the pre-ictal time. For that, irrelevant or redundant features must be excluded, as they can degrade the discriminative power of the features set and can lead to overfitting. This reduction should occur without significant loss of information. It is a very challenging problem and plays an important role in classification [45, 54].

Table 3.7 shows which studies use feature selection methods. It is worth mentioning that some researchers did not extract a huge set of resources, so feature selection methods were not carried out in these studies.

Principal Component Analysis (PCA) was also added, since it is used as a dimension reductional method. This method projects the data onto a orthogonal space and select the projections with higher variance values [11, 53, 68].

Some studies have employed methods such as maximum Difference Amplitude Distribution of histogram (mDAD) [54, 55, 61], minimum Redundance Maximum Relevance (mRMR) [54, 56, 61, 65] or RelieF [37, 57]. mDAD, as the name implies, is based on amplitude distributions histograms. The features that contributed with less

Table 3.7: Feature selection methods and PCA adopted for some of the studies in recent years.

Study	Feature selection methods				PCA
	mDAD	mRMR	GA	ReliefF	
Bandarabadi et al.[54] (2015)	x	x			
Bandarabadi et al.[55] (2015)	x				
Bou Assi et al.[52] (2015)			x		
Rasekhi et al.[56] (2015)		x			
Mporas et al.[57] (2015)				x	
Moghim et al.[37] (2014)				x	
Bandarabadi et al.[61] (2012)	x	x			
Direito et al.[65] (2011)		x	x		
Mirowksi et al.[68] (2009)					x
D'Alessandro et al.[70] (2003)			x		
Geva et al.[53] (1998)					x

superposition of the histograms for each class are considered the most discriminative [54]. mRMR is a well-known method that ranks features by two parameters. This is accomplished by minimizing the redundancy among the features while maximizing their relevance [11]. ReliefF ranks the features in order of importance by repeatedly sampling a random data instance and finding the value of a feature for the closest instance of the same and different classes. The output of the method is a vector of weights that allows ranking the features according to their discriminative power between classes [37, 57].

Bou Assi et al. [52], Direito et al.[65] and D'Alessandro et al.[70] had reported the use of genetic algorithms. Genetic Algorithm (GA) is an evolutionary algorithm inspired by the process of natural selection. It relies on some biological principles: from an initial population, the strongest ones will survive and recombine in order to adapt to external changes [52].

3.6 Classification

After extracting and selecting features, it is assumed that the remaining feature (or feature set) is suitable to distinguish between inter-ictal and pre-ictal periods. Therefore, a classification algorithm will then use this information to decide which class it belongs to. Firstly, it must be trained on how to make decisions and only after that phase must be employed to unseen data [3, 8].

There is a significant heterogeneity of algorithms that can be used, from simpler to more complex. Table 3.8 presents the classification decisions taken in the studies in recent years. Earlier studies adopted mainly classical Machine Learning

approaches while later studies started to introduce Deep Learning. Studies carried out over the past few years revealed that Support Vector Machines (SVM) is one of the most used classifiers for forecasting seizures, as it outperform others classification methods regarding the number of false positives produced [58]. Among the most recent studies addressing Deep Learning, CNNs [10, 19, 46, 47] are the most commonly used classification models.

Support Vector Machines

One of the prominent classifiers in seizure prediction is Support Vector Machines (SVMs). SVM is a supervised model that finds a linear separating hyperplane, in an N-dimensional space, which try to maximize the distance between the nearest points of different classes (see Fig. 3.4 a)) [3, 8].

Despite being a native linear classifier, SVMs can perform a non-linear classification using kernel functions, such as Gaussian Radial Basis Function (RBF). The original input space can be mapped into a higher dimensional space allowing the generation of a hyperplane with a largest margin separation between the two classes (see Fig. 3.4 b)) [3, 8, 11].

SVM is a binary classification algorithm although it is possible to extend it to multi-class problems. For this, the multi-class problem have to be reduced to several two-class problems. There are several strategies and the most used are "one vs rest" and "one vs one" [8, 11, 58].

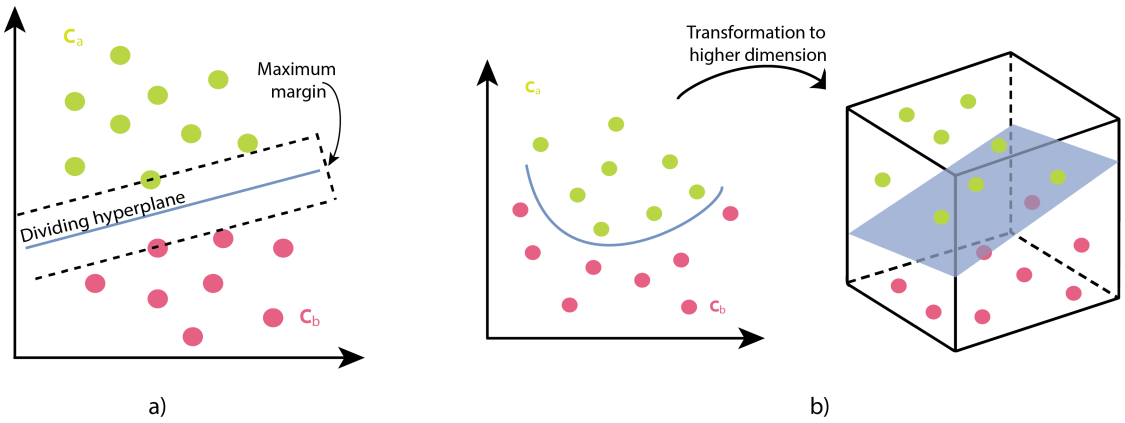

Figure 3.4: Classification of data by support vector machines. (a) is a linearly separable problem and (b) is an non-linearly separable case. Adapted from [3].

Table 3.8: Classification aspects adopted for some of the studies in recent years.

Study	Classifier	Patient specific	Train and Test
Agarwal et al.[48] (2018)	SVM	N.A.	N.A.
Chamseddine et al.[19] (2018)	LSTM GRU CNN	Yes	Train: 80% data samples Test: the remaining
Sun et al.[47] (2018)	LDA LR CNN RNN	Yes	N.A.
Khan et al.[46] (2018)	CNN	Yes	10-fold cross validation
Kiral-Kornek et al.[10] (2017)	CNN	Yes	Train: first 2 months Test: the remaining
Aggarwal et al.[50] (2017)	Thresholding	Yes	N.A.
Direito et al.[23] (2017)	SVM	Yes	Train: 2-3 seizures Test: the remaining
Bandarabadi et al.[54] (2015)	SVM	Yes	Train: first 3 seizures Test: the remaining
Bandarabadi et al.[55] (2015)	N.A.	Yes	N.A.
Bou Assi et al.[52] (2015)	SVM ANFIS	Yes	Train: 80% seizures Test: the remaining
Rasekhi et al.[56] (2015)	SVM	Yes	Train: first 3 seizures Test: the remaining
Mporas et al.[57] (2015)	SVM	Yes	10-fold cross validation
Teixeira et al.[58] (2014)	SVM ANN	Yes	Train: 2-3 seizures Test: the remaining
Moghim et al.[37] (2014)	SVM	Yes	10-fold cross validation Train: 70% data samples (random) Test: the remaining
Alvarado Rojas et al.[41] (2014)	Thresholding	Yes	Train: 2-4 seizures Test: the remaining
Rabbi et al.[59] (2013)	ANFIS	Yes	Train: 1 seizure Check: 1 seizure Test: the remaining
Rasekhi et al.[60] (2013)	SVM	Yes	Train: 3 first seizures Test: the remaining
Bandarabadi et al.[61] (2012)	SVM	Yes	Train: 3 seizures Test: the remaining
Teixeira et al.[62] (2012)	ANN SVM	Yes	Train: first 3 seizures Test: the remaining
Direito et al.[63] (2012)	HMM	Yes	N.A.
Valderrama et al.[64] (2012)	SVM	Yes	Train: fist half of seizures Test: the remaining
Acharya et al.[51] (2012)	FSC SVM KNN PNN DT GMM NBC	No	Data divided into 3 sets with classes in the same proportion Train: 2 sets Test: 1 set
Direito et al.[65] (2011)	SVM	Yes	Train: 3 seizures Test: the remaining
Park et al.[66] (2011)	SVM	Yes	Train: the remaining Test: 1 pre-ictal + 1 inter-ictal segment (random)
Chisci et al.[67] (2010)	SVM	Yes	Train: 15min preceding fisrt 2-3 seizures +30min interictal Test: the remaining
Mirowksi et al.[68] (2009)	SVM CNN Logistic regression	Yes	Train: the remaining Test: last 1-2 seizures + 33% of inter-ictal
Mormann et al.[69] (2005)	N.A.	N.A.	N.A.
Kreuz et al.[40] (2004)	Thresholding	Yes	N.A.
D'Alessandro et al.[70] (2003)	PNN	Yes	Train: 70% seizures Test: the remaining
Mormann et al.[71] (2003)	Thresholding	No	18-fold cross validation "leaving-one-out" for test
Mormann et al.[72] (2003)	Thresholding	No	10-fold cross validation "leaving-one-out" for test
Le Van Quyen et al.[73] (1999)	Thresholding	Yes	N.A.
Geva et al.[53] (1998)	UFC	Yes	N.A.

Hidden Markov Model (HMM), Fuzzy Sugeno Classifier (FSC), Support Vector Machine (SVM), K-Nearest Neighbour (KNN), Probabilistic Neural Network (PNN), Decision Tree (DT), Gaussian Mixture Model (GMM), Naive Bayes Classifier (NBC), Unsupervised Fuzzy Clustering (UFC), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Network (CNN), Linear Regression (LR), Linear Discriminant Analysis (LDA), Recurrent Neural Network (RNN)

Training, testing and validation

In addition, it is necessary to divide the data into subsets, one for training the model, other for validate (choose the model parameters) and the third one for testing the algorithm.

Concerning data splitting, once again there is great heterogeneity between studies. Data used for training should not be included in the validation and testing dataset. Nor should samples be selected at random from the entire data set for training, validation and testing. Test data should not be linked to the same events (seizures) as those used in training.

Authors make different assumptions while dividing the data. Some consider the seizure generation as a patient-independent process, so they gather all the data from all patients and only then divide it, as the case of Acharya et al.[51] and Mormann et al.[71, 72] studies. Therefore, a certain number of the total seizures from all patients is used for training, the remaining for testing.

Others researchers assume that seizure generation process differs from patient to patient. In this situation, a model is trained and tested for each patient [14]. In recent years, several authors opted for this assumption, such as Direito et al.[23, 65], Bou Assi et al.[52] and Moghim et al.[37]. Currently, patient-specific algorithms are the most widely used.

Some studies go further and in addition to assuming a specific algorithm for each patient, they also take into account the concept drift of seizures. Thus, they considered the order of seizures to split the data [10, 54, 56, 60, 62]. For example, in the study of Kiral-Kornek et al. [10] the first two months of data was used in the training phase and the remaining for testing and tuning the parameters.

Training data should be balanced, as the number of pre-ictal samples is considerably lower than non-pre-ictal samples. One way around this problem is to reduce the number of inter-ictal samples randomly (downsampling) [52, 54, 58, 60, 64], other is to adjust the weights of each class, such as the case of cost-sensitive SVMs [66].

3.7 Post-Processing

Post-processing is a step performed after classification that can employ regularization methods (e.g. Firing Power) or Deep Learning methods such as Long Short-Term Memory (LSTM) that deals with signal temporality. It is also possible to choose other approaches, such as the signals fusion (e.g. EEG and ECG) through

the combination of several classifiers outputs and the establishment of a set of rules to define the final output [78]. In Table 3.9 is possible to see which post-processing decisions were taken in the studies in recent years.

In seizure prediction studies, regularization methods are the most adopted. Regularization reduces the number of false alarms by smoothing the classifiers output (e.g. by reducing the noise). The most frequent regularization methods are Firing power and Kalman filter, which take into account the temporal signal dynamics. However, these techniques are ad-hoc approaches, which do not take into account any biological link over time, lacking this way a more clinical view of the seizure generation process [8].

Concerning Deep Learning approachs, Chamseddine et al. [19] compared the two more traditional regularization methods with several deep architectures: Gated Recurrent Unit (GRU), LSTM, CNN, and Multi-Layer Perceptron (MLP). The main conclusion of the study was that LSTMs can be successfully applied as post-processing regularizer for the classifier output. However, the study was conducted on only one individual, so it must be evaluated on larger scale data. In addition, the decisions of these architectures are difficult to explain, compromising their interpretability and clinical implementation [18, 79].

Kiral-Kornek et al. [10] implemented an artificial integrate-and-fire neuron to ensure that only in the presence of several pre-ictal class predictions in close temporal proximity to a seizure, an alarm would be triggered. This approach is based on the same idea as the Firing Power filter.

Firing power

Firing power was proposed by Teixeira et al.[11] to quantify the number of samples classified as pre-ictal, using a sliding window technique.

Mathematically, the Firing power is defined by (3.1)

$$fp\,[n] = \frac{\sum_{k=n-\tau}^{n} O\,[k]}{\tau} \tag{3.1}$$

where n is a given discrete time instant, τ is the number of samples in each window, with size equal to the pre-ictal period and $O\,[k]$ is the binary classifier output at sample k. $O\,[k]$ can only take two values, 1 for a samples classified as pre-ictal and 0 otherwise. $fp\,[n]$ ranges between 0 and 1. The alarm is raised when $fp\,[n]$ exceeds a certain threshold. After the generation of an alarm, is only possible to raise a new one after a time equal to the pre-ictal period and if $fp\,[n]$ exceeds the threshold in an ascending way. No optimal threshold has yet been determined

[8, 11, 62]. Fig 3.5 illustrates how Firing power operates.

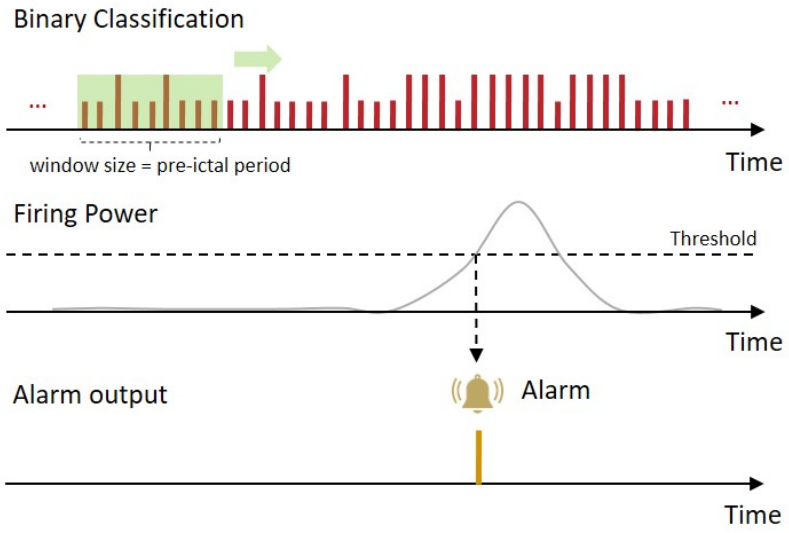

Figure 3.5: Practical definition of Firing Power Filter. Adapted from [11].

In short, firing power can be seen simply as a moving average filter, a low pass to eliminate noise. Thus an alarm is raised, when over time there is a great consensus of voting of the classifier in the pre-ictal class.

Kalman filter

Chisci et al.[67] were the first to use the Kalman filter both in the field of epilepsy and for classification purposes. Kalman filter is a regularization strategy operating on a continuous variable, inside the classification procedure. This filter reduces the signal chattering behaviour by imposing smooth transitions [66, 67].

For this, it is needed to estimate the states vector $s_k = [d_k, \dot{d}_k]$, where d_k denote the noiseless decision variable and $\dot{d}_k$ its rate of change. Therefore, the continuous SVM output z_k, can be represented by the state-space model (3.2)

$$\begin{cases} s_{k+1} = \begin{bmatrix} 1 & T_p \\ 0 & 1 \end{bmatrix} s_k + w_k \\ z_k = \begin{bmatrix} 1 & 0 \end{bmatrix} s_k + v_k \end{cases} \tag{3.2}$$

where T_p is the prediction interval, w_k and v_k are zero mean white noise vectors[67].

The Kalman filter can be applied to recursively provide a filtered estimate of d_k which will then be used instead of z_k in the classification process. The parameter

that the filter can adjust is the standard deviation for the random fluctuations of $\dot{d}_k$ which is equivalent to the Kalman gain $\sigma_{KF} \triangleq \sigma_w$ [62, 67].

The alarm is raised when the Kalman filter smoothed output is classified as a pre-ictal sample. A new alarm can only be generated when the output crosses the threshold in an ascending way [62].

In conclusion, the Kalman filter is a less conservative method as it does not maintain a long memory of classification dynamics and has no time restrictions for generating alarms, and thus presenting a higher number of alarms [8, 62].

Generally, both regularization techniques are related to the decisions of a "static" classifier over time, lacking some clinical knowledge related to the problem.

3.8 Performance Evaluation

As already mentioned in the previous chapter, the most appropriate metrics to evaluate performance in seizure prediction are sensitivity and False Positive Rate per Hour (FPR/h) [6, 8]. However, some authors instead of using FPR/h opted for specificity [37, 51, 52, 63–65]. Others have chosen other measures to evaluate performance, such as accuracy [23, 37, 51, 57, 63, 65].

As can be seen in Table 3.9, few studies achieved an FPR/h less than 0.15, the reference value for warning systems established by Winterhalder et al. for pre-surgical patients [54, 60, 61]. Nonetheless, it is complicated to make a comparison between studies, since both the number of patients and the number of records varies significantly, as well as the values adopted for SOP and SPH. Nevertheless, even using the same database, there is no convergence of results. Moreover, some authors only report the best performance values despite having trained and tested different models for the same patient. Additionally, different studies have different application goals making it difficult to compare the consequences for patients.

Given the complexity of seizures prediction, in addition to the calculation of metrics for performance evaluation (sensitivity and FPR/h), statistical validation should also be implemented to understand if the proposed methods perform above chance level [15]. Although in recent years it started to be something more frequent, most older prediction studies did not perform it.

Table 3.9: Performance and post-processing decisions from some of the studies in recent years.

Study	Post-processing	SS(%)	FPR/h	Statistical validation
Agarwal et al.[48] (2018)	N.A.	96.75	N.A.	N.A.
Chamseddine et al.[19] (2018)	LSTM, GRU, CNN, MLP, Firing Power, Kalman filter	88.70	N.A.	N.A.
Sun et al.[47] (2018)	N.A.	N.A.	N.A.	N.A.
Khan et al.[46] (2018)	N.A.	87.80	0.14	Yes (Random predictor)
Kiral-Kornek et al.[10] (2017)	Integrate-and-fire neuron	68.60	N.A.	Yes (Random predictor)
Aggarwal et al.[50] (2017)	N.A.	N.A.	N.A.	N.A.
Direito et al.[23] (2017)	Firing Power	35.34	0.20	Yes (Random predictor)
Bandarabadi et al.[54] (2015)	Firing Power	75.80	0.10	Yes (Random predictor)
Bandarabadi et al.[55] (2015)	N.A.	N.A.	N.A.	N.A.
Bou Assi et al.[52] (2015)	N.A.	82.02*	N.A.	N.A.
Rasekhi et al.[56] (2015)	Firing Power	60.90	0.11	Yes (Random predictor)
Mporas et al.[57] (2015)	N.A.	N.A.	N.A.	N.A.
Teixeira et al.[58] (2014)	Firing Power	70.61*	0.34*	N.A.
Moghim et al.[37] (2014)	N.A.	91.14	N.A.	N.A.
Alvarado Rojas et al.[41] (2014)	Kalman filter	68	0.33	Yes (Random predictor)
Rabbi et al.[59] (2013)	N.A.	80^{Θ}	0.46^{Θ}	N.A.
Rasekhi et al.[60] (2013)	Firing Power	73.90	0.15	N.A.
Bandarabadi et al.[61] (2012)	Firing Power	67.39*	0.12*	N.A.
Teixeira et al.[62] (2012)	Firing Power and Kalman filter	77	0.20	N.A.
Direito et al.[63] (2012)	N.A.	94.59	N.A.	N.A.
Valderrama et al.[64] (2012)	N.A.	33.38	N.A.	Yes (Surrogate predictor)
Acharya et al.[51] (2012)	N.A.	97.60*	N.A.	N.A.
Direito et al.[65] (2011)	N.A.	43.35*	N.A.	N.A.
Park et al.[66] (2011)	Kalman filter	94.40*	0.25*	N.A.
Chisci et al.[67] (2010)	Kalman filter	100	0.41*	N.A.
Mirowksi et al.[68] (2009)	N.A.	71^{Θ}	0^{Θ}	Yes (Random predictor)
Mormann et al.[69] (2005)	N.A.	N.A.	N.A.	N.A.
Kreuz et al.[40] (2004)	N.A.	N.A.	N.A.	Yes (Surrogate predictor)
D'Alessandro et al.[70] (2003)	N.A.	62.50	0.28	N.A.
Mormann et al.[71] (2003)	N.A.	N.A.	N.A.	N.A.
Mormann et al.[72] (2003)	N.A.	N.A.	N.A.	N.A.
Le Van Quyen et al.[73] (1999)	N.A.	N.A.	N.A.	N.A.
Geva et al.[53] (1998)	N.A.	N.A.	N.A.	N.A.

Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Convolutional Neural Networks (CNN)
Multi-Layer Perceptron (MLP), Average performance (*), Best performance ($^{\Theta}$)

3.9 Summary

Most seizure prediction studies follow a similar pipeline constituted by several main steps. Data acquisition plays a critical role in the seizure prediction field as it provides the "raw material". The vast majority of existing databases contain records of patients undergoing pre-surgical monitoring. This may be an obstacle to real-life application devices. In the future, data of several months or years should be acquired in a normal day-to-day scenario.

After the acquisition of the EEG signal, many authors choose to pre-process it by filtering, to remove power line interference, frequencies related to noise, and other physiological phenomena. Then, the most heterogeneous step is performed: a sliding window technique is usually used for feature extraction, where the univariate linear are the most chosen ones. Feature selection can be an optional step and aims to increase the discriminative power of the feature set.

Subsequently, a classification algorithm is used to detect pre-ictal changes. Several strategies are adopted at this stage, including Deep Learning models, where SVM is the most widely used in studies. To reduce the number of false alarms, the output of the classifier must go through a post-processing phase. The most common technique is the signal regularization through filters (Firing Power Filter and Kalman Filter). Some authors have also used Deep Learning methods, such as LSTMs that handle signal temporality. However, these strategies are difficult to interpret, compromising the clinical application. Therefore, efforts must be made to develop post-processing methods capable of dealing with the temporality of brain activity prioritizing clinical interpretability.

Finally, performance must be assessed using sensitivity, FPR/h and statistical validation. The lack of a correct evaluation through the appropriate metrics and statistical validation is frequent in several studies. Another problem is the presentation of performance results only for the optimal SOP value and not for a range of values, biasing the results positively and making comparison between studies difficult.

4

Methodology

The present work aims to develop a patient-tailored algorithm for epileptic seizure prediction while exploring the chronology of brain activity as a post-processing stage. For this, two different post-processing methodologies were adopted along with a control one, inspired in the most common framework presented in the state of the art. The details of each one will be presented in section 4.7 of this chapter.

The general framework of the adopted seizure prediction methodologies is summarized in Fig. 4.1. As outlined, it consists of several stages: pre-processing, feature extraction, data splitting, feature selection, classification, post-processing and performance evaluation. The proposed pipeline will be applied for three different Seizure Occurrence Period (SOP) values: 10, 20 and 30 minutes.

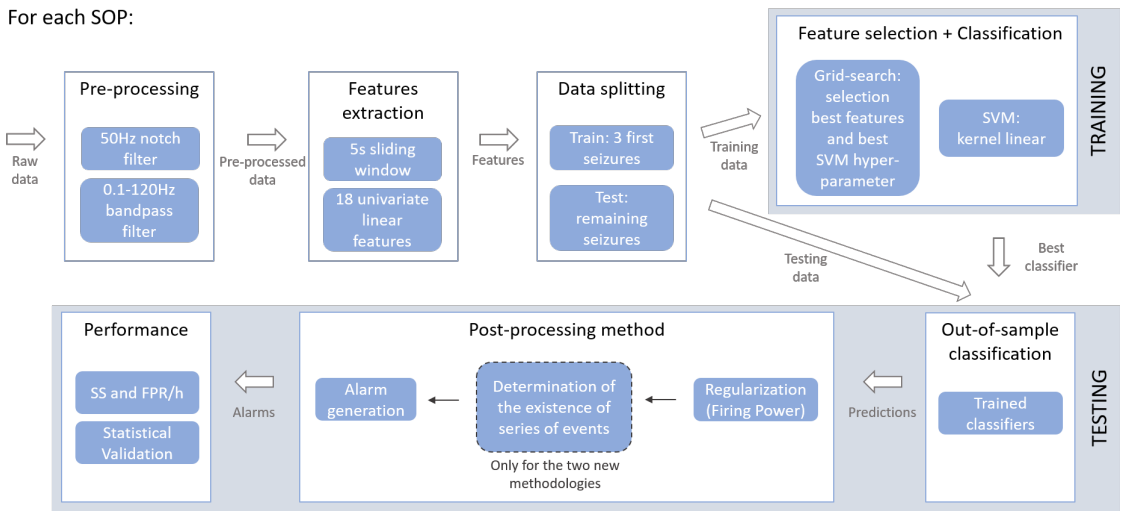

Figure 4.1: General outline of the proposed pipeline for each SOP.

The raw Electroencephalogram (EEG) data is filtered and then segmented to extract features. The extracted features are divided into two groups: the training and testing sets. Training features are used to fit the parameters of the classifiers and feature selection. Once the models are completely trained, they are applied to the testing dataset to make predictions based on the observations. Successively, the

post-processing phase is carried out. Firstly, a regularization filter (Firing Power) is applied to the predictions. This filter was chosen since it has a simple and intuitive understanding and has already been used in previous studies. Subsequently, in the proposed methodologies an extra step is used to find out if there is a chronological series of events that will trigger an alarm. Finally, it is evaluated the performance of the model. Since the algorithm is patient-specific, all these steps were performed for each patient.

4.1 Data

For this study, 36 Drug-Resistant Epilepsy (DRE) patients (15 females and 21 males, aged 40.08 ± 15.62 years) were selected from the EPILEPSIAE database. These patients' data were acquired by the University Medical Center of Freiburg in Germany. All the selected patients contain seizures localized in the temporal lobe and have EEG scalp records acquired with a sampling rate of 256Hz. For all patients, 19 electrodes belonging to the 10-20 system were analyzed: FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz and Pz. Table 4.1 presents the demographic data of each patient (gender and age) and their seizures information (number of seizures, their classification, state of vigilance during the seizure and the period of the day it occurred).

For the selection of patients, the number of registered seizures was investigated and those with at least 5 independent seizures were selected. For a seizure to be considered independent, it has to be separated from another at least 4 hours. When two seizures occur in an interval shorter than 4 hours, they are considered to belong to the same cluster [14, 15]. So although the total dataset contains 345 seizures, only 214 were considered as independent and therefore suitable for analysis.

Finally, the hours before the selected seizures were selected. For the first seizure of each patient, data was collected up to one day before. This is because from the beginning of the capture of EEG records until the first seizures, there may be a period of adaptation of the brain, since the patients are under pre-surgical monitoring and the medication may still have some effect. The value was set based on articles of the state of the art which indicate that for evaluation purposes, 24 hours of inter-ictal data per seizure are necessary [6, 16]. For some patients, adjustments were made to the time collected before certain seizures, as there were missing files. In these cases, either the time has been reduced or the seizure has been ruled out. This was the case for patients 56402, 60002, 93902, 16202 and 98202.

As the data belong to patients undergoing pre-surgical monitoring, the present

work can only be seen as a proof-of-concept that, if successful, should be tested on more reliable real-life data.

4.2 Pre-processing

First, the data was filtered with a 50Hz notch filter to remove the interference from the power supply followed by a bandpass filter between 0.1-120Hz. Next, a highpass fourth-order butterworth filter was applied with a cutoff frequency equal to 0.1Hz and then a lowpass filter with a cutoff frequency of 120Hz.

In this phase, the segments in which there were gaps in the EEG data were also identified, since the data are kept in records of approximately 1 hour and may not be fully continuous. When collecting features, these gaps were discarded, as they are not considered to significantly affect performance [15].

4.3 Feature extraction

After pre-processing, the next step is to extract features from the EEG signal. A 5s sliding window without overlap was used to extract the 18 features. This window was considered adequate to represent EEG variations and was chosen based on the state of the art of seizure prediction in Epilepsy. In general, 5s is a reasonable window in terms of trade-off between stationarity, temporal and spectral resolution [8, 23].

The univariate linear features were computed on 19 electrodes of the EEG segments and are listed below:

- Statistics (mean, variance, skewness and kurtosis);
- Relative spectral power (delta ($< 4Hz$), theta ($4 - 7Hz$), alpha ($8 - 12Hz$), beta ($12 - 30Hz$) and gamma ($> 30Hz$));
- Hjorth (mobility and complexity);
- Energy of wavelet coefficients (6 levels of decomposition);
- Accumulated energy.

It was considered that the EEG patterns that lead to the seizure event can be recognized with linear and univariate features. Linear features, as mentioned in section 3.4, are better understood and perform similarly to non-linear features [69]. These were computed based only on single EEG channels to make the process faster.

As mentioned before, the notion of network theory was used, so it was considered that all brain regions may have features that lead to the development of a

Table 4.1: Information for the 36 studied patients.

Patient	Gender	Age	# seizures (training/ testing)	Record time (h)	Seizure classification	Seizure activity pattern	Vigilance state	Period of the day
402	f	55	3	74.9	FOIA, FBTC, FOIA	t, t, t	A, A, A	N, D, N
			2	25.0	FBTC, FOIA	t, t	A, A	D, D
8902	f	67	3	53.3	UC, FOIA, FOIA	a, b, a	A, A, A	N, N, N
			2	23.2	FOIA, FOIA	m, a	A, A	N, N
11002	m	41	3	38.9	UC, FOIA, FOIA	?, s, a	A, R, A	N, N, D
			2	16.2	FOIA, FOA	t, t	A, A	D, D
16202	f	46	3	61.4	UC, FBTC, UC	r, ?, r	A, A, A	N, N, N
			4	34.5	FOIA, FOIA, FOIA, FOIA	r, r, ?, r	A, A, A, A	D, N, D, D
23902	m	36	3	71.4	FOA, FOA, FOA	t, t, t	A, A, A	D, D, D
			2	34.6	FOA, FOA	d, t	A, A	D, N
30802	m	28	3	51.1	FOA, FOA, FOA	t, t, t	R, A, 2	N, N, D
			5	63.4	FOA, FOA, FOA, FOA, FOA	t, t, t, t, t	A, A, R, 2, 2	N, N, N, N, N
32702	f	62	3	73.4	FOIA, FOIA, FOIA	t, t, t	A, A, A	D, D, D
			2	20.5	FOIA, FOIA	r, a	A, A	D, D
46702	f	15	3	47.9	FOA, FOIA, FOIA	a, a, t	A, 2, A	D, N, D
			2	13.0	FBTC, FOIA	b, t	2, A	N, N
50802	m	43	3	129.3	FOIA, UC, UC	t, t, t	A, 2, 2	N, N, D
			2	36.3	FOIA, FBTC	t, t	2, A	N, N
53402	m	39	3	38.3	FOA, FOA, FOA	?, ?, ?	A, A, 2	D, D, N
			2	51.1	FOA, FOIA	?, t	A, A	D, D
55202	f	17	3	47.8	FOIA, FOIA, FOA	t, d, t	A, A, A	N, D, D
			5	67.0	UC, UC, FOA, UC, FOIA	t, t, t, r, r	A, A, A, A, A	D, D, D, D, D
56402	m	47	3	61.3	UC, UC, UC	t, ?, b	A, A, A	D, D, N
			3	71.8	UC, FBTC, FBTC	?, a, t	A, A, A	D, N, D
58602	m	32	3	79.9	FOIA, FOIA, FOIA	r, t, t	A, R, A	D, N, D
			4	28.6	FOIA, FOIA, FOIA, FOIA	r, r, r, t	A, A, A, 2	D, D, D, N
59102	m	47	3	65.2	FOA, FOIA, FOIA	?, t, t	A, A, A	D, D, D
			2	82.5	FOIA, FOA	t, t	A, A	D, D
60002	m	55	3	35.7	FOIA, FOIA, FOIA	d, c, t	1, A, A	N, N, D
			3	50.0	UC, FOIA, FOIA	t, d, d	R, R, 1	N, N, N
64702	m	51	3	75.6	FOA, FBTC, FBTC	?, m, t	A, A, A	D, N, D
			2	32.3	FBTC, FBTC	t, t	A, 2	D, N
75202	m	13	3	69.4	FOA, FOA, UC	t, t, t	2, 2, A	N, N, D
			4	53.9	FOA, FOA, FOA, FOA	t, t, ?, t	A, A, A, A	D, D, D, N
80702	f	22	3	50.1	FOIA, FOIA, UC	b, b, ?	A, A, A	N, D, D
			3	30.4	FOIA, FBTC, FOIA	c, c, c	A, A, A	N, D, D
85202	f	54	3	45.6	FOIA, FOIA, UC	m, c, m	2, A, A	N, D, N
			2	21.0	UC, UC	m, m	A, A	D, N
93402	m	67	3	45.4	FBTC, FOIA, FOIA	t, t, t	2, 2, 2	N, D, N
			2	54.5	UC, UC	t, t	2, 2	N, N
93902	m	50	3	129.2	FOA, FOIA, FBTC	t, t, d	A, A, 2	D, D, N
			2	14.1	FOIA, UC	d, d	2, A	N, D
94402	f	37	3	95.8	FOA, UC, FOIA	?, d, b	A, A, A	D, D, D
			4	32.0	UC, FOA, UC, FOA	t, ?, b, ?	2, A, 2, A	N, D, N, D
95202	f	50	3	44.3	FBTC, FOIA, FOIA	b, b, b	2, 2, 2	N, D, N
			4	91.6	FOIA, UC, FOIA, UC	m, b, b, t	2, 2, 2, 2	D, N, N, N
96002	m	58	3	46.3	FOIA, FOIA, FOIA	t, t, t	A, A, A	D, D, D
			4	74.8	FOIA, UC, FOIA, FOIA	d, a, t, a	A, A, A, A	N, N, D ,N
98102	m	36	3	108.0	FOA, UC, UC	?, ?, ?	A, A, A	N, D, N
			2	46.2	UC, FBTC	?, ?	A, A	N, N
98202	m	39	3	74.0	FOIA, FOIA, FOIA	t, a, t	A, A, A	N, D, N
			4	37.3	FBTC, FOIA, FOIA, UC	t, t, t, t	A, A, A, A	D, N, N, D
101702	m	52	3	29.0	FOIA, FOIA, FOIA	t, t, t	A, A, A	N, D, D
			2	24.4	FOIA, FOIA	r, r	2, A	N, D
102202	m	17	3	50.5	FOA, UC, FOIA	b, ?, t	2, A, 2	N, D, N
			4	52.7	UC, FOA, FOIA, UC	?, t, t, t	A, A, 2, A	N, D, N, D
104602	f	17	3	38.2	FOIA, FBTC, FBTC	t, a, t	A, 2, 2	D, N, N
			2	16	FBTC, UC	t, d	2, 2	D, N
109502	m	50	3	34.6	FOIA, FOIA, FOIA	t, t, t	A, A, 1	D, D, N
			2	47.7	UC, UC	t, t	A, A	N, D
110602	m	56	3	69.3	FOIA, FOIA, FOIA	t, t, t	A, A, A	D, D, D
			2	26.5	FOIA, FOA	t, t	A, A	D, D
112802	m	52	3	45.9	UC, FOIA, UC	t, t, t	A, A, A	D, N, D
			3	110.4	FOIA,FOIA,UC	t, t, t	A, A, A	N, D, D
113902	f	29	3	52.9	UC, FOIA, FOIA	t, d, t	A, A, 2	N, D, N
			4	27.8	FOIA, UC, UC, FOIA	t, t, t, t	A, 2, A, A	D, N, D, D
114702	f	22	3	45.0	FOIA, FOIA, UC	t, t, t	A, A, A	D, D, N
			6	39.8	FOIA, FOIA, FOIA, FOIA, FOIA, FOIA	t, d, t, t, d, t	A, A, A, 2, A, A	D, D, D, N, D, D
114902	f	16	3	27.6	FOA, FOIA, FOIA	s, b, s	A, A, A	D, D, D
			4	51.9	FBTC, UC, FOIA, FOIA	t, r, a, t	2, A, A, A	N, D, D, D
123902	f	25	3	94.6	FBTC, FBTC, FOIA	t, t, t	2, 2, R	N, N, N
			2	30.8	FOIA, FOA	t, t	A, A	D, D

seizures: number of seizures; Gender: female (f), male (m); Seizure classification: unclassified (UC), Focal Onset Aware (FOA), Focal Onset Impaired Awareness (FOIA), Focal to Bilateral Tonic-Clonic (FBTC); Seizure activity pattern: unclear (?), rhythmic sharp waves (s), rhythmic alpha waves (a), rhythmic delta waves (d), rhythmic theta waves (t), rhythmic beta waves (b), repetitive spiking (r), cessation of inter-ictal activity (c), amplitude depression (m); Vigilance state: awake (A), REM sleep stage (R), Non-REM sleep stage I (1), Non-REM sleep stage II (2); Period of the day:day, between 8am and 22pm (D), night, between 22pm and 8am (N).

seizure, therefore, there was no previous selection of the electrodes to be used. Thus, in total for each patient, 18 features were collected in each of the 19 EEG channels, making a total of 342 features extracted [15].

4.4 Data splitting

The feature set was split into 2 sets for each patient. The first set, the training set, containing the first 3 seizures was used for classifier training and parameter optimization. The second set was constituted by the remaining seizures and used as testing set.

The seizures were chronologically divided respecting the order in which they occurred to take into account the concept drifts. In addition, this division aimed to simulate as possible a real seizure prediction scenario. First, the predictor algorithm would use a determined number of seizures collected in an initial period and only then applied online to upcoming data.

Overall, 108 seizures were used in the training phase and 106 in testing, comprising a total of 2201.1 hours and 1533.8 hours respectively.

4.5 Feature Selection and Classification

Classes labeling

After feature extraction, the samples of feature sets are labeled into two classes: event or pre-ictal (depending on the approaches defined in the section 4.7) and inter-ictal. Three different event/pre-ictal durations (10, 20 and 30 minutes) were used for the correspondent 10, 20 and 30 minutes of SOP.

The labeling took into account a Seizure Prediction Horizon (SPH) of 5 min, a duration considered sufficient for a possible intervention by the patient as a form of prevention through a timely raised warning.

Class balancing

In the training stage, one must define the class balancing strategy. As the inter-ictal class contains the majority of samples, unbalanced data could lead to biased and inaccurate results. Thus, the inter-ictal class was undersampled, for a total number of samples equal to the number of samples of the pre-ictal samples. This technique is called random undersampling.

This process (see Fig. 4.2) was carried out at each seizure and maintained the sequential chronology of the inter-ictal samples. Therefore, the total set of inter-ictal samples was divided into n groups, with n being the number of pre-ictal class. Finally, a sample was chosen at random from each of these groups. This was performed in order to maintain points from all inter-ictal intervals and therefore, to guarantee a better representativeness.

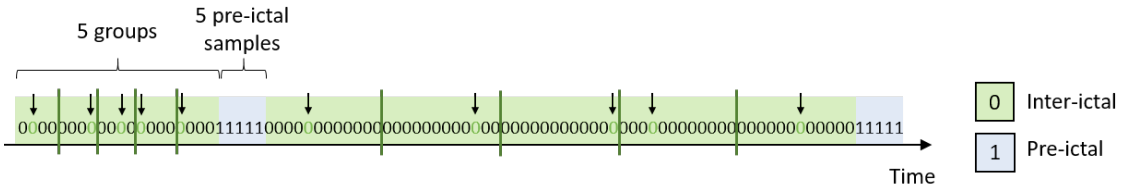

Figure 4.2: Random undersampling of inter-ictal class respecting the sequential chronology of samples. Green colored samples with a black arrow above correspond to the ones randomly chosen from each group.

Feature selection

For the selection of features, a filter-based method was used. Filter-based methods are simple and computationally fast and they rank features according to their relationship with the target based on a univariate metric. The metric used was the Pearson correlation coefficient that assumes the linearity of the data [80, 81].

The number of features was reasonably defined for 10, 20, or 30. The final value was chosen by the grid-search method.

Support Vector Machines (SVM)

For classification, the SVM was used since it has few parameters for optimization and has shown good results for seizure prediction in previous studies [23, 37, 51, 52, 54, 56–58, 60–62, 64–68]. The kernel chosen was the linear one as it is simpler and computationally lighter, presenting similar performance to others [23]. For the implementation of linear SVM it was used python scikit-learn library [82].

Hyper-parameter C (cost), a regularization parameter to balance the number of misclassifications and the size of the class separation margin was selected with a grid-search procedure.

Ensemble learning

In order to deal with the stochasticity inherent to downsampling, an ensemble learning was adopted, where 31 SVM classifiers were trained with different data samples due to class balancing. When tested, the final prediction will result from a vote between the predictions of all classifiers.

Training: Grid-search

To obtain the optimal parameters for training the classifiers, a grid-search approach was adopted in order to incorporate both the search for the best set of features (feature selection) and the search for the best value of the SVM hyperparameter. The number of features (k) was 10, 20 or 30, and eight values were used for the SVM cost (C) ($[2^{-20}, 2^{-16}, 2^{-12}, 2^{-8}, 2^{-4}, 2^{0}, 2^{4}, 2^{8}]$).

In order to optimize these parameters a similar approach to Direito et al. [23] was implemented. A 3-fold cross-validation was used with an evaluation metric based on sensitivity and specificity ($\sqrt{SS^2 + SP^2}$).

The division of the 3 groups (folds) was made respecting the data for each seizure as represented in the Fig. 4.3. The dataset contains data from the first three seizures (pre-ictal period and the corresponding inter-ictal period) and each seizure was associated with a group. Thus, folds 1, 2, and 3 are composed of data from the first (S1), second (S2), and third (S3) seizures, respectively, ensuring that each group contains samples from both pre-ictal and inter-ictal classes. For a given iteration i of the cross-validation procedure, the test set corresponds to fold i and the training set to the remaining folds.

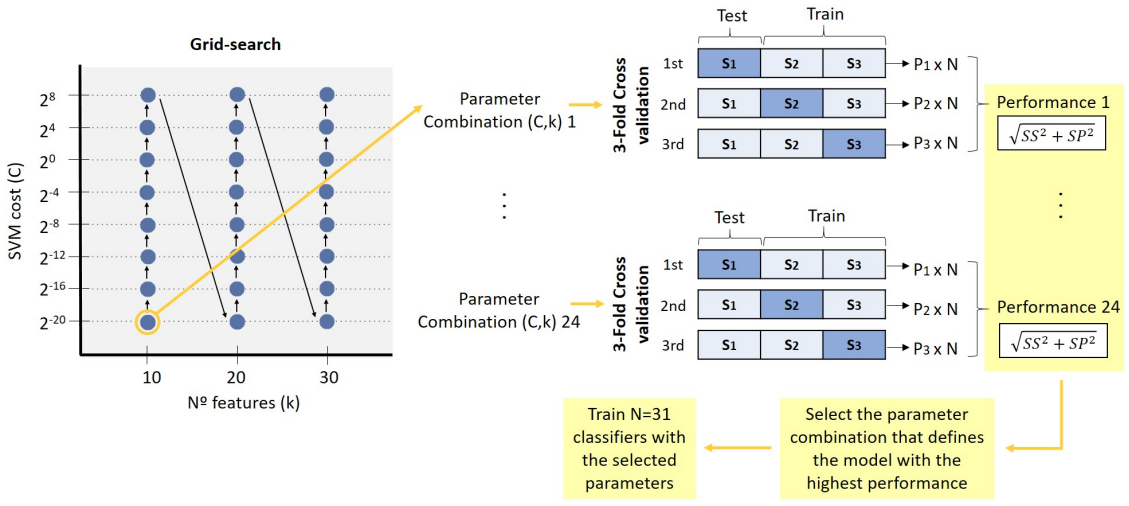

Figure 4.3: Selection process of the best training parameters.

Due to the ensemble learning, each of the three iterations was performed 31 times (N = 31) and the average of their performances was used to calculate the final metric. The final performances of the 24 parameter combinations were compared and the set of parameters with the highest performance was selected.

Finally, 31 classifiers were trained with the chosen parameters.

4.6 Out-of-sample classification

After training the classifiers, the testing data is used to make predictions. But first, the procedure applied to the training data is applied to this data, excluding the class balancing step (see Fig. 4.4). Thus, the test features are standardized using the z-score parameters (mean and standard deviation) of the training data. Then, the features identified in the training phase are selected. Finally, this set of features is used to provide an unbiased evaluation of the trained classifiers.

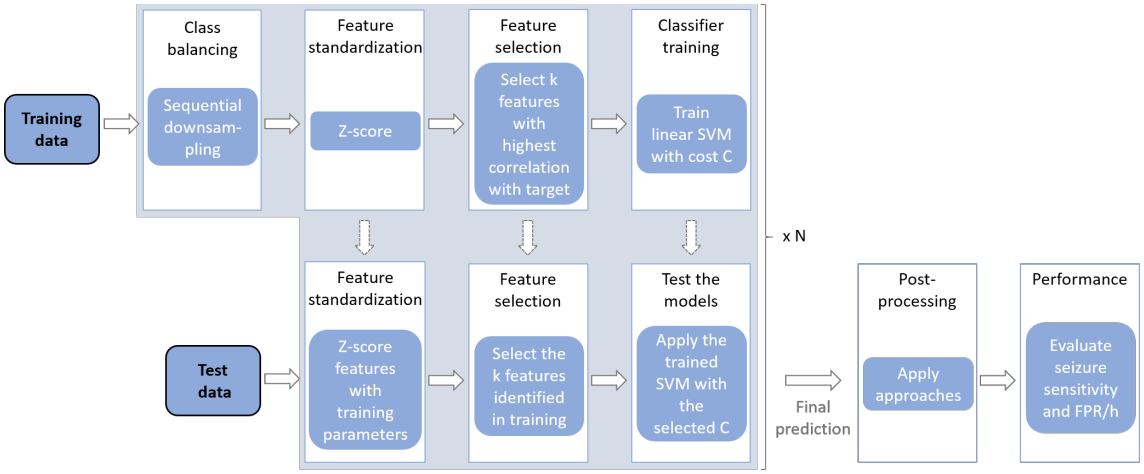

Figure 4.4: Procedure applied to training an test data.

This process is repeated for each of the 31 trained classifiers (N = 31), resulting in 31 predictions. Those are used in a voting system, in which for a given sample the predominant class in all predictions is chosen for the final output.

4.7 Post-processing methods

It may be unrealistic to simply consider the output of the classifiers as a good alarm generator, since not only it is unlikely that it will classify all samples correctly and may trigger false alarms, but also can be susceptible to noise contained in the data. In order to prevent this, regularization filters are used. In this work, the three different post-processing approaches were based on the same regularization method,

the Firing Power filter which was used in order to take into account the temporal dynamics of the classification and to reduce the noise.

Firing Power was obtained according to the procedure described in the section 3.7. An alarm is triggered when the Firing Power exceeds a certain predefined threshold (L) [62]. The alarm series $a[n]$ is represented in equation 4.1. The threshold values vary depending on the approach.

$$a[n] = \begin{cases} \text{alarm, if } fp\,[n] \geq L \\ \text{no alarm, if } fp\,[n] < L \end{cases} \tag{4.1}$$

As already mentioned, the present work intends to embrace the brain as a biological system. To consider the generation of seizures as a chronological process of events in brain activity approaches A and B were developed. Each one has a different post-processing which depends on the trained classifiers. The methods developed were based on a sequence of three events. Approach C assumes that the generation of seizures is a key event fixed in time, as in state of the art articles, thus serving for comparison. The three approaches are presented below.

It is worth mentioning that three different time durations were analyzed: 10, 20, and 30 minutes. Figures 4.5, 4.6, 4.7 were only shown for a 10-minute duration, just for exemplification. The figures do not show all the steps of the methodology set out above, as they intend to emphasize only what changes among approaches.

4.7.1 Approach A

In the first approach, three events separated in time were considered. The duration of each event is equal to the SOP value.

The aforementioned Machine Learning pipeline was applied to each event, resulting in three model outputs. Then, a Firing Power filter was used for each output. A threshold (L) of 0.5 was established. When each Firing Power exceeds this defined threshold, a given event is considered to have occurred.

An alarm is triggered by the patient warning system only when the three events occur chronologically within a stipulated period of time (three times the value of the SOP under analysis).

Figure 4.5 shows an example for a 10-min SOP pipeline. The first event is associated with the time interval of 35 to 25 minutes before the seizure onset, the second between 25-15 minutes and the third between 15-5 minutes, with the SPH being 5 minutes. After the various stages of the pipeline, three outputs are obtained to which a Firing Power filter is applied. The Firing Power associated with the first

event (FP_1) exceeds the threshold L three times, resulting in three cases. For each of those, a search for two other events will be carried out in the second and third Firing Power, within a period of 30 minutes. In all cases, the FP_2 also crossed the threshold. Finally, the same inspection is done for the FP_3 between the time when the second event occurred and the end of the established period (30 minutes after the first event). In the example shown, this only occurred in the first case, in which an alarm was activated.

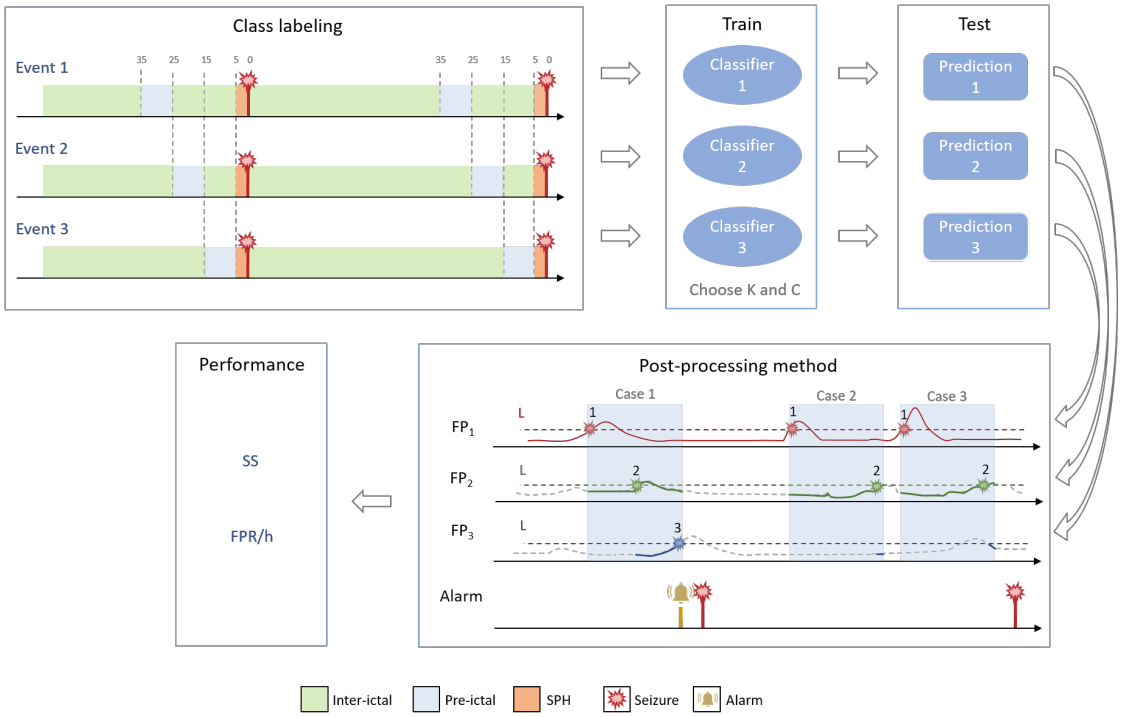

Figure 4.5: Approach A.

4.7.2 Approach B

Approach B deals with chronology in a different way. Its purpose is to inspect whether a cumulative Firing Power can correctly trigger alarms. For this purpose, three events of different duration and overlapping in time are considered. Thus, the third event is shorter, of equal duration to the SOP value, and closer to seizure onset. The other two have a duration two and three times longer than the first one.

After performing the proposed pipeline for each of the events, the Firing Power filter is applied. Hence three resulting Firing Power parameters are added together to form a single Firing Power. It is in the latter that the existence of a chronological sequence of events will be evaluated. For this, three thresholds of 0.5, 1.0 and 1.5 are used. Thus, for an alarm to be released, the total Firing Power must exceed the

three thresholds from the smallest to the largest in a predefined period (three times the value of the SOP adopted). That is, after the threshold of 0.5 is exceeded, it is necessary that the threshold of 1.0 and 1.5 also be overpassed in the given time.

Figure 4.6 presents an example of this situation for a 10-minute SOP. The first event corresponds to the period between 5 to 35 minutes before seizure, the second between 5 to 25 minutes and the third between 5 and 15 minutes. After performing the pipeline, it is evaluated when the total Firing Power exceeds the first threshold, the red dashed line. Once these points are found, it is searched for up to a maximum time of 30 minutes later, if that same Firing Power exceeds a second threshold (L_2). If it happens, it is still seen if the FP_{total} exceeds the third threshold (L_3) between the time the second threshold was passed and the final time. The three events only occurred, within the stipulated period, for the second case, so only one alarm was raised. In the first case the three event did not occur within the 30 minutes, so no alarm was raised to warn about that seizure.

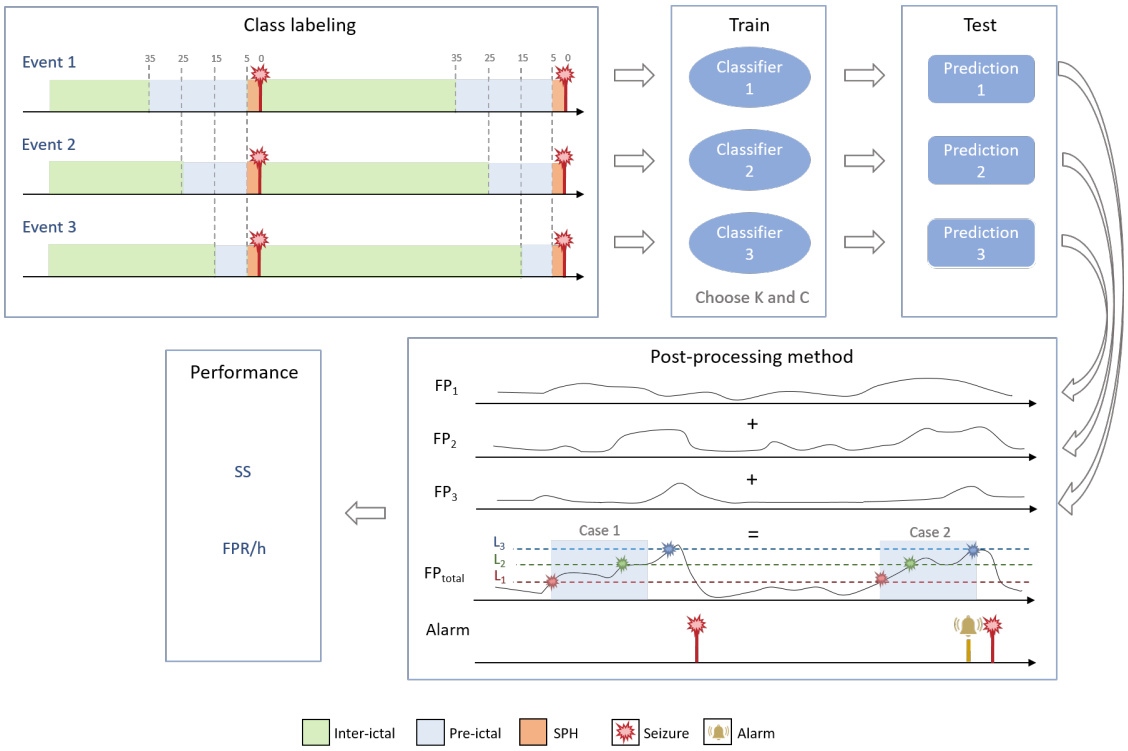

Figure 4.6: Approach B.

4.7.3 Approach C

Approach C aims to simulate a recurring state of the art approach in the epileptic seizure prediction field. It is considered a pre-ictal of equal duration to the adopted SOP. The pipeline developed is applied to this pre-ictal resulting in one

model output. After, it is applied the Firing Power filter and when a threshold (L) of 0.5 is exceeded, an alarm is raised. Finally, the performance is evaluated.

Figure 4.7 represents this approach for a 10-minute SOP.

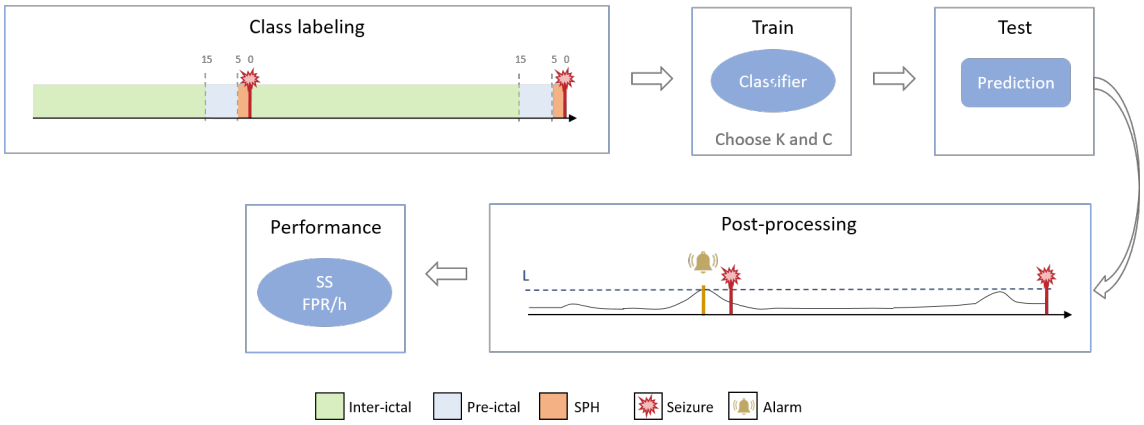

Figure 4.7: Approach C.

4.8 Performance

Prediction performance is based on two measures: seizure sensitivity and False Positive Rate per Hour (FPR/h) (formulations can be found in section 2.3.2). Their optimal values are 100% and 0 FPR/h respectively, which are associated with the correct forecast of all seizures without any false alarms. The used FPR/h definition incorporates the refractory behavior to simulate a real situation. Thus, for the same seizure, two near-consecutive alarms should only correspond to one, in order to reduce stress and confusion in the patient.

Concerning statistical validation, the methodology's performance was evaluated by comparing it against the surrogate time series analysis. For each test seizure, the pre-ictal labels were shifted randomly throughout the inter-ictal time and the associated seizure sensitivity was calculated. This process was carried out seizure by seizure so that the generated seizure times respect the seizure distribution over time. Thus considering a test set consisting of two seizures, with a total of 6 and 15 hours of data recorded respectively, the first seizure time (pre-ictal samples) will be randomly generated within the 6 hours associated with it and the second one in its corresponding 15 hours, instead of two times generated in the total recording time (21 hours).

For each seizure, times were shuffled 30 times and the average sensitivity was compared against the one obtained with the methodology under study. For a model to perform above chance level, it must achieve greater sensitivity with statistical

significance. Such significance is only reached when the null hypothesis "the performance of the proposed method is not superior to the performance of the surrogate predictor" is rejected. A significance level of 5% was considered.

A random predictor could also be used, but it is not based on EEG information and although the surrogate predictor is computationally heavier it allows to obtain a higher level of confidence when a model outperforms it [13].

5

Results and Discussion

This chapter presents the results obtained from the proposed methodology, both from the training phase (section 5.1) and the testing phase (section 5.2), as well as its interpretative analysis and discussion.

5.1 Training phase

The methodology described in the previous chapter was applied for three different time periods (10, 20 and 30 minutes) to the data of each patient. Therefore, for each event duration, three classifiers were trained for the first two approaches (A and B), and one classifier for the third approach (C).

The following Tables 5.1, 5.2, 5.3 summarize, for each classifier, the selected training parameters in the grid-search (cost of Support Vector Machines (SVM) and number of features) along with the obtained validation results (sample sensitivity and sample specificity). Those parameters were chosen with data from training seizures, which were the first three seizures of each patient. Only the results related to an event duration of 10 minutes will be presented in this chapter. The Tables B.1, B.2, B.3, B.4, B.5 and B.6 that contain the remaining results for 20-minute and 30-minute Seizure Occurrence Period (SOP)s can be found in Appendix B.

In order to facilitate the comparison between these approaches, the average sample sensitivity and specificity, for all the event durations, is presented in Table 5.4. Overall, although the average values of sensitivity and specificity are similar between the approaches, specificity was relatively higher than sensitivity, which means that there is a better classification of the inter-ictal samples.

In addiction, some standard deviations are considerably large, particularly for sensitivity. This means that although there may be some solutions with a high performance, there will also be others that perform poorly (some even below 0.5).

Table 5.1: Training parameters and performance of Approach A with events duration of 10 minutes.

	Approach A with 10-min events											
	Event 1				**Event 2**				**Event 3**			
Patient	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}
402	2^{-4}	10	0.67	0.54	2^{0}	20	0.67	0.58	2^{-16}	30	0.67	0.54
8902	2^{-8}	30	0.67	0.90	2^{-8}	10	0.61	0.83	2^{-8}	30	0.63	0.88
11002	2^{4}	20	0.81	0.68	2^{0}	10	0.67	0.74	2^{4}	10	0.67	0.75
16202	2^{0}	30	0.99	0.51	2^{4}	30	0.97	0.54	2^{-4}	10	1.00	0.50
23902	2^{8}	10	0.67	0.80	2^{-8}	10	0.67	0.78	2^{-8}	10	0.67	0.79
30802	2^{-20}	30	0.67	0.74	2^{-20}	30	0.66	0.74	2^{-12}	30	0.67	0.75
32702	2^{8}	10	0.35	0.72	2^{8}	10	0.35	0.71	2^{8}	30	0.42	0.63
46702	2^{-8}	30	0.71	0.84	2^{4}	20	0.67	0.90	2^{-12}	10	0.99	0.71
50802	2^{0}	20	0.94	0.58	2^{0}	30	0.81	0.66	2^{0}	30	0.85	0.63
53402	2^{8}	30	0.35	0.83	2^{4}	20	0.64	0.57	2^{8}	30	0.67	0.60
55202	2^{4}	10	0.65	0.61	2^{4}	20	0.57	0.80	2^{8}	30	0.69	0.61
56402	2^{-16}	20	0.67	0.60	2^{-20}	30	0.67	0.61	2^{-8}	10	0.67	0.61
58602	2^{8}	30	0.35	0.90	2^{8}	10	0.54	0.80	2^{0}	10	0.64	0.73
59102	2^{-20}	30	1.00	0.68	2^{-20}	30	1.00	0.71	2^{-20}	30	0.99	0.73
60002	2^{-12}	10	0.63	0.66	2^{4}	10	0.39	0.81	2^{8}	30	0.69	0.38
64702	2^{-20}	10	1.00	0.56	2^{-8}	10	1.00	0.56	2^{-4}	20	0.97	0.68
75202	2^{-8}	20	0.85	0.84	2^{8}	10	0.72	0.84	2^{8}	30	0.66	0.89
80702	2^{0}	10	0.32	0.90	2^{8}	20	0.47	0.79	2^{-20}	20	0.36	0.79
85202	2^{4}	20	0.65	0.86	2^{-12}	10	0.67	0.81	2^{-20}	10	0.67	0.81
93402	2^{-4}	10	0.67	0.67	2^{-16}	10	0.67	0.69	2^{8}	10	0.67	0.71
93902	2^{0}	20	0.67	0.72	2^{4}	30	0.60	0.75	2^{-4}	20	0.34	0.93
94402	2^{0}	30	0.48	0.60	2^{0}	20	0.40	0.65	2^{4}	30	0.65	0.59
95202	2^{-20}	10	0.67	0.60	2^{8}	20	0.69	0.63	2^{-20}	10	0.67	0.61
96002	2^{-12}	30	0.81	0.62	2^{-12}	20	0.99	0.58	2^{-16}	20	1.00	0.61
98102	2^{0}	10	0.64	0.63	2^{8}	10	0.62	0.61	2^{8}	10	0.62	0.62
98202	2^{-4}	10	0.64	0.80	2^{0}	10	0.66	0.76	2^{-4}	10	0.66	0.79
101702	2^{-20}	20	0.67	0.68	2^{-20}	20	0.67	0.68	2^{-20}	20	0.67	0.69
102202	2^{0}	20	0.66	0.72	2^{4}	20	0.57	0.88	2^{4}	20	0.65	0.81
104602	2^{4}	10	0.49	0.82	2^{4}	10	0.67	0.81	2^{8}	20	0.55	0.94
109502	2^{4}	10	0.67	0.8	2^{4}	30	0.26	0.90	2^{8}	20	0.33	0.91
110602	2^{-20}	20	0.62	0.76	2^{-16}	30	0.46	0.77	2^{-20}	30	0.66	0.76
112802	2^{-20}	20	0.92	0.59	2^{-20}	20	0.91	0.59	2^{-20}	20	0.84	0.59
113902	2^{-4}	10	0.59	0.81	2^{0}	20	0.65	0.74	2^{-4}	10	0.67	0.82
114702	2^{-4}	20	0.33	0.67	2^{-4}	20	0.33	0.67	2^{0}	30	0.34	0.67
114902	2^{-4}	30	0.40	0.71	2^{4}	30	0.57	0.82	2^{4}	10	0.51	0.60
123902	2^{-8}	30	0.67	0.75	2^{-4}	30	0.67	0.74	2^{-4}	30	0.66	0.74
Avg			0.65	0.71			0.64	0.72			0.67	0.71

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table 5.2: Training parameters and performance of Approach B with events duration of 10 minutes.

	Approach B with 10-min events											
	Event 1				**Event 2**				**Event 3**			
Patient	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}
402	2^{-4}	10	0.67	0.54	2^{-20}	30	0.67	0.54	2^{-20}	30	0.67	0.54
8902	2^{-8}	30	0.65	0.94	2^{-8}	30	0.67	0.92	2^{-8}	30	0.64	0.88
11002	2^{0}	10	0.67	0.76	2^{0}	10	0.67	0.75	2^{4}	10	0.67	0.75
16202	2^{-8}	10	1.00	0.56	2^{-4}	10	1.00	0.50	2^{-4}	10	1.00	0.50
23902	2^{-12}	30	0.66	0.79	2^{-12}	30	0.66	0.79	2^{-8}	10	0.67	0.79
30802	2^{-12}	30	0.67	0.74	2^{-12}	30	0.67	0.74	2^{-12}	30	0.67	0.75
32702	2^{0}	10	0.33	0.72	2^{4}	10	0.33	0.71	2^{8}	30	0.42	0.64
46702	2^{-4}	30	0.66	0.91	2^{-4}	20	0.67	0.88	2^{-12}	10	1.00	0.71
50802	2^{0}	20	0.91	0.62	2^{0}	20	0.94	0.60	2^{0}	20	0.87	0.61
53402	2^{-4}	20	0.66	0.60	2^{8}	20	0.67	0.59	2^{8}	30	0.67	0.60
55202	2^{4}	20	0.67	0.62	2^{0}	10	0.67	0.61	2^{8}	30	0.69	0.61
56402	2^{-8}	10	0.67	0.66	2^{-20}	20	0.67	0.62	2^{-8}	10	0.67	0.61
58602	2^{0}	30	0.67	0.61	2^{0}	10	0.67	0.63	2^{0}	10	0.60	0.73
59102	2^{0}	30	0.70	0.95	2^{-20}	30	0.99	0.74	2^{-20}	30	0.99	0.73
60002	2^{0}	30	0.67	0.51	2^{0}	30	0.67	0.52	2^{4}	30	0.70	0.40
64702	2^{-8}	10	1.00	0.61	2^{-8}	10	1.00	0.59	2^{-4}	20	0.98	0.69
75202	2^{-8}	30	0.73	0.88	2^{8}	30	0.67	0.89	2^{8}	20	0.66	0.89
80702	2^{-8}	20	0.41	0.78	2^{-8}	20	0.43	0.77	2^{-20}	30	0.37	0.79
85202	2^{-20}	10	0.75	0.82	2^{-12}	10	0.67	0.81	2^{-16}	10	0.67	0.81
93402	2^{8}	10	0.67	0.71	2^{8}	10	0.67	0.71	2^{8}	10	0.67	0.71
93902	2^{0}	10	0.46	0.86	2^{-4}	20	0.33	0.94	2^{-4}	20	0.33	0.94
94402	2^{8}	30	0.67	0.41	2^{4}	30	0.67	0.50	2^{4}	30	0.63	0.60
95202	2^{-20}	20	0.67	0.54	2^{-4}	10	0.66	0.51	2^{-20}	10	0.67	0.61
96002	2^{-12}	20	0.95	0.64	2^{-12}	20	1.00	0.61	2^{-16}	20	1.00	0.61
98102	2^{8}	30	0.67	0.55	2^{8}	20	0.67	0.54	2^{8}	10	0.63	0.62
98202	2^{-4}	10	0.66	0.77	2^{8}	10	0.67	0.75	2^{-4}	10	0.66	0.79
101702	2^{-20}	20	0.67	0.70	2^{-20}	20	0.67	0.69	2^{-20}	20	0.67	0.69
102202	2^{4}	10	0.66	0.85	2^{4}	30	0.67	0.76	2^{4}	20	0.67	0.80
104602	2^{0}	30	0.34	0.94	2^{4}	10	0.65	0.86	2^{4}	10	0.63	0.89
109502	2^{0}	20	0.33	0.88	2^{-4}	20	0.32	0.90	2^{8}	20	0.34	0.89
110602	2^{-20}	20	0.67	0.76	2^{-16}	20	0.67	0.76	2^{-20}	30	0.66	0.76
112802	2^{-20}	20	0.97	0.61	2^{-20}	20	0.90	0.60	2^{-20}	20	0.84	0.59
113902	2^{-4}	10	0.65	0.78	2^{-4}	10	0.66	0.79	2^{-4}	10	0.67	0.82
114702	2^{0}	30	0.33	0.73	2^{-4}	10	0.34	0.67	2^{0}	30	0.34	0.67
114902	2^{0}	10	0.61	0.65	2^{8}	20	0.53	0.82	2^{0}	10	0.51	0.60
123902	2^{-8}	30	0.67	0.75	2^{8}	10	0.66	0.74	2^{-4}	30	0.66	0.74
Avg			0.66	0.72			0.67	0.70			0.67	0.70

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table 5.3: Training parameters and performance of Approach C with 10-min pre-ictal.

	Approach C with 10-min pre-ictal			
Patient	C	k	SS_{sample}	SP_{sample}
402	2^{-20}	30	0.67	0.54
8902	2^{-8}	30	0.63	0.88
11002	2^{4}	10	0.67	0.75
16202	2^{-4}	10	1.00	0.50
23902	2^{-8}	10	0.67	0.79
30802	2^{-12}	30	0.67	0.75
32702	2^{8}	20	0.51	0.58
46702	2^{-12}	10	1.00	0.71
50802	2^{0}	20	0.86	0.62
53402	2^{8}	30	0.67	0.60
55202	2^{8}	30	0.69	0.60
56402	2^{-8}	10	0.67	0.61
58602	2^{4}	10	0.67	0.69
59102	2^{-20}	30	0.99	0.73
60002	2^{8}	30	0.70	0.38
64702	2^{-4}	20	0.97	0.69
75202	2^{8}	20	0.66	0.89
80702	2^{-20}	20	0.36	0.79
85202	2^{-20}	10	0.67	0.81
93402	2^{8}	10	0.67	0.71
93902	2^{-4}	20	0.34	0.94
94402	2^{4}	30	0.65	0.59
95202	2^{-20}	10	0.67	0.61
96002	2^{-16}	20	1.00	0.61
98102	2^{8}	10	0.60	0.63
98202	2^{-4}	10	0.65	0.79
101702	2^{-20}	20	0.67	0.69
102202	2^{4}	20	0.66	0.81
104602	2^{8}	20	0.63	0.93
109502	2^{8}	20	0.33	0.90
110602	2^{-20}	30	0.66	0.76
112802	2^{-20}	20	0.84	0.59
113902	2^{-4}	10	0.66	0.82
114702	2^{0}	20	0.34	0.67
114902	2^{0}	10	0.51	0.60
123902	2^{-4}	30	0.67	0.74
Avg			0.67	0.70

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Additionally, it is worth noting that, as expected, the results related to the third classifiers (third event) of approaches A and B and the classifier of approach C are quite similar, since they relate to the same pre-ictal samples, so only differing in the inter-ictal samples. They show small fluctuations in both uncertainties and averages due to the stochasticity inherent in the class balancing process via undersampling.

One would expect more satisfactory results, and that the classifiers would be able to better distinguish the two classes. However, only in the testing phase will it be possible to ascertain their predictive power, since sometimes high performances in training can mean problems such as overfitting.

Table 5.4: Average validation results according to the approach and event /pre-ictal duration.

Event / pre-ictal duration (min)	Approach	Event 1		Event 2		Event 3		Pre-ictal	
		$\mathbf{SS}_{sample}$	$\mathbf{SP}_{sample}$	$\mathbf{SS}_{sample}$	$\mathbf{SP}_{sample}$	$\mathbf{SS}_{sample}$	$\mathbf{SP}_{sample}$	$\mathbf{SS}_{sample}$	$\mathbf{SP}_{sample}$
10	A	0.65 ± 0.18	0.71 ± 0.11	0.64 ± 0.18	0.72 ± 0.10	0.67 ± 0.18	0.71 ± 0.12	-	-
	B	0.66 ± 0.17	0.72 ± 0.13	0.67 ± 0.17	0.70 ± 0.13	0.67 ± 0.18	0.70 ± 0.12	-	-
	C	-	-	-	-	-	-	0.67 ± 0.17	0.70 ± 0.13
20	A	0.66 ± 0.18	0.70 ± 0.11	0.68 ± 0.19	0.71 ± 0.12	0.67 ± 0.17	0.70 ± 0.13	-	-
	B	0.71 ± 0.16	0.70 ± 0.10	0.70 ± 0.17	0.71 ± 0.11	0.67 ± 0.18	0.70 ± 0.13	-	-
	C	-	-	-	-	-	-	0.67 ± 0.17	0.70 ± 0.13
30	A	0.67 ± 0.19	0.70 ± 0.12	0.67 ± 0.18	0.71 ± 0.13	0.66 ± 0.16	0.72 ± 0.13	-	-
	B	0.71 ± 0.17	0.71 ± 0.11	0.71 ± 0.16	0.69 ± 0.10	0.67 ± 0.17	0.71 ± 0.14	-	-
	C	-	-	-	-	-	-	0.67 ± 0.17	0.71 ± 0.14

SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

5.2 Testing phase

After the training phase, the proposed methodology was tested on a new set of data (testing data), unknown to the created Machine Learning models. In total 106 seizures were tested comprising 1533.8 hours of data. At this phase, performance was accessed in terms of seizure sensitivity, False Positive Rate per Hour (FPR/h), and statistical validation by using a surrogate analysis to access performance above chancel level.

Performance was evaluated for a range of SOP values. Tables 5.5, 5.6, 5.7 show the testing results of each approach for the 36 patients, for a 10-minute SOP and a 5-minute Seizure Prediction Horizon (SPH). As in the section above, the remaining results for the 20 (Tables B.7, B.8 and B.9) and 30-minute (Tables B.10, B.11 and B.12) SOPs can be found in Appendix B.

Averaging across all 36 patients, seizure sensitivity and FPR/h for approach A with a 10-minute SOP are 0.20 ± 0.28 and 3.13 ± 7.19, respectively. The sensitivity of surrogate predictors averaged at 0.12 ± 0.03, where only 13 out of 36 patients (36%) beat the surrogate predictor with statistical significance.

Table 5.5: Seizure prediction performance and statistical validation results of Approach A with 10-min SOP.

Approach A with 10-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.50	0.08	0.02 ± 0.09	< 0.01	Yes
8902	2	0.00	0.09	0.00 ± 0.00	-	No
11002	2	0.00	0.07	0.00 ± 0.00	-	No
16202	4	0.00	0.00	0.00 ± 0.00	-	No
23902	2	0.50	20.38	0.43 ± 0.36	0.32	No
30802	5	0.00	7.24	0.33 ± 0.15	-	No
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.50	0.00	0.00 ± 0.00	< 0.01	Yes
50802	2	0.00	0.69	0.10 ± 0.20	-	No
53402	2	0.00	0.06	0.00 ± 0.00	-	No
55202	5	0.00	0.21	0.02 ± 0.06	-	No
56402	3	0.00	0.18	0.02 ± 0.08	-	No
58602	4	0.00	0.00	0.00 ± 0.00	-	No
59102	2	0.50	10.61	0.28 ± 0.25	< 0.01	Yes
60002	3	0.33	0.58	0.11 ± 0.16	< 0.01	Yes
64702	2	0.00	3.59	0.25 ± 0.25	-	No
75202	4	0.00	1.02	0.08 ± 0.13	-	No
80702	3	0.00	0.00	0.00 ± 0.00	-	No
85202	2	0.00	0.00	0.00 ± 0.00	-	No
93402	2	0.00	2.70	0.17 ± 0.24	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.00	0.00 ± 0.00	-	No
95202	4	0.25	0.72	0.18 ± 0.16	0.03	Yes
96002	4	0.50	7.94	0.38 ± 0.25	0.01	Yes
98102	2	0.00	0.36	0.05 ± 0.15	-	No
98202	4	0.25	2.11	0.10 ± 0.12	< 0.01	Yes
101702	2	1.00	8.52	0.48 ± 0.35	< 0.01	Yes
102202	4	0.50	1.68	0.17 ± 0.16	< 0.01	Yes
104602	2	0.50	2.07	0.33 ± 0.24	< 0.01	Yes
109502	2	0.00	2.42	0.17 ± 0.24	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	1.00	38.18	0.43 ± 0.21	< 0.01	Yes
113902	4	0.25	0.55	0.05 ± 0.10	< 0.01	Yes
114702	6	0.00	0.14	0.01 ± 0.04	-	No
114902	4	0.50	0.41	0.08 ± 0.15	< 0.01	Yes
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Regarding approach B, sensitivity averaged at 0.41 ± 0.30 and FPR/h at 8.39 ± 25.54. Mean sensitivity for the surrogate predictors was 0.20 ± 0.03. For 27 patients (75%), the sensitivity of the approach outperformed the average sensitivity of the surrogate predictor, where in 25 (69%) of them was possible to observe a performance above chance.

Table 5.6: Seizure prediction performance and statistical validation results of Approach B with 10-min SOP.

Approach B with 10-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.50	0.46	0.07 ± 0.17	< 0.01	Yes
8902	2	0.50	1.06	0.12 ± 0.28	< 0.01	Yes
11002	2	0.50	0.50	0.03 ± 0.12	< 0.01	Yes
16202	4	0.25	0.62	0.07 ± 0.11	< 0.01	Yes
23902	2	1.00	18.38	0.57 ± 0.36	< 0.01	Yes
30802	5	0.40	7.50	0.32 ± 0.17	0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.50	0.35	0.07 ± 0.17	< 0.01	Yes
50802	2	1.00	2.82	0.42 ± 0.26	< 0.01	Yes
53402	2	0.50	1.98	0.28 ± 0.25	< 0.01	Yes
55202	5	0.60	5.25	0.31 ± 0.16	< 0.01	Yes
56402	3	0.00	0.29	0.04 ± 0.11	-	No
58602	4	0.25	0.04	0.00 ± 0.00	< 0.01	Yes
59102	2	0.50	11.29	0.32 ± 0.24	< 0.01	Yes
60002	3	0.67	153.26	0.63 ± 0.28	0.51	No
64702	2	0.50	5.08	0.23 ± 0.25	< 0.01	Yes
75202	4	0.50	1.11	0.09 ± 0.14	< 0.01	Yes
80702	3	0.33	3.71	0.21 ± 0.16	< 0.01	Yes
85202	2	0.50	3.45	0.28 ± 0.28	< 0.01	Yes
93402	2	0.00	2.75	0.18 ± 0.24	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.28	0.05 ± 0.01	-	No
95202	4	0.50	5.72	0.33 ± 0.22	< 0.01	Yes
96002	4	0.50	10.07	0.47 ± 0.23	0.55	No
98102	2	0.50	1.27	0.07 ± 0.17	< 0.01	Yes
98202	4	0.25	2.17	0.08 ± 0.12	< 0.01	Yes
101702	2	1.00	9.88	0.53 ± 0.31	< 0.01	Yes
102202	4	0.50	2.69	0.28 ± 0.22	< 0.01	Yes
104602	2	0.50	2.22	0.27 ± 0.25	< 0.01	Yes
109502	2	0.00	2.48	0.15 ± 0.23	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	1.00	40.31	0.36 ± 0.27	< 0.01	Yes
113902	4	0.75	1.04	0.11 ± 0.14	< 0.01	Yes
114702	6	0.17	0.14	0.03 ± 0.08	< 0.01	Yes
114902	4	0.25	4.02	0.27 ± 0.21	-	No
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Concerning approach C, overall a seizure sensitivity of 0.32 ± 0.29 and an FPR/h of 5.39 ± 10.01 were obtained. The sensitivity of surrogate predictors averaged at 0.21 ± 0.03 and 20 patients (56%) were statistical validated.

Table 5.7: Seizure prediction performance and statistical validation results of Approach C with 10-min SOP.

Approach C with 10-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.00	0.46	0.02 ± 0.09	-	No
8902	2	0.50	0.80	0.02 ± 0.09	< 0.01	Yes
11002	2	0.00	0.58	0.07 ± 0.17	-	No
16202	4	0.25	0.58	0.07 ± 0.11	< 0.01	Yes
23902	2	0.50	20.38	0.47 ± 0.41	0.66	No
30802	5	0.40	7.37	0.31 ± 0.17	0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.50	0.65	0.05 ± 0.15	< 0.01	Yes
50802	2	0.50	5.25	0.48 ± 0.30	0.76	No
53402	2	0.50	1.98	0.38 ± 0.25	0.02	Yes
55202	5	0.80	3.21	0.29 ± 0.18	< 0.01	Yes
56402	3	0.00	0.34	0.03 ± 0.10	-	No
58602	4	0.00	0.07	0.01 ± 0.04	-	No
59102	2	0.00	10.77	0.30 ± 0.24	-	No
60002	3	1.00	42.33	0.60 ± 0.26	< 0.01	Yes
64702	2	0.50	3.47	0.20 ± 0.24	< 0.01	Yes
75202	4	0.25	1.08	0.11 ± 0.14	< 0.01	Yes
80702	3	0.33	3.71	0.23 ± 0.15	< 0.01	Yes
85202	2	0.00	3.45	0.33 ± 0.27	-	No
93402	2	0.00	2.70	0.13 ± 0.22	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.20	0.03 ± 0.08	-	No
95202	4	0.25	5.72	0.46 ± 0.19	-	No
96002	4	0.50	8.71	0.48 ± 0.23	0.7	No
98102	2	0.50	0.55	0.13 ± 0.22	< 0.01	Yes
98202	4	0.25	2.17	0.09 ± 0.12	< 0.01	Yes
101702	2	1.00	9.40	0.53 ± 0.31	< 0.01	Yes
102202	4	0.50	2.69	0.29 ± 0.18	< 0.01	Yes
104602	2	0.50	2.22	0.35 ± 0.23	< 0.01	Yes
109502	2	0.50	5.04	0.25 ± 0.25	< 0.01	Yes
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	0.67	43.94	0.41 ± 0.24	< 0.01	Yes
113902	4	0.25	0.60	0.07 ± 0.11	< 0.01	Yes
114702	6	0.17	0.05	0.01 ± 0.03	< 0.01	Yes
114902	4	0.50	3.65	0.30 ± 0.20	< 0.01	Yes
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

5.3 Comparative analysis between approaches

A summary in terms of sensitivity and performance above chance for each of the three approaches and SOP values is presented for all patients in Figure 5.1. Table 5.8 presents the overall performance of the implemented methodology. Figure 5.2

shows the information presented in Table 5.8, allowing a more intuitive view of the obtained performance.

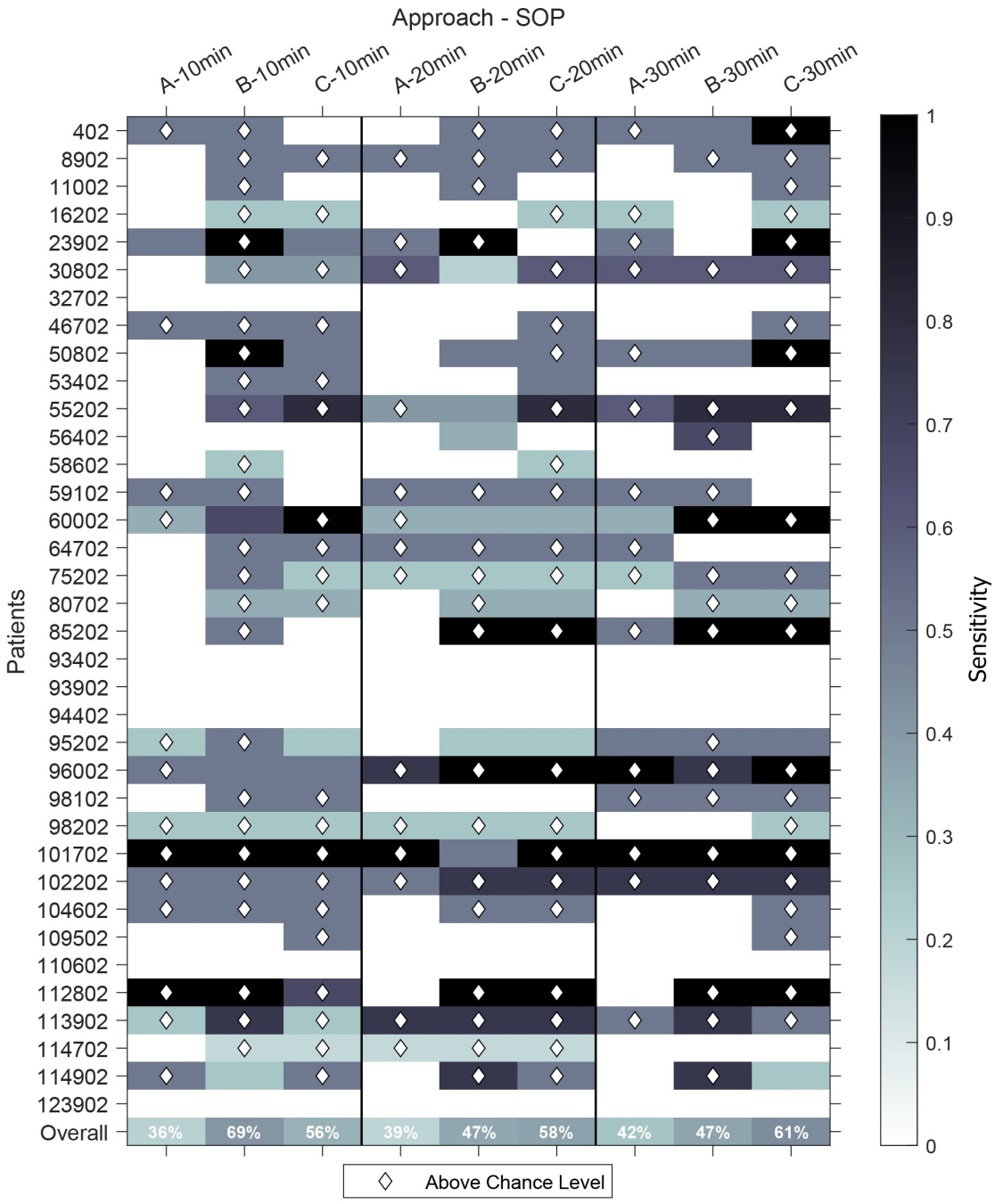

Figure 5.1: Sensitivity of seizure prediction and results of statistical validation of all post-processing approaches for each patient. The blue color scale refers to the sensitivity achieved and the diamond shape is present in the models that outperformed the surrogate predictor. On the overall column, one can see the percentage of patients whose models perform above chance level in white and the average sensitivity by the colors of the cells.

There are patients who have not been validated by any method, regardless of SOP values or approaches: 32702, 93402, 93902, 94402, 110602, 123902 (6 patients, 17%). For 83% of patients, there was at least one approach that was better than chance level.

The best result among all the adopted methodologies and SOP values was obtained for approach B for a SOP of 10 minutes. There are 8 patients (402, 11002, 23902, 50802, 58602, 59102, 85202, 95202) in which this approach managed to be better than the chance level, where the control approach did not. The opposite, that is, the methodology C being better than chance level but the approach B not, only occurred in three cases (60002, 109502, 114902).

Table 5.8: Seizure prediction average performance according to the approach and SOP duration.

SOP duration (min)	Approach	All Patients			Validated patients		
		$SS_{seizure}$	FPR/h	Surrogate $SS_{seizure}$	%	$SS_{seizure}$	FPR/h
10	A	0.20 ± 0.28	3.13 ± 7.19	0.12 ± 0.03	36.11%	0.51 ± 0.23	5.65 ± 10.01
	B	0.41 ± 0.30	8.39 ± 25.54	0.20 ± 0.03	69.44%	0.54 ± 0.23	5.16 ± 8.31
	C	0.32 ± 0.29	5.39 ± 10.01	0.21 ± 0.03	55.56%	0.49 ± 0.23	6.77 ± 12.34
20	A	0.19 ± 0.28	0.86 ± 1.58	0.10 ± 0.02	38.89%	0.50 ± 0.22	1.91 ± 2.10
	B	0.35 ± 0.33	2.98 ± 4.30	0.26 ± 0.03	47.22%	0.60 ± 0.27	3.76 ± 5.56
	C	0.37 ± 0.33	9.97 ± 45.24	0.24 ± 0.03	58.33%	0.57 ± 0.26	3.01 ± 4.85
30	A	0.26 ± 0.31	0.84 ± 1.31	0.17 ± 0.03	41.67%	0.56 ± 0.21	1.54 ± 1.66
	B	0.36 ± 0.37	2.30 ± 3.32	0.27 ± 0.03	47.22%	0.71 ± 0.19	3.59 ± 3.78
	C	0.44 ± 0.38	2.11 ± 3.22	0.27 ± 0.03	61.11%	0.68 ± 0.27	2.68 ± 3.67

$SS_{seizure}$: seizure sensitivity; Surrogate $SS_{seizure}$: seizure sensitivity from surrogate predictor;
%: percentage of validated patients

A good seizure predictor must have a high sensitivity and FPR/h as low as possible. For patients in pre-surgical monitoring, as is the case, a FPR/h of 0.15 is considered reasonable for a warning system [6]. By a quick examination of the Table 5.8 it is easy to note that the values obtained for the FPR/h do not satisfy this condition. This means that the number of false alarms raised by the models may be higher than what is considered acceptable for a warning system. This goes for the average values as to its considerably large standard deviations. However, these high values may have resulted from the fact that all available test data were used, totaling hours and sometimes even days under analysis.

Once again, the sensitivity values show a wide dispersion of the results, indicating that although there may be solutions with a good performance, there are others that perform very poorly. This is easily detected by the number of patients with zero sensitivity in the previous Figure 5.1.

In addition to having a high sensitivity and low FPR/h, the methodology must

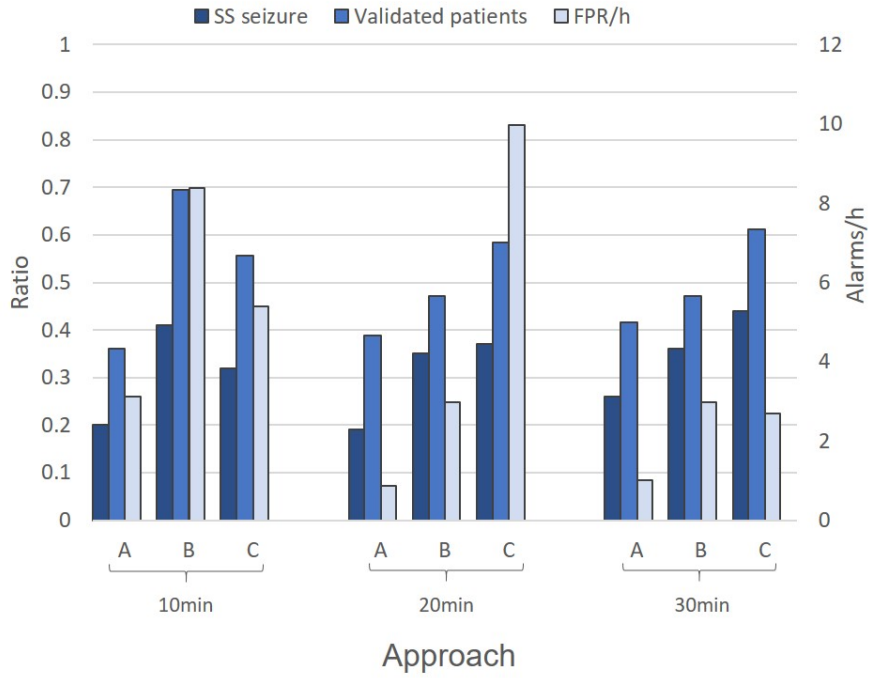

Figure 5.2: Average seizure sensitivity, average FPR/h and percentage of validated patients obtained for all approaches and SOP durations.

also be able to outperform a surrogate predictor. Mean sensitivity obtained were higher than the average sensitivity surrogate predictor. That is, even with high FPR/h values, some of the created models were able to perform above chance.

For 20-minute and 30-minute SOPs, approach C was the one with the best sensitivity and the largest number of validated patients. Regarding FPR/h, approach A showed the lowest values. This was due to a greater number of patients not having triggered any type of alarm (false or true), thus lowering the overall value, but also of having more patients with low FPR/h than in other approaches.

For the 10-minute SOP, approach B achieved the best performance in terms of validated patients (about 69%) and the second best sensitivity value (around 41%). However, the value of FPR/h is quite high, which may or may not be a problem depending on the final application. In the case of a warning device, the FPR/h value obtained is not adequate. For a brain electrostimulation device, which turns out to have no significant side effects for the patient, the FPR/h may possibly be higher. However, this issue needs to be further studied in order to discover the maximum limit of FPR/h for the different applications.

Regarding the percentage of validated patients, approach B with a 10-minute SOP was the best performance in relation to all the attempted approaches with all SOPs. Although it does not perform better for all SOPs, it clearly appears to have

a superior advantage for small ones. It is a relevant result for application reasons, making it more suitable for implementation in warning device, since the patient spends less time waiting for the supposed seizure. For a treatment system, a small period means shorter intervention time, therefore fewer side effects, if any.

In relation to control C, the performance improves as the SOP increases.This is consistent with what has been observed in the literature and hence there are more and more studies with larger SOPs. But does it make sense, considering an application or even the will of the patients? Patients clearly prefer small SOPs, as they have less influence on their daily lives when on the verge of a seizure [83].

5.3.1 Comparative analysis with other studies

These results can be compared with the performance of previous articles, mentioned in Chapter 3. For this purpose, two studies were chosen. They are both based on the EPILEPSIAE database and incorporate Temporal Lobe Epilepsy (TLE) patients into their data. Table 5.9 summarizes the seizure prediction metrics of those studies. Only performance measures that concern TLE patients will be compared since the dataset evaluated in this study contained this restriction.

Table 5.9: Seizure prediction performance for studies under comparison.

	All patients			**TLE patients**			**Validated TLE patients**		
Study	**#**	**SS$_{seizure}$**	**FPR/h**	**#**	**SS$_{seizure}$**	**FPR/h**	**%**	**SS$_{seizure}$**	**FPR/h**
Direito et al.[23] (2017)	216	0.38	0.20	130	0.38	0.23	N.A.	N.A.	N.A.
Alvarado Rojas et al.[41] (2014)	53	0.73	0.29	32	N.A.	N.A.	10%	0.66	0.33
This study (Approach B with 10-min SOP)	N.A.	N.A.	N.A.	36	0.41	8.39	69%	0.54	5.16

SS$_{seizure}$: seizure sensitivity; %: percentage of validated patients; #: number of patients

In Direito et al.[23] study, for the 130 temporal lobe epileptic patients, a sensitivity of 0.38 and an FPR/h of 0.23 were achieved. Considering the results of Table 5.8, it is clear that only approach B for a 10-minute SOP and approach C for a 30-minute SOP were able to overcome this performance in terms of sensitivity. As for the values of FPR/h neither approached such value. The authors only indicate that the implemented methodology performance reached statistical significance for 11% of the patients, without specifying the percentage of validated TLE patients.

It is also important to compare our control approach (approach C) with the Direito et al. [23] study, since the methodologies are similar. In the referenced study, a pre-selection of three sets of six channels was made to which 22 linear features were extracted from each Electroencephalogram (EEG) channel. In our work, the EEG electrodes selection was not performed, instead a selection of the features with best relationship with the target was accomplished. The training phases were similar,

except that Direito et al.[23] opted for a multiclass SVM thus using data from the four classes (pre-ictal, ictal, post-ictal and inter-ictal) while, in this work, it was opted to only build a target of two classes, inter-ictal and pre-ictal. In addition, in our study, due to computational limitations, the highest value for C used was 2^8, not having such high values nor its search divided into two parts. Furthermore, in the present work, the stochasticity of random undersampling of the inter-ictal class was dealt through the use of essemble learning, reason why, possibly, our results are more robust to noise. When presenting the results, it was decided to show the results obtained for each SOP value. However, Direito et al.[23] made a selection of the best SVM model among all those created (with different SOP values and different sets of electrodes) based on the best testing performance. This positively biased the results presented, since neither a general average of the performance of all classifiers was made nor an average according to the value of SOP. Despite the very limited comparison, approach C for a 30-minute SOP achieved a better performance in terms of sensitivity than that reported in the study. Regarding the number of validated patients, our approach (using a surrogate predictor) reached much higher values.

For the Alvarado-Rojas et al.[41] study, comparing the performance of validated TLE patients, only approach B and C for a 30-minute SOP outperformed the sensitivity value of 0.66, but none of the methods achieved a better FPR/h. In this study, 13.2% of all patients were validated, where only 10% were significant for cases of temporal lobe epilepsy which is a lower value than the ones obtained with the developed approaches.

It is important to highlight that the comparison with other studies is not a simple task since there are many factors and decisions made differently: from the number of patients, the number of EEG records, the way the training is performed and even the values chosen for SOP and SPH, among others.

The SPH value was only reported for the first study and is 30 times smaller (10 seconds) than the SPH adopted in this methodology. Unlike the SPH used by Direito et al.[23], an SPH of 5 minutes allows a reasonable time for the patient to be able to prepare.

SOP values are another factor under discussion, as short durations are preferable in real-life applications. However, in state of the art studies the best results were often reported for higher SOPs, especially in relation to FPR/h (this can also be observed in our results). The range of values is enormous, and the ones used in these two cases were 10, 20, 30 and 40 minutes for Direito et al.[23] and 60 minutes for Alvarado-Rojas et al.[41] study.

In addition, it should be noted the two studies did not present the results against a surrogate predictor, which is the case presented in our study. So the direct comparison may not be fair. However, the surrogate predictor reveals to be more flexible [15], adapting to the study conditions, the used system, and the data (in terms of the number of seizures and recording time), making the validation more solid.

5.4 Patient stratification

As mentioned in chapter 3, the EPILEPSIAE database, in addition to the long-term records of the EEG, also incorporates metadata information and clinical annotations. This allowed grouping the patients with similar characteristics to ascertain if there are patterns or differences with the overall results.

Stratification was carried out according to seizure classification, seizure activity pattern, vigilance state, and circadian cycle (present in Table 4.1 and followed the criteria:

1. patients whose seizures were classified as Focal Onset Aware (FOA) and Focal Onset Impaired Awareness (FOIA);
2. patients who only showed rhythmic activity patterns in their seizures;
3. patients who were awake during seizures;
4. patients who had seizures occurring during the day in both sets of data;
5. patients who had seizures during night in both training and testing data.

There were 21, 23, 14, 27 and 24 patients who met the first, second, third, fourth and fifth criteria, respectively. The constitution of each group is shown in Appendix B in Table B.13.

Table 5.10 shows the results obtained for each of the mentioned groups according to the method and the value of SOP duration. The Figure 5.3 displays the average sensitivities and FPR/h for each patient group.

Analyzing both the table and figure, it appears that the group of patients with only FOA and/or FOIA seizure has higher sensitivity averages than the average for all patients, for all approaches, and all SOP durations. This could be considered an improvement in performance, if the FPR/h values were maintained or decreased compared to the values for all patients, which however is not the case, having in some cases suffered quite significant variations.

The second group, made up of patients with rhythmic activity patterns in their seizures, has FPR/h values similar to or less than the mean values for some of

the approaches, however, has lower sensitivity values. Regarding the percentage of validated patients, some of the stratifications showed a substantial improvement. The first and fifth groups showed values that were above the average of all patients, with the fifth group having the largest number of validated patients: about 79% for method B and a SOP of 10 minutes. The second group always presented values that were lower or similar to the general average.

No significant changes were found in the average sensitivity of the surrogate predictors.

In general, even after stratification, no viable values were obtained for medical application with the proposed methodology. In addition, standard deviations remained high.

Table 5.10: Patient stratification results according to the approach and SOP duration.

SOP duration (min)	Approach	Stratification	$SS_{seizure}$	FPR/h	Surrogate $SS_{seizure}$	% Validated patients
10	A	FOA/FOIA	0.22 ± 0.32	4.75 ± 9.03	0.14 ± 0.04	33.33%
		Rhythmic patterns	0.22 ± 0.28	4.27 ± 8.62	0.13 ± 0.03	21.74%
		Awake	0.23 ± 0.31	5.72 ± 10.67	0.12 ± 0.04	64.29%
		Day	0.23 ± 0.30	3.74 ± 8.11	0.13 ± 0.03	40.74%
		Night	0.19 ± 0.30	3.04 ± 7.64	0.13 ± 0.03	37.50%
	B	FOA/FOIA	0.47 ± 0.31	12.89 ± 32.66	0.23 ± 0.05	71.43%
		Rhythmic patterns	0.39 ± 0.33	4.73 ± 8.79	0.18 ± 0.04	65.22%
		Awake	0.42 ± 0.31	6.78 ± 10.68	0.19 ± 0.05	71.43%
		Day	0.41 ± 0.30	4.91 ± 8.16	0.20 ± 0.04	70.37%
		Night	0.44 ± 0.30	10.42 ± 30.82	0.20 ± 0.04	79.17%
	C	FOA/FOIA	0.36 ± 0.32	7.75 ± 12.49	0.24 ± 0.04	57.14%
		Rhythmic patterns	0.26 ± 0.23	5.09 ± 9.51	0.19 ± 0.04	47.83%
		Awake	0.31 ± 0.27	6.83 ± 11.70	0.18 ± 0.05	50.00%
		Day	0.30 ± 0.28	4.96 ± 8.83	0.21 ± 0.04	51.85%
		Night	0.37 ± 0.28	6.01 ± 11.46	0.22 ± 0.04	70.83%
20	A	FOA/FOIA	0.30 ± 0.31	1.17 ± 1.94	0.12 ± 0.03	57.14%
		Rhythmic patterns	0.19 ± 0.26	1.00 ± 1.68	0.11 ± 0.03	39.13%
		Awake	0.21 ± 0.26	1.03 ± 1.91	0.09 ± 0.03	42.86%
		Day	0.21 ± 0.29	0.89 ± 1.68	0.10 ± 0.03	40.74%
		Night	0.20 ± 0.29	0.78 ± 1.27	0.11 ± 0.03	41.67%
	B	FOA/FOIA	0.42 ± 0.36	3.79 ± 5.31	0.28 ± 0.05	52.38%
		Rhythmic patterns	0.36 ± 0.35	3.24 ± 5.05	0.24 ± 0.04	52.17%
		Awake	0.42 ± 0.36	4.23 ± 6.06	0.26 ± 0.05	57.14%
		Day	0.42 ± 0.34	3.12 ± 4.69	0.27 ± 0.04	59.26%
		Night	0.34 ± 0.30	3.16 ± 4.63	0.27 ± 0.04	45.63%
	C	FOA/FOIA	0.45 ± 0.37	16.24 ± 58.42	0.27 ± 0.04	61.90%
		Rhythmic patterns	0.34 ± 0.33	2.56 ± 4.82	0.20 ± 0.04	56.52%
		Awake	0.37 ± 0.35	3.64 ± 5.86	0.21 ± 0.04	57.14%
		Day	0.40 ± 0.35	2.71 ± 4.48	0.23 ± 0.03	59.26%
		Night	0.40 ± 0.32	13.86 ± 54.96	0.26 ± 0.04	66.67%
30	A	FOA/FOIA	0.31 ± 0.34	1.00 ± 1.39	0.20 ± 0.04	47.62%
		Rhythmic patterns	0.22 ± 0.30	0.67 ± 1.23	0.14 ± 0.03	39.13%
		Awake	0.28 ± 0.31	0.71 ± 1.39	0.13 ± 0.04	50.00%
		Day	0.26 ± 0.33	0.87 ± 1.41	0.17 ± 0.03	40.74%
		Night	0.26 ± 0.29	0.85 ± 1.22	0.17 ± 0.03	41.67%
	B	FOA/FOIA	0.44 ± 0.41	2.99 ± 4.07	0.31 ± 0.05	57.14%
		Rhythmic patterns	0.28 ± 0.34	2.35 ± 3.85	0.24 ± 0.04	34.78%
		Awake	0.40 ± 0.33	3.5 ± 4.61	0.32 ± 0.06	57.14%
		Day	0.36 ± 0.38	2.49 ± 3.63	0.28 ± 0.04	48.15%
		Night	0.40 ± 0.38	2.16 ± 3.42	0.27 ± 0.04	54.17%
	C	FOA/FOIA	0.51 ± 0.40	2.69 ± 3.92	0.29 ± 0.05	66.67%
		Rhythmic patterns	0.42 ± 0.39	2.20 ± 3.73	0.25 ± 0.04	60.87%
		Awake	0.47 ± 0.40	2.94 ± 4.48	0.28 ± 0.05	71.43%
		Day	0.42 ± 0.39	2.20 ± 3.50	0.27 ± 0.04	55.56%
		Night	0.47 ± 0.35	2.23 ± 3.52	0.28 ± 0.04	70.83%

$SS_{seizure}$: seizure sensitivity; Surrogate $SS_{seizure}$: seizure sensitivity from surrogate predictor;
% Validated patients: percentage of validated patients

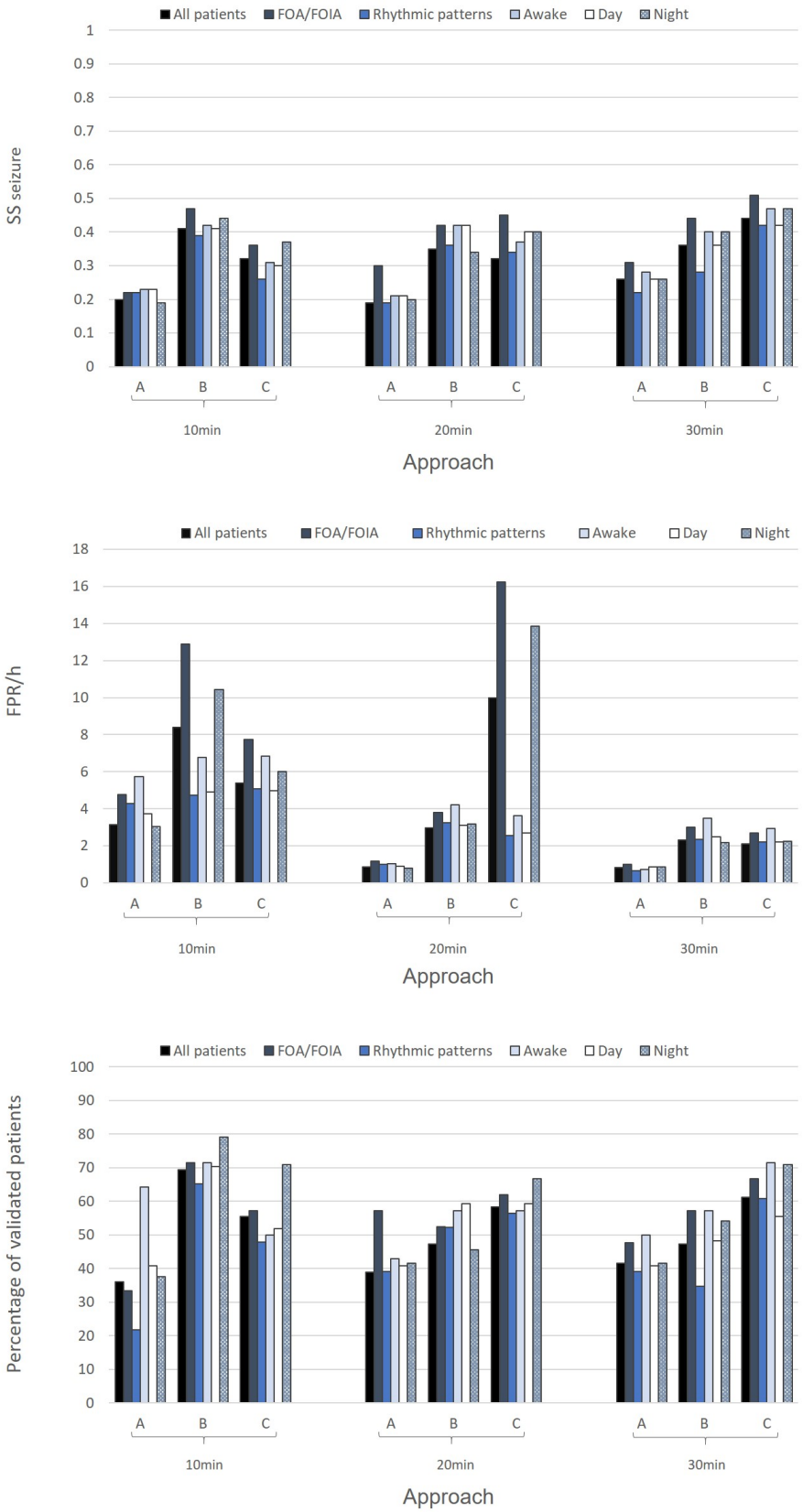

Figure 5.3: Average seizure sensitivity, average FPR/h and percentage of validated patients obtained from patient stratification for all approaches and SOP durations.

6

Conclusion

The proposed patient-specific seizure prediction algorithm was intended to explore the concept of chronology of brain activity in a post-processing phase to increase the complexity of the models while maintaining interpretability. To this end, two different post-processing methods were built based on the Firing Power regularization method. Therefore, approaches A and B were developed and compared to control approach (C).

Between the two post-processing methods, approach B achieved better performance in terms of sensitivity and percentage of validated patients. In relation to all methodologies and Seizure Occurrence Period (SOP)s, for a 10-minute SOP, this approach achieved the best result, in terms of performance above chance level (69% of patients) and the second best result in terms of sensitivity (41%). The False Positive Rate per Hour (FPR/h) values were not very satisfactory for implementation in patient warning systems. However, they may be valid for intervention systems, such as the case of electrostimulation, if these do not provoke negative consequences to the patient's health.

Lately, state of the art seizure prediction studies have tended to opt for larger SOPs in order to escape worse performance. The fact that in the present work, smaller SOPs (10, 20 and 30 minutes) were used prioritizes the patients' will and reduces the waiting/intervention time for a seizure. It is also concluded with the control approach that the surrogate predictor is able to validate more patients than the random predictor used in other previous studies with Temporal Lobe Epilepsy (TLE) patients.

In the future, to assess the applicability of this methodology, it remains to be studied which maximum FPR/h value can be applied in each intervention strategy without causing harmful consequences for patients. Furthermore, in the developed post-processing methods, one could change the thresholds imposed by thresholds adapted to each patient and evaluate the differences in performance.

In addition, the study must be reproduced in real data instead of pre-surgical

monitoring data, that is, in long-term daily data containing seizures occurring naturally. The application of this methodology in databases closer to real life, such as the case of Neurovista, used in studies by Cook et al.[16] and Kiral-Kornek et al.[10] would allow not only the comparison with more studies, but also the more rigorous evaluation of the algorithm.

Bibliography

[1] R. S. Fisher, J. H. Cross, J. A. French, N. Higurashi, E. Hirsch, F. E. Jansen, L. Lagae, S. L. Moshé, J. Peltola, E. Roulet Perez, *et al.*, "Operational classification of seizure types by the international league against epilepsy: Position paper of the ilae commission for classification and terminology," *Epilepsia*, vol. 58, no. 4, pp. 522–530, 2017.

[2] I. E. Scheffer, S. Berkovic, G. Capovilla, M. B. Connolly, J. French, L. Guilhoto, E. Hirsch, S. Jain, G. W. Mathern, S. L. Moshé, *et al.*, "Ilae classification of the epilepsies: position paper of the ilae commission for classification and terminology," *Epilepsia*, vol. 58, no. 4, pp. 512–521, 2017.

[3] A. Varsavsky, I. Mareels, and M. Cook, *Epileptic seizures and the EEG: measurement, models, detection and prediction*. CRC Press, 2016.

[4] A. K. Health, "Invasive electroencephalography (eeg) monitoring before epilepsy surgery," 2017. Last accessed 16 June 2020.

[5] L. D. Iasemidis, "Epileptic seizure prediction and control," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 549–558, 2003.

[6] M. Winterhalder, T. Maiwald, H. Voss, R. Aschenbrenner-Scheibe, J. Timmer, and A. Schulze-Bonhage, "The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods," *Epilepsy & Behavior*, vol. 4, no. 3, pp. 318–325, 2003.

[7] B. Schelter, R. G. Andrzejak, and F. Mormann, "Can your prediction algorithm beat a random predictor," *Seizure prediction in epilepsy: from basic mechanisms to clinical applications. New York: Wiley*, pp. 237–48, 2008.

[8] E. B. Assi, D. K. Nguyen, S. Rihana, and M. Sawan, "Towards accurate prediction of epileptic seizures: A review," *Biomedical Signal Processing and Control*, vol. 34, pp. 144–157, 2017.

[9] EPILEPSIAE, "The european epilepsy database." Last accessed 14 July 2020.

[10] I. Kiral-Kornek, S. Roy, E. Nurse, B. Mashford, P. Karoly, T. Carroll, D. Payne, S. Saha, S. Baldassano, T. O'Brien, *et al.*, "Epileptic seizure prediction using big data and deep learning: toward a mobile system," *EBioMedicine*, vol. 27, pp. 103–111, 2018.

[11] C. Teixeira, B. Direito, H. Feldwisch-Drentrup, M. Valderrama, R. Costa, C. Alvarado-Rojas, S. Nikolopoulos, M. Le Van Quyen, J. Timmer, B. Schelter, *et al.*, "Epilab: A software package for studies on the prediction of epileptic seizures," *Journal of Neuroscience Methods*, vol. 200, no. 2, pp. 257–271, 2011.

[12] E. Foundation, "About epilepsy: The basics," 2020. Last accessed 25 April 2020.

[13] L. Kuhlmann, K. Lehnertz, M. P. Richardson, B. Schelter, and H. P. Zaveri, "Seizure prediction—ready for a new era," *Nature Reviews Neurology*, vol. 14, no. 10, pp. 618–630, 2018.

[14] D. R. Freestone, P. J. Karoly, and M. J. Cook, "A forward-looking review of seizure prediction," *Current opinion in neurology*, vol. 30, no. 2, pp. 167–173, 2017.

[15] F. Mormann, R. G. Andrzejak, C. E. Elger, and K. Lehnertz, "Seizure prediction: the long and winding road," *Brain*, vol. 130, no. 2, pp. 314–333, 2007.

[16] M. J. Cook, T. J. O'Brien, S. F. Berkovic, M. Murphy, A. Morokoff, G. Fabinyi, W. D'Souza, R. Yerra, J. Archer, L. Litewka, *et al.*, "Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study," *The Lancet Neurology*, vol. 12, no. 6, pp. 563–571, 2013.

[17] B. Goodman and S. Flaxman, "European union regulations on algorithmic decision-making and a "right to explanation"," *AI magazine*, vol. 38, no. 3, pp. 50–57, 2017.

[18] I. Lage, E. Chen, J. He, M. Narayanan, B. Kim, S. Gershman, and F. Doshi-Velez, "An evaluation of the human-interpretability of explanation," *arXiv preprint arXiv:1902.00006*, 2019.

[19] A. Chamseddine and M. Sawan, "Deep learning based method for output regularization of the seizure prediction classifier," in *2018 IEEE Life Sciences Conference (LSC)*, pp. 118–121, IEEE, 2018.

[20] R. S. Fisher, W. V. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and

J. Engel Jr, "Epileptic seizures and epilepsy: definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe)," *Epilepsia*, vol. 46, no. 4, pp. 470–472, 2005.

[21] R. S. Fisher, C. Acevedo, A. Arzimanoglou, A. Bogacz, J. H. Cross, C. E. Elger, J. Engel Jr, L. Forsgren, J. A. French, M. Glynn, *et al.*, "Ilae official report: a practical clinical definition of epilepsy," *Epilepsia*, vol. 55, no. 4, pp. 475–482, 2014.

[22] E. Foundation, "Temporal lobe epilepsy (tle)," 2019. Last accessed 08 June 2020.

[23] B. Direito, C. A. Teixeira, F. Sales, M. Castelo-Branco, and A. Dourado, "A realistic seizure prediction study based on multiclass svm," *International journal of neural systems*, vol. 27, no. 03, p. 1750006, 2017.

[24] E. Foundation, "Types of epilepsy syndromes," 2008. Last accessed 05 November 2019.

[25] J. Engel Jr, "Ilae classification of epilepsy syndromes," *Epilepsy research*, vol. 70, pp. 5–10, 2006.

[26] J. W. C. Medithe and U. R. Nelakuditi, "Study of normal and abnormal eeg," in *2016 3rd International conference on advanced computing and communication systems (ICACCS)*, vol. 1, pp. 1–4, IEEE, 2016.

[27] I. Osorio, H. P. Zaveri, M. G. Frei, and S. Arthurs, *Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics*. CRC press, 2016.

[28] T. Kirschstein and R. Köhling, "What is the source of the eeg?," *Clinical EEG and neuroscience*, vol. 40, no. 3, pp. 146–149, 2009.

[29] S. Shorvon, R. Guerrini, M. Cook, and S. Lhatoo, *Oxford textbook of epilepsy and epileptic seizures*. OUP Oxford, 2012.

[30] D. Wang, D. Ren, K. Li, Y. Feng, D. Ma, X. Yan, and G. Wang, "Epileptic seizure detection in long-term eeg recordings by using wavelet-based directed transfer function," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 11, pp. 2591–2599, 2018.

[31] U. Herwig, P. Satrapi, and C. Schönfeldt-Lecuona, "Using the international 10-20 eeg system for positioning of transcranial magnetic stimulation," *Brain topography*, vol. 16, no. 2, pp. 95–99, 2003.

[32] M. H. Libenson, *Practical Approach to Electroencephalography E-Book*. Elsevier Health Sciences, 2012.

[33] J. Wellmer, F. von der Groeben, U. Klarmann, C. Weber, C. E. Elger, H. Urbach, H. Clusmann, and M. von Lehe, "Risks and benefits of invasive epilepsy surgery workup with implanted subdural and depth electrodes," *Epilepsia*, vol. 53, no. 8, pp. 1322–1332, 2012.

[34] A. Valentín, N. Hernando-Quintana, J. Moles-Herbera, D. Jimenez-Jimenez, S. Mourente, I. Malik, R. Selway, and G. Alarcón, "Depth versus subdural temporal electrodes revisited: Impact on surgical outcome after resective surgery for epilepsy," *Clinical Neurophysiology*, vol. 128, no. 3, pp. 418–423, 2017.

[35] E. Foundation, "Video eeg monitoring with invasive electrodes," 2018. Last accessed 16 June 2020.

[36] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated eeg analysis of epilepsy: a review," *Knowledge-Based Systems*, vol. 45, pp. 147–165, 2013.

[37] N. Moghim and D. W. Corne, "Predicting epileptic seizures in advance," *PloS one*, vol. 9, no. 6, 2014.

[38] B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, A. Schulze-Bonhage, and J. Timmer, "Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 16, no. 1, p. 013108, 2006.

[39] R. G. Andrzejak, F. Mormann, T. Kreuz, C. Rieke, A. Kraskov, C. E. Elger, and K. Lehnertz, "Testing the null hypothesis of the nonexistence of a preseizure state," *Physical Review E*, vol. 67, no. 1, p. 010901, 2003.

[40] T. Kreuz, R. G. Andrzejak, F. Mormann, A. Kraskov, H. Stögbauer, C. E. Elger, K. Lehnertz, and P. Grassberger, "Measure profile surrogates: a method to validate the performance of epileptic seizure prediction algorithms," *Physical Review E*, vol. 69, no. 6, p. 061915, 2004.

[41] C. Alvarado-Rojas, M. Valderrama, A. Fouad-Ahmed, H. Feldwisch-Drentrup, M. Ihle, C. Teixeira, F. Sales, A. Schulze-Bonhage, C. Adam, A. Dourado, *et al.*, "Slow modulations of high-frequency activity (40–140 hz) discriminate preictal changes in human focal epilepsy," *Scientific reports*, vol. 4, p. 4545, 2014.

[42] T. R. Hoens, R. Polikar, and N. V. Chawla, "Learning from streaming data with

concept drift and imbalance: an overview," *Progress in Artificial Intelligence*, vol. 1, no. 1, pp. 89–101, 2012.

[43] A. Tsymbal, M. Pechenizkiy, P. Cunningham, and S. Puuronen, "Dynamic integration of classifiers for handling concept drift," *Information fusion*, vol. 9, no. 1, pp. 56–68, 2008.

[44] J. Klatt, H. Feldwisch-Drentrup, M. Ihle, V. Navarro, M. Neufang, C. Teixeira, C. Adam, M. Valderrama, C. Alvarado-Rojas, A. Witon, *et al.*, "The epilepsiae database: An extensive electroencephalography database of epilepsy patients," *Epilepsia*, vol. 53, no. 9, pp. 1669–1676, 2012.

[45] S. Marsland, *Machine learning: an algorithmic perspective*. CRC press, 2015.

[46] H. Khan, L. Marcuse, M. Fields, K. Swann, and B. Yener, "Focal onset seizure prediction using convolutional networks," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 2109–2118, 2017.

[47] M. Sun, F. Wang, T. Min, T. Zang, and Y. Wang, "Prediction for high risk clinical symptoms of epilepsy based on deep learning algorithm," *IEEE Access*, vol. 6, pp. 77596–77605, 2018.

[48] P. Agarwal, H.-C. Wang, and K. Srinivasan, "Epileptic seizure prediction over eeg data using hybrid cnn-svm model with edge computing services," in *MATEC Web of Conferences*, vol. 210, p. 03016, EDP Sciences, 2018.

[49] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[50] G. Aggarwal and T. K. Gandhi, "Prediction of epileptic seizures based on mean phase coherence," *bioRxiv*, p. 212563, 2017.

[51] U. R. Acharya, F. Molinari, S. V. Sree, S. Chattopadhyay, K.-H. Ng, and J. S. Suri, "Automated diagnosis of epileptic eeg using entropies," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 401–408, 2012.

[52] E. B. Assi, M. Sawan, D. Nguyen, and S. Rihana, "A hybrid mrmr-genetic based selection method for the prediction of epileptic seizures," in *2015 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pp. 1–4, IEEE, 2015.

[53] A. B. Geva and D. H. Kerem, "Forecasting generalized epileptic seizures from the eeg signal by wavelet analysis and dynamic unsupervised fuzzy clustering," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 10, pp. 1205–1216, 1998.

[54] M. Bandarabadi, C. A. Teixeira, J. Rasekhi, and A. Dourado, "Epileptic seizure prediction using relative spectral power features," *Clinical Neurophysiology*, vol. 126, no. 2, pp. 237–248, 2015.

[55] M. Bandarabadi, J. Rasekhi, C. A. Teixeira, M. R. Karami, and A. Dourado, "On the proper selection of preictal period for seizure prediction," *Epilepsy & Behavior*, vol. 46, pp. 158–166, 2015.

[56] J. Rasekhi, M. R. K. Mollaei, M. Bandarabadi, C. A. Teixeira, and A. Dourado, "Epileptic seizure prediction based on ratio and differential linear univariate features," *Journal of medical signals and sensors*, vol. 5, no. 1, p. 1, 2015.

[57] I. Mporas, V. Tsirka, E. I. Zacharaki, M. Koutroumanidis, M. Richardson, and V. Megalooikonomou, "Seizure detection using eeg and ecg signals for computer-based monitoring, analysis and management of epileptic patients," *Expert systems with applications*, vol. 42, no. 6, pp. 3227–3233, 2015.

[58] C. A. Teixeira, B. Direito, M. Bandarabadi, M. Le Van Quyen, M. Valderrama, B. Schelter, A. Schulze-Bonhage, V. Navarro, F. Sales, and A. Dourado, "Epileptic seizure predictors based on computational intelligence techniques: A comparative study with 278 patients," *Computer methods and programs in biomedicine*, vol. 114, no. 3, pp. 324–336, 2014.

[59] A. F. Rabbi, L. Azinfar, and R. Fazel-Rezai, "Seizure prediction using adaptive neuro-fuzzy inference system," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2100–2103, IEEE, 2013.

[60] J. Rasekhi, M. R. K. Mollaei, M. Bandarabadi, C. A. Teixeira, and A. Dourado, "Preprocessing effects of 22 linear univariate features on the performance of seizure prediction methods," *Journal of neuroscience methods*, vol. 217, no. 1-2, pp. 9–16, 2013.

[61] M. Bandarabadi, C. A. Teixeira, B. Direito, and A. Dourado, "Epileptic seizure prediction based on a bivariate spectral power methodology," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5943–5946, IEEE, 2012.

[62] C. Teixeira, B. Direito, M. Bandarabadi, and A. Dourado, "Output regularization of svm seizure predictors: Kalman filter versus the "firing power" method," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6530–6533, IEEE, 2012.

[63] B. Direito, C. Teixeira, B. Ribeiro, M. Castelo-Branco, F. Sales, and A. Dourado, "Modeling epileptic brain states using eeg spectral analysis and topographic mapping," *Journal of neuroscience methods*, vol. 210, no. 2, pp. 220–229, 2012.

[64] M. Valderrama, C. Alvarado, S. Nikolopoulos, J. Martinerie, C. Adam, V. Navarro, and M. Le Van Quyen, "Identifying an increased risk of epileptic seizures using a multi-feature eeg–ecg classification," *Biomedical Signal Processing and Control*, vol. 7, no. 3, pp. 237–244, 2012.

[65] B. Direito, F. Ventura, C. Teixeira, and A. Dourado, "Optimized feature subsets for epileptic seizure prediction studies," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1636–1639, IEEE, 2011.

[66] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of eeg using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[67] L. Chisci, A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio, and F. Fuggetta, "Real-time epileptic seizure prediction using ar models and support vector machines," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 5, pp. 1124–1132, 2010.

[68] P. Mirowski, D. Madhavan, Y. LeCun, and R. Kuzniecky, "Classification of patterns of eeg synchronization for seizure prediction," *Clinical neurophysiology*, vol. 120, no. 11, pp. 1927–1940, 2009.

[69] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *Clinical neurophysiology*, vol. 116, no. 3, pp. 569–587, 2005.

[70] M. D'Alessandro, R. Esteller, G. Vachtsevanos, A. Hinson, J. Echauz, and B. Litt, "Epileptic seizure prediction using hybrid feature selection over multiple intracranial eeg electrode contacts: a report of four patients," *IEEE transactions on biomedical engineering*, vol. 50, no. 5, pp. 603–615, 2003.

[71] F. Mormann, T. Kreuz, R. G. Andrzejak, P. David, K. Lehnertz, and C. E. Elger, "Epileptic seizures are preceded by a decrease in synchronization," *Epilepsy research*, vol. 53, no. 3, pp. 173–185, 2003.

[72] F. Mormann, R. G. Andrzejak, T. Kreuz, C. Rieke, P. David, C. E. Elger, and K. Lehnertz, "Automated detection of a preseizure state based on a decrease in

synchronization in intracranial electroencephalogram recordings from epilepsy patients," *Physical Review E*, vol. 67, no. 2, p. 021912, 2003.

[73] M. Le Van Quyen, J. Martinerie, M. Baulac, and F. Varela, "Anticipating epileptic seizures in real time by a non-linear analysis of similarity between eeg recordings," *Neuroreport*, vol. 10, no. 10, pp. 2149–2155, 1999.

[74] 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*, pp. 1864–1867, IEEE, 2009.

[75] L. D. Iasemidis, J. C. Sackellares, H. P. Zaveri, and W. J. Williams, "Phase space topography and the lyapunov exponent of electrocorticograms in partial seizures," *Brain topography*, vol. 2, no. 3, pp. 187–201, 1990.

[76] L. D. Iasemidis, P. Pardalos, J. C. Sackellares, and D.-S. Shiau, "Quadratic binary programming and dynamical system approach to determine the predictability of epileptic seizures," *Journal of combinatorial optimization*, vol. 5, no. 1, pp. 9–26, 2001.

[77] J. Arnhold, P. Grassberger, K. Lehnertz, and C. E. Elger, "A robust method for detecting interdependences: application to intracranially recorded eeg," *Physica D: Nonlinear Phenomena*, vol. 134, no. 4, pp. 419–430, 1999.

[78] R. Gravina, P. Alinia, H. Ghasemzadeh, and G. Fortino, "Multi-sensor fusion in body sensor networks: State-of-the-art and research challenges," *Information Fusion*, vol. 35, pp. 68–80, 2017.

[79] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, 2019.

[80] N. Sánchez-Marono, A. Alonso-Betanzos, and M. Tombilla-Sanromán, "Filter methods for feature selection–a comparative study," in *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 178–187, Springer, 2007.

[81] K. Hoyos-Osorio, J. Castañeda-Gonzaiez, and G. Daza-Santacoloma, "Automatic epileptic seizure prediction based on scalp eeg and ecg signals," in *2016 XXI Symposium on signal processing, images and artificial vision (STSIVA)*, pp. 1–7, IEEE, 2016.

[82] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel,

M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[83] A. Schulze-Bonhage, F. Sales, K. Wagner, R. Teotonio, A. Carius, A. Schelle, and M. Ihle, "Views of patients with epilepsy on seizure prediction devices," *Epilepsy & behavior*, vol. 18, no. 4, pp. 388–396, 2010.

[84] B. Litt, R. Esteller, J. Echauz, M. D'Alessandro, R. Shor, T. Henry, P. Pennell, C. Epstein, R. Bakay, M. Dichter, *et al.*, "Epileptic seizures may begin hours in advance of clinical onset: a report of five patients," *Neuron*, vol. 30, no. 1, pp. 51–64, 2001.

[85] B. Hjorth, "Eeg analysis based on time domain properties," *Electroencephalography and clinical neurophysiology*, vol. 29, no. 3, pp. 306–310, 1970.

[86] D. Schwender, M. Daunderer, S. Mulzer, S. Klasing, U. Finsterer, and K. Peter, "Spectral edge frequency of the electroencephalogram to monitor "depth" of anaesthesia with isoflurane or propofol," *British journal of anaesthesia*, vol. 77, no. 2, pp. 179–184, 1996.

[87] M. Akin, "Comparison of wavelet transform and fft methods in the analysis of eeg signals," *Journal of medical systems*, vol. 26, no. 3, pp. 241–247, 2002.

[88] T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, and F. E. A. El-Samie, "Eeg seizure detection and prediction algorithms: a survey," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, p. 183, 2014.

[89] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical review letters*, vol. 50, no. 5, p. 346, 1983.

[90] Y.-C. Lai, I. Osorio, M. A. F. Harrison, and M. G. Frei, "Correlation-dimension and autocorrelation fluctuations in epileptic seizure dynamics," *Physical review E*, vol. 65, no. 3, p. 031921, 2002.

Appendices

A

Features description

A.1 Univariate Linear features

Statistical Moments

The first four statistical moments are the most commonly used in seizure prediction. The following Table (A.1) summarizes those moments. In the expressions, N is the number of samples and x the vector with the values for analysis [15].

Table A.1: Statistical moments

Order	Formula	Definition
First (Mean)	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$ (A.1)	Measures the central tendency of the amplitude of the samples
Second (Variance)	$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^2$ (A.2)	Measures the dispersion of the amplitude of the samples around its mean
Third (Skewness)	$\chi = \frac{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^3}{\sigma^3}$ (A.3)	Measures the degree of asymmetries of the amplitude distribution
Fourth (Kurtosis)	$\kappa = \frac{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^4}{\sigma^4} - 3$ (A.4)	Measures the relative flatness or peakedness of the amplitude distribution

Accumulated energy

To calculate the accumulated energy, instantaneous energy is obtained first by squaring the Electroencephalogram (EEG) values. Then, an averaging window with length M is applied on the resulting instantaneous energy sequence (A.5).

$$E_{avg} = \frac{1}{M}\sum_{i=1}^{M} x_i^2 \tag{A.5}$$

Finally, the accumulated energy (A.6) is given by the average of the E_{avg} se-

quence using other window with length N:

$$AE_m = \frac{1}{N}\sum_{j=1}^{N} E_{avg} + AE_{m-1} \tag{A.6}$$

where AE_m is m^{th} value of the accumulated energy and AE_{m-1} is the previous value [56, 84].

Hjörth parameters

Hjörth [85] formulated these time domain parameters taking into account the standard deviations. Let us consider x the input signal and x' and x'' the first and second derivatives of x. Table A.2 presents the three Hjorth parameters in mathematical terms along with their description.

Table A.2: Hjörth parameters

Hjorth parameters	**Formula**	**Definition**
Activity	$\mathrm{H}_a = \sigma^2(x)$(A.7)	Measure of the squared standard deviation of the amplitude.
Mobility	$\mathrm{H}_m = \sqrt{\frac{\sigma^2(x')}{\sigma^2(x)}}$(A.8)	Measure of the standard deviation of the slope with reference to the standard deviation of the amplitude.
Complexity	$\mathrm{H}_c = \sqrt{\frac{\sigma^2(x'')\sigma^2(x)}{(\sigma^2(x'))^2}}$(A.9)	Measure of the standard slopes generated during the average time required for generation of one standard amplitude as given by the mobility.

Decorrelation time

Decorrelation time is defined as the first zero crossing of a given EEG signal's autocorrelation function. The autocorrelation function of a time series x is defined as A.10 where $\tau = 0,..., N-1$ [8, 15].

$$A(\tau) = \frac{1}{(N-1)\sigma^2}\sum_{i=1}^{N-\tau} x_i x_{i-\tau} \tag{A.10}$$

$A(\tau)$ ranges between -1 and 1, if $A(\tau) = 0$ the values being compared (x_i and $x_{i-\tau}$) are uncorrelated and when those values are similar, $A(\tau) = 1$. The decorrelation time can be used to estimate the strength of linear correlations and is

defined as the following equation (A.11) [15]:

$$\tau_0 = min\{\tau | A(\tau) = 0\} \tag{A.11}$$

For lower values of decorrelation time, the signal is less correlated [11].

Linear modeling

Autoregressive (AR) models assumes that the stationarity of the signal. In this modelling approach, the model's output value Y_i is determined by employing a weighted sum of a certain number of p previous output values Y_{i-l} plus a constant c and noise ε_l [15, 23, 36, 60].

An AR process is defined by equation A.12 where a_l are the parameters of the model.

$$Y_i = c + \sum_{l=1}^{p} a_l Y_{i-l} + \varepsilon_l \tag{A.12}$$

Several studies mention measures computed from this type of models as part of their features set. Direito et al.[23, 65], Rasekhi et al.[56, 60], Teixeira et al.[58] and Valderrama et al.[64] calculated the mean square prediction error derived from an autoregressive model. The prediction error is the difference between the value predicted by an AR model and the actual value [23]. As seizure onset approaches, the decrease in the prediction error is assumed, since the EEG signals become better predicted by the AR model. After seizure onset, in the ictal period, the error tends to increase [8, 58]. Other studies [57, 62] used the model coefficients. Chisci et al.[67] developed a study based only on features of autoregressive models to predict seizures.

Relative Spectral band power

Power Spectral Density (PSD) or power spectrum represents the distribution of strength of signal components across different frequencies. PSD can be calculated by applying firstly the Fast Fourier Transform (FFT) to the EEG time series and then to average the squared coefficients of the frequency range of interest. Since there is more power at EEG low frequencies than at high frequencies, is performed a normalization of the spectral power. Thus, the relative spectral power (A.13) is defined as the power of a given frequency band (p_i) divided by the total power of

the EEG signal (P) [8, 15, 56].

$$RSP = \frac{p_i}{P} \tag{A.13}$$

More specifically, for example for the gamma band, the relative power is given by,

$$\gamma = \frac{\sum_{f=30Hz}^{100Hz} p_f}{P} \tag{A.14}$$

where p_f is the power spectrum.

Spectral edge frequency and power

Spectral Edge Frequency (SEF) is generally defined as the minimum frequency up to which a certain percent of the total power is contained in the signal. Spectral Edge Power (SEP) is the power corresponding to the percent defined of the total power. For example, SEF 50, SEF 90, SEF 95 are the minimum frequencies up to which 50%, 90% and 95%, respectively, of the total power is located [15, 86].

In the case of the EEG signal, most of the power is comprised between 0Hz and 40Hz. Thus, in seizure prediction studies SEF 50 is commonly used. Therefore, SEF 50 is the frequency below to which 50% of the total power of the signal up to 40 Hz is located and SEP 50 is the corresponding half power of the signal up to 40 Hz [15, 58].

Wavelet Coefficients

Wavelet method can be used, as an alternative to the FFT, for both spectral and temporal analysis of the brain signals. This type of analysis has a significant role in the processing of non-stationary signals [60, 87].

Wavelet transform enables the decomposition of a certain signal in different levels into basis functions called wavelets. Those wavelets are obtained by scaled and translated variations from a mother wavelet. The higher frequencies of the signal are associated with the first decomposition levels and the lower frequencies with the last levels [58, 60, 87].

Wavelets have been widely used for EEG signal analysis. The main purpose of using wavelet transform is to extract discriminating features from different sub-bands [88]. Rasekhi et al. [56, 60], Direito et al.[23, 65], D'Alessandro et al.[70] studies computed the energy of the signals originated from the decomposition. The most used family of mother wavelets is Daubechies due to their good localization

properties in time and frequency domains. They are used in studies such as Mporas et al.[57], Direito et al. [23, 65] and Rasekhi et al.[56, 60].

A.2 Univariate Non-linear features

Correlation sum and correlation dimension

The correlation sum is a measure used to estimate the correlation integral, originally proposed by Grassberger and Procaccia [89]. It determines the number of pairs of vectors in state space that are closer than a certain hypothetical hypersphere of radius ε. The correlation integral can be approximated by the following correlation sum (A.15):

$$C(\varepsilon) = \frac{2}{N(N-1)} \sum_{j=1}^{N} \sum_{i=j+1}^{N} \Theta(\varepsilon - |x_i - x_j|) \tag{A.15}$$

where Θ is the Heaviside function and $|x_i - x_j|$ is the norm between two points. The points are obtained from a reconstructed vector time series ($x(t)$) of length N using the time-delay embedding technique [15, 89, 90].

The correlation dimension estimate the number of active degrees of freedom of random points within a state space and is defined from local slope of the correlation sum (equation A.16) [15, 37].

$$D_2 = \lim_{N \to \infty} \lim_{\varepsilon \to} \frac{ln(C(\varepsilon))}{ln(\varepsilon)} \tag{A.16}$$

Lyapunov exponent

Lyapunov exponents quantify the exponential divergence of nearby trajectories in state-space, which is a space where all the possible system states are represented. The exponential divergence, one basic indicator of deterministic chaos, depends on the initial state conditions. [8, 37, 74].

One may compute the value of the largest Lyapunov exponent (L_{max}) in order to determine the predictability of a dynamical system [74]. L_{max} can be estimated from (A.17)

$$d_j(i) = C_j e^{L_{max} i \Delta t} \tag{A.17}$$

where $d_j(i)$ is the average difference between two trajectory segments at time

t_i, C_j is the initial distance of a reference vector $\vec{z}_j$ in state-space and its closet neighbor, with $j = 1,...,M$ [15].

Lyapunov exponents have been used in some studies as Moghim et al. [37], Mormann et al.[69] and also in the Mirowski et al.[68] to compute the dynamical entrainment. The first reports [75] revealed a decrease in the largest Lyapunov exponent several minutes before seizure onset, however, a further investigation [69] observed the opposite indicating an increase of L_{max} 30 min prior seizure.

Dynamical similarity index

Dynamic Similarity Index (DSI) was proposed by Le Van Quyen et al.[73] to measure the dynamical similarity between two segments of the EEG signal. One of the segments is a reference segment, typically of a few minutes, selected during a distant interval in time from a seizure, i.e. an inter-ictal segment. The other one is a segment obtained from a moving test window [8, 15, 73].

Let us consider r the reference segment and t the test window. Mathematically, DSI is defined as (A.18)

$$\gamma_t = \frac{C_{rt}}{\sqrt{C_{rr}C_{tt}}} \tag{A.18}$$

where C_{rt} is the cross-correlation sum and C_{rr} and C_{tt} are correlation sums of the reference and test window, respectively [15, 73].

γ_t gives a sensitive measure of closeness and ranges from 0 to 1. If the reference segment and the test segment share identical dynamics, γ_t will be near 1. Otherwise it will be less than 1 [73].

In their investigations, Le Van Quyen et al.[73] achieved good seizure forecasting results. However, other researchers, such as Winterhalder et al.[6] have not achieved such optimistic results for the DSI [8, 74].

Entropy measures

Entropy is a measure of the unpredictability and uncertainty of a signal [51]. There are several measures related to entropy, such as approximate entropy, sample entropy, permutation entropy and spectral entropy.

Acharya et al.[51] extracted four entropy measures from EEG signals. Mporas et al.[57] presented a methodology based on a vector of 55 features including also entropy features like spectral entropy and logarithm energy entropy.

A.3 Multivariate Linear features

Maximum linear cross-correlation

Maximum linear cross-correlation is a measure for the degree of lag synchronization, which is the condition in which two systems are identical but shifted by a time lag t. Therefore, this feature quantifies the similarity between two time series [8, 72].

The maximum of normalized cross-correlation keeps the measure independent of the variance of the signals. It is confined to the interval [0,1]. Values close to 1 indicate high similarity between the two signals in analysis although with a possible time lag t, while unsynchronized signals will return values close to 0 [8, 15, 72].

A.4 Multivariate Non-linear features

Mean phase coherence

Mean Phase Coherence (MPC) quantifies the degree of phase synchronization for two time series. For its calculation, firstly it is necessary to extract the phases of the series. Let $\phi_a(t)$ and $\phi_b(t)$ be the phases of the series a and b with length N at instant t [40, 50]. Mathematically, MPC is given by equation A.19.

$$R = \left| \frac{1}{N} \sum_{j=1}^{N} e^{j[\phi_a(t_j) - \phi_b(t_j)]} \right| \tag{A.19}$$

The MPC is confined between 0 and 1, where values close to 1 indicate a high degree of synchronization [15].

Mormann et al. [71] and Aggarwal et al.[50] reported the decrease in mean phase coherence values before seizure onset while the pre-ictal period. Aggarwal et al. [50] also evaluated a patient without epilepsy and observed no changes in MPC value.

Dynamical entrainment

Dynamical entrainment is based on chaos theory since it is a multichannel version of the Lyapunov exponent. This measure was proposed by Iasemidis et al.[76] and it calculates the difference between the largest Lyapunov exponents of two series [8, 68]. Mirowski et al.[68] also reported the use of this feature.

Non-linear interdependence

Non-linear interdependence has been used is some studies [59, 68, 69] as a measure of generalized synchronization. Arnhold et al.[77] proposed two measures for non-linear interdependence, termed S and H. These bivariate features determine the Euclidean distance between trajectories previously reconstructed from two EEG signals in the state-space [15].

B

Supplementary Results

Table B.1: Training parameters and performance of Approach A with events duration of 20 minutes.

	Approach A with 20-min events											
	Event 1				**Event 2**				**Event 3**			
Patient	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}
402	2^{0}	20	0.67	0.58	2^{-4}	10	0.67	0.54	2^{-20}	30	0.67	0.54
8902	2^{-8}	20	0.65	0.92	2^{-8}	10	0.66	0.91	2^{-8}	30	0.67	0.92
11002	2^{-4}	10	0.67	0.75	2^{0}	20	0.67	0.74	2^{0}	10	0.67	0.75
16202	2^{0}	20	0.97	0.52	2^{0}	20	0.99	0.53	2^{-4}	10	1.00	0.50
23902	2^{0}	10	0.67	0.82	2^{4}	10	0.67	0.81	2^{-12}	30	0.66	0.78
30802	2^{-12}	30	0.67	0.74	2^{-12}	30	0.67	0.74	2^{-12}	30	0.67	0.75
32702	2^{8}	30	0.66	0.57	2^{8}	10	0.37	0.72	2^{4}	10	0.33	0.71
46702	2^{-8}	10	0.67	0.84	2^{0}	20	1.00	0.86	2^{-4}	20	0.67	0.88
50802	2^{4}	10	1.00	0.53	2^{0}	20	0.94	0.59	2^{0}	20	0.94	0.60
53402	2^{-4}	20	0.57	0.71	2^{4}	10	0.34	0.87	2^{8}	20	0.67	0.59
55202	2^{8}	30	0.66	0.75	2^{4}	10	0.67	0.61	2^{8}	20	0.65	0.62
56402	2^{-20}	20	0.67	0.60	2^{-8}	10	0.67	0.62	2^{-8}	10	0.67	0.62
58602	2^{0}	20	0.67	0.61	2^{4}	30	0.63	0.81	2^{0}	10	0.67	0.63
59102	2^{-20}	20	0.99	0.68	2^{-20}	30	1.00	0.71	2^{-20}	30	0.99	0.74
60002	2^{8}	30	0.33	0.90	2^{8}	20	0.60	0.78	2^{0}	30	0.68	0.51
64702	2^{-8}	20	1.00	0.62	2^{-8}	10	1.00	0.64	2^{-8}	10	1.00	0.59
75202	2^{-8}	10	0.67	0.85	2^{-8}	20	0.72	0.86	2^{8}	30	0.67	0.89
80702	2^{4}	30	0.33	0.84	2^{-8}	30	0.54	0.76	2^{-8}	20	0.42	0.77
85202	2^{0}	10	1.00	0.63	2^{4}	20	0.63	0.88	2^{-12}	10	0.67	0.81
93402	2^{-16}	10	0.67	0.69	2^{-4}	10	0.67	0.67	2^{8}	10	0.67	0.71
93902	2^{0}	30	0.66	0.74	2^{0}	30	0.67	0.69	2^{-4}	20	0.33	0.94
94402	2^{8}	30	0.37	0.67	2^{4}	10	0.33	0.75	2^{4}	30	0.67	0.51
95202	2^{0}	30	0.37	0.66	2^{-20}	20	0.67	0.55	2^{-4}	10	0.66	0.51
96002	2^{-20}	10	0.67	0.70	2^{-16}	20	0.93	0.54	2^{-12}	20	1.00	0.61
98102	2^{4}	30	0.67	0.61	2^{4}	10	0.67	0.69	2^{8}	20	0.67	0.54
98202	2^{4}	10	0.67	0.83	2^{-8}	10	0.67	0.79	2^{-4}	10	0.67	0.77
101702	2^{-20}	20	0.67	0.68	2^{-20}	20	0.67	0.69	2^{-20}	20	0.67	0.69
102202	2^{0}	10	0.67	0.86	2^{0}	10	0.67	0.75	2^{4}	10	0.67	0.73
104602	2^{-8}	20	0.34	0.81	2^{8}	30	0.33	0.97	2^{4}	10	0.64	0.86
109502	2^{0}	20	0.67	0.49	2^{8}	20	0.97	0.55	2^{-4}	20	0.33	0.89
110602	2^{-20}	20	0.67	0.74	2^{-16}	20	0.67	0.75	2^{-20}	20	0.67	0.76
112802	2^{8}	10	0.91	0.54	2^{-20}	20	0.88	0.59	2^{-20}	20	0.90	0.60
113902	2^{-4}	10	0.67	0.79	2^{-4}	10	0.61	0.78	2^{-4}	10	0.66	0.79
114702	2^{-4}	20	0.33	0.70	2^{-4}	30	0.34	0.67	2^{4}	10	0.33	0.67
114902	2^{0}	10	0.67	0.59	2^{4}	30	0.66	0.50	2^{8}	20	0.53	0.83
123902	2^{-16}	10	0.67	0.74	2^{-20}	30	0.67	0.74	2^{8}	10	0.67	0.75
Avg			0.66	0.70			0.68	0.71			0.67	0.70

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table B.2: Training parameters and performance of Approach B with events duration of 20 minutes.

	Approach B with 20-min events											
	Event 1				Event 2				Event 3			
Patient	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}
402	2^{-4}	10	0.67	0.54	2^{-4}	10	0.67	0.54	2^{-20}	30	0.67	0.54
8902	2^{-12}	30	0.66	0.85	2^{-12}	30	0.66	0.84	2^{-8}	30	0.67	0.92
11002	2^{-20}	20	0.67	0.77	2^{-20}	20	0.67	0.75	2^{0}	10	0.67	0.75
16202	2^{-4}	30	1.00	0.56	2^{-4}	20	1.00	0.55	2^{-4}	10	1.00	0.50
23902	2^{-12}	10	0.67	0.79	2^{4}	10	0.67	0.83	2^{-12}	30	0.66	0.78
30802	2^{0}	10	0.65	0.77	2^{-20}	20	0.67	0.74	2^{-12}	30	0.67	0.74
32702	2^{8}	10	0.65	0.60	2^{4}	10	0.47	0.68	2^{4}	10	0.33	0.71
46702	2^{8}	30	0.94	0.87	2^{4}	30	1.00	0.84	2^{-4}	20	0.67	0.88
50802	2^{-4}	10	0.96	0.60	2^{0}	10	1.00	0.50	2^{0}	20	0.93	0.60
53402	2^{0}	20	0.67	0.59	2^{0}	20	0.67	0.58	2^{8}	20	0.67	0.59
55202	2^{-4}	20	0.68	0.63	2^{4}	10	0.67	0.64	2^{8}	20	0.67	0.62
56402	2^{-20}	20	0.67	0.63	2^{-20}	20	0.67	0.62	2^{-8}	10	0.67	0.62
58602	2^{8}	30	0.67	0.61	2^{4}	20	0.67	0.60	2^{0}	10	0.67	0.63
59102	2^{-20}	20	1.00	0.70	2^{-20}	30	0.99	0.75	2^{-20}	30	0.99	0.74
60002	2^{-16}	10	0.65	0.64	2^{4}	30	0.67	0.53	2^{0}	30	0.65	0.53
64702	2^{-12}	20	1.00	0.63	2^{-8}	10	1.00	0.62	2^{-8}	10	1.00	0.59
75202	2^{-8}	10	0.63	0.87	2^{-8}	20	0.71	0.89	2^{8}	30	0.67	0.89
80702	2^{-12}	20	0.58	0.79	2^{-8}	20	0.47	0.78	2^{-8}	20	0.43	0.77
85202	2^{-20}	10	0.80	0.83	2^{-20}	10	0.82	0.82	2^{-12}	10	0.67	0.81
93402	2^{0}	10	0.67	0.69	2^{0}	10	0.67	0.69	2^{8}	10	0.67	0.71
93902	2^{0}	10	0.67	0.68	2^{-20}	10	0.67	0.87	2^{-4}	20	0.33	0.94
94402	2^{8}	30	0.67	0.57	2^{-16}	10	0.33	0.67	2^{4}	30	0.67	0.50
95202	2^{4}	10	0.51	0.65	2^{0}	10	0.63	0.63	2^{-4}	10	0.66	0.51
96002	2^{8}	30	1.00	0.47	2^{-12}	30	0.91	0.68	2^{-12}	20	1.00	0.61
98102	2^{-12}	30	0.53	0.71	2^{4}	20	0.67	0.54	2^{8}	30	0.67	0.54
98202	2^{8}	10	0.67	0.76	2^{-4}	20	0.67	0.77	2^{-4}	10	0.66	0.77
101702	2^{-20}	20	0.67	0.72	2^{-20}	20	0.67	0.71	2^{-20}	20	0.67	0.69
102202	2^{8}	10	0.67	0.80	2^{0}	10	0.67	0.67	2^{4}	10	0.67	0.73
104602	2^{-12}	20	0.34	0.80	2^{8}	20	0.44	0.93	2^{4}	10	0.63	0.86
109502	2^{0}	10	0.85	0.70	2^{8}	30	0.67	0.73	2^{4}	20	0.31	0.90
110602	2^{-20}	20	0.67	0.76	2^{-16}	20	0.67	0.76	2^{-16}	20	0.66	0.76
112802	2^{-20}	20	0.86	0.62	2^{-20}	20	0.92	0.61	2^{-20}	20	0.90	0.60
113902	2^{-4}	10	0.67	0.74	2^{-4}	10	0.67	0.76	2^{-4}	10	0.67	0.79
114702	2^{-20}	10	0.33	0.67	2^{-4}	30	0.33	0.75	2^{0}	10	0.33	0.67
114902	2^{8}	20	0.83	0.70	2^{4}	30	0.61	0.81	2^{8}	20	0.53	0.81
123902	2^{-16}	10	0.67	0.75	2^{-20}	30	0.67	0.75	2^{-4}	20	0.67	0.74
Avg			0.71	0.70			0.70	0.71			0.67	0.70

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table B.3: Training parameters and performance of Approach C with 20-min pre-ictal.

	Approach C with 20-min pre-ictal			
Patient	C	k	SS_{sample}	SP_{sample}
402	2^{-20}	30	0.67	0.54
8902	2^{-8}	30	0.67	0.92
11002	2^{0}	10	0.67	0.75
16202	2^{-4}	10	1.00	0.50
23902	2^{-12}	30	0.66	0.78
30802	2^{-12}	30	0.67	0.75
32702	2^{4}	10	0.33	0.71
46702	2^{-4}	20	0.67	0.88
50802	2^{0}	30	0.91	0.63
53402	2^{8}	20	0.67	0.59
55202	2^{8}	20	0.65	0.62
56402	2^{-8}	10	0.67	0.62
58602	2^{0}	10	0.67	0.63
59102	2^{-20}	30	0.99	0.74
60002	2^{0}	30	0.68	0.51
64702	2^{-8}	10	1.00	0.59
75202	2^{8}	30	0.67	0.89
80702	2^{-8}	20	0.44	0.77
85202	2^{-12}	10	0.67	0.81
93402	2^{8}	10	0.67	0.71
93902	2^{-4}	20	0.33	0.94
94402	2^{4}	30	0.67	0.50
95202	2^{-4}	10	0.66	0.51
96002	2^{-12}	20	1.00	0.61
98102	2^{8}	20	0.67	0.54
98202	2^{-4}	10	0.66	0.77
101702	2^{-20}	20	0.67	0.69
102202	2^{4}	10	0.67	0.73
104602	2^{4}	10	0.66	0.86
109502	2^{-4}	20	0.33	0.89
110602	2^{-16}	20	0.67	0.76
112802	2^{-20}	20	0.89	0.60
113902	2^{-4}	10	0.67	0.79
114702	2^{0}	10	0.34	0.67
114902	2^{8}	20	0.58	0.81
123902	2^{-8}	30	0.67	0.74
Avg			0.67	0.70

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table B.4: Training parameters and performance of Approach A with events duration of 30 minutes.

	Approach A with 30-min events											
	Event 1				Event 2				Event 3			
Patient	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}
402	2^{-4}	10	0.67	0.57	2^{-4}	10	0.67	0.55	2^{-4}	10	0.67	0.54
8902	2^{0}	20	0.66	0.79	2^{-8}	30	0.58	0.94	2^{-8}	30	0.65	0.94
11002	2^{-20}	20	0.67	0.73	2^{-4}	10	0.67	0.76	2^{0}	10	0.67	0.76
16202	2^{0}	20	1.00	0.49	2^{0}	20	0.97	0.53	2^{-8}	10	1.00	0.56
23902	2^{0}	10	0.66	0.94	2^{-8}	10	0.67	0.82	2^{-12}	30	0.66	0.79
30802	2^{-8}	30	0.66	0.77	2^{-20}	30	0.67	0.74	2^{-12}	30	0.67	0.75
32702	2^{4}	10	0.35	0.73	2^{8}	30	0.66	0.57	2^{0}	10	0.33	0.72
46702	2^{-8}	10	0.67	0.85	2^{0}	20	0.66	0.92	2^{-4}	30	0.66	0.90
50802	2^{4}	10	0.97	0.56	2^{4}	10	1.00	0.54	2^{4}	20	0.61	0.83
53402	2^{4}	30	0.33	0.88	2^{0}	20	0.67	0.60	2^{-4}	20	0.67	0.59
55202	2^{-8}	30	0.62	0.70	2^{8}	20	0.67	0.62	2^{4}	20	0.67	0.62
56402	2^{-20}	20	0.67	0.62	2^{-20}	20	0.67	0.62	2^{-8}	10	0.67	0.66
58602	2^{4}	30	0.67	0.73	2^{4}	30	0.67	0.63	2^{0}	30	0.67	0.61
59102	2^{-12}	10	1.00	0.61	2^{-20}	20	1.00	0.68	2^{0}	30	0.74	0.96
60002	2^{8}	20	0.42	0.69	2^{-8}	10	0.33	0.92	2^{0}	30	0.67	0.51
64702	2^{-8}	10	1.00	0.63	2^{-8}	10	1.00	0.70	2^{-8}	10	1.00	0.61
75202	2^{-8}	30	0.78	0.81	2^{-8}	10	0.67	0.85	2^{-8}	30	0.73	0.88
80702	2^{8}	20	1.00	0.58	2^{-8}	20	0.61	0.75	2^{-8}	20	0.44	0.77
85202	2^{-12}	10	0.67	0.78	2^{-8}	10	0.67	0.79	2^{-20}	10	0.75	0.82
93402	2^{-20}	10	0.67	0.68	2^{-16}	10	0.67	0.68	2^{8}	10	0.67	0.71
93902	2^{0}	10	0.55	0.81	2^{0}	10	0.52	0.83	2^{0}	30	0.67	0.69
94402	2^{4}	20	0.67	0.66	2^{8}	20	0.33	0.80	2^{8}	30	0.67	0.41
95202	2^{-20}	20	0.67	0.55	2^{-12}	30	0.66	0.48	2^{-20}	20	0.67	0.53
96002	2^{-16}	20	1.00	0.56	2^{-20}	20	0.95	0.55	2^{-12}	30	0.96	0.64
98102	2^{4}	30	0.67	0.58	2^{-4}	10	0.55	0.76	2^{8}	20	0.67	0.54
98202	2^{0}	10	0.33	0.87	2^{4}	10	0.59	0.87	2^{-4}	10	0.67	0.77
101702	2^{-20}	20	0.67	0.69	2^{-20}	20	0.67	0.69	2^{-20}	20	0.67	0.70
102202	2^{-4}	10	0.59	0.87	2^{0}	10	0.67	0.88	2^{4}	30	0.67	0.85
104602	2^{8}	10	0.67	0.55	2^{-8}	20	0.33	0.82	2^{0}	30	0.34	0.94
109502	2^{-4}	10	0.67	0.55	2^{4}	10	0.69	0.48	2^{0}	20	0.32	0.88
110602	2^{-20}	10	0.67	0.76	2^{-20}	20	0.67	0.74	2^{-20}	20	0.67	0.76
112802	2^{0}	20	0.33	0.97	2^{8}	10	0.97	0.54	2^{-20}	20	0.97	0.61
113902	2^{-4}	10	0.67	0.77	2^{-4}	10	0.62	0.76	2^{-4}	10	0.65	0.78
114702	2^{4}	20	0.84	0.57	2^{4}	30	0.33	0.70	2^{0}	30	0.33	0.74
114902	2^{-16}	30	0.34	0.63	2^{0}	10	0.67	0.57	2^{0}	10	0.62	0.62
123902	2^{-12}	10	0.67	0.75	2^{-12}	10	0.67	0.75	2^{-8}	30	0.67	0.75
Avg			0.67	0.70			0.67	0.71			0.66	0.72

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table B.5: Training parameters and performance of Approach B with events duration of 30 minutes.

	Approach B with 30-min events											
	Event 1				Event 2				Event 3			
Patient	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}	C	k	SS_{sample}	SP_{sample}
402	2^{-8}	10	0.67	0.54	2^{-4}	10	0.67	0.54	2^{-4}	10	0.67	0.54
8902	2^{-8}	30	0.59	0.96	2^{-12}	30	0.66	0.85	2^{-8}	30	0.65	0.94
11002	2^{-20}	20	0.66	0.79	2^{-20}	20	0.67	0.77	2^{0}	10	0.67	0.76
16202	2^{0}	20	1.00	0.53	2^{-4}	30	1.00	0.56	2^{-8}	10	1.00	0.56
23902	2^{8}	10	0.67	0.85	2^{-12}	10	0.67	0.79	2^{-12}	30	0.66	0.79
30802	2^{-4}	20	0.67	0.90	2^{0}	10	0.65	0.77	2^{-12}	30	0.67	0.74
32702	2^{8}	30	0.42	0.71	2^{8}	10	0.65	0.59	2^{0}	10	0.33	0.72
46702	2^{4}	30	0.67	0.91	2^{8}	30	0.97	0.86	2^{-4}	30	0.66	0.91
50802	2^{0}	10	1.00	0.55	2^{-4}	10	0.96	0.60	2^{0}	10	1.00	0.50
53402	2^{4}	30	0.66	0.62	2^{0}	20	0.67	0.60	2^{-4}	20	0.67	0.60
55202	2^{-8}	30	0.66	0.66	2^{-4}	20	0.68	0.63	2^{4}	20	0.67	0.62
56402	2^{-20}	20	0.67	0.63	2^{-20}	20	0.67	0.63	2^{-8}	10	0.67	0.66
58602	2^{-4}	10	0.67	0.68	2^{8}	30	0.67	0.61	2^{0}	30	0.67	0.61
59102	2^{-12}	10	1.00	0.66	2^{-20}	20	1.00	0.70	2^{0}	30	0.70	0.96
60002	2^{-4}	20	0.61	0.64	2^{-20}	10	0.67	0.62	2^{0}	30	0.67	0.51
64702	2^{-12}	20	1.00	0.62	2^{-12}	20	1.00	0.63	2^{-8}	10	1.00	0.61
75202	2^{-12}	30	0.81	0.84	2^{-8}	10	0.63	0.87	2^{-8}	30	0.74	0.88
80702	2^{-12}	20	0.65	0.78	2^{-12}	20	0.58	0.79	2^{-8}	20	0.43	0.77
85202	2^{-16}	10	0.75	0.83	2^{-20}	10	0.80	0.83	2^{-20}	10	0.75	0.82
93402	2^{-4}	10	0.67	0.71	2^{0}	10	0.67	0.69	2^{8}	10	0.67	0.71
93902	2^{-4}	20	0.67	0.83	2^{0}	10	0.67	0.68	2^{0}	10	0.48	0.84
94402	2^{8}	20	0.67	0.56	2^{8}	30	0.67	0.57	2^{8}	30	0.67	0.41
95202	2^{-8}	10	0.67	0.65	2^{4}	10	0.52	0.64	2^{-20}	20	0.67	0.53
96002	2^{-20}	20	1.00	0.61	2^{8}	30	1.00	0.47	2^{-12}	30	0.96	0.64
98102	2^{8}	30	0.67	0.53	2^{-12}	30	0.53	0.71	2^{8}	20	0.67	0.54
98202	2^{-16}	10	0.67	0.81	2^{8}	10	0.67	0.76	2^{-4}	10	0.67	0.77
101702	2^{-20}	20	0.67	0.74	2^{-20}	20	0.67	0.72	2^{-20}	20	0.67	0.70
102202	2^{-4}	10	0.67	0.76	2^{8}	10	0.67	0.78	2^{4}	30	0.67	0.85
104602	2^{-12}	20	0.34	0.80	2^{-12}	20	0.34	0.80	2^{0}	20	0.33	0.94
109502	2^{0}	10	0.75	0.75	2^{0}	10	0.85	0.70	2^{0}	20	0.33	0.88
110602	2^{-16}	20	0.67	0.76	2^{-20}	20	0.67	0.75	2^{-20}	20	0.67	0.76
112802	2^{4}	30	1.00	0.63	2^{-20}	20	0.87	0.62	2^{-20}	20	0.97	0.61
113902	2^{-4}	10	0.67	0.74	2^{-4}	10	0.67	0.74	2^{-4}	10	0.67	0.78
114702	2^{-12}	20	0.34	0.67	2^{-8}	20	0.33	0.67	2^{0}	20	0.33	0.71
114902	2^{8}	30	0.98	0.58	2^{8}	30	0.81	0.71	2^{8}	10	0.67	0.57
123902	2^{-12}	10	0.67	0.77	2^{-16}	10	0.67	0.75	2^{-4}	10	0.67	0.75
Avg			0.71	0.71			0.71	0.69			0.67	0.71

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table B.6: Training parameters and performance of Approach C with 30-min pre-ictal.

	Approach C with 30-min pre-ictal			
Patient	C	k	SS_{sample}	SP_{sample}
402	2^{-4}	10	0.67	0.54
8902	2^{-8}	30	0.65	0.94
11002	2^{0}	10	0.67	0.76
16202	2^{-8}	10	1.00	0.56
23902	2^{-12}	30	0.66	0.79
30802	2^{-12}	30	0.67	0.74
32702	2^{8}	10	0.34	0.71
46702	2^{-4}	30	0.66	0.91
50802	2^{0}	10	1.00	0.50
53402	2^{-4}	20	0.66	0.60
55202	2^{4}	20	0.67	0.62
56402	2^{-8}	10	0.67	0.66
58602	2^{0}	30	0.67	0.61
59102	2^{0}	30	0.71	0.96
60002	2^{0}	30	0.67	0.51
64702	2^{-8}	10	1.00	0.61
75202	2^{-8}	30	0.74	0.88
80702	2^{-8}	20	0.43	0.77
85202	2^{-20}	10	0.75	0.82
93402	2^{8}	10	0.67	0.71
93902	2^{0}	30	0.67	0.69
94402	2^{8}	30	0.67	0.41
95202	2^{-20}	20	0.67	0.53
96002	2^{-12}	30	0.96	0.64
98102	2^{8}	30	0.67	0.55
98202	2^{-4}	10	0.67	0.77
101702	2^{-20}	20	0.67	0.70
102202	2^{4}	10	0.66	0.85
104602	2^{0}	20	0.33	0.94
109502	2^{0}	20	0.33	0.87
110602	2^{-20}	20	0.67	0.76
112802	2^{-20}	20	0.97	0.61
113902	2^{-4}	10	0.66	0.78
114702	2^{0}	30	0.33	0.74
114902	2^{0}	10	0.63	0.62
123902	2^{-8}	30	0.67	0.75
Avg			0.67	0.71

C: SVM cost; k: number of features; SS_{sample}: sample sensitivity; SP_{sample}: sample specificity

Table B.7: Seizure prediction performance and statistical validation results of Approach A with 20-min SOP.

Approach A with 20-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.00	0.04	0.00 ± 0.00	-	No
8902	2	0.50	0.14	0.02 ± 0.09	< 0.01	Yes
11002	2	0.00	0.00	0.00 ± 0.00	-	No
16202	4	0.00	0.00	0.00 ± 0.00	-	No
23902	2	0.50	0.55	0.02 ± 0.09	< 0.01	Yes
30802	5	0.60	4.01	0.33 ± 0.17	< 0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.00	0.00	0.00 ± 0.00	-	No
50802	2	0.00	0.84	0.30 ± 0.24	-	No
53402	2	0.00	0.19	0.08 ± 0.19	-	No
55202	5	0.40	0.35	0.07 ± 0.11	< 0.01	Yes
56402	3	0.00	0.00	0.00 ± 0.00	-	No
58602	4	0.00	0.00	0.00 ± 0.00	-	No
59102	2	0.50	6.49	0.30 ± 0.24	< 0.01	Yes
60002	3	0.33	0.02	0.00 ± 0.00	< 0.01	Yes
64702	2	0.50	2.08	0.30 ± 0.24	< 0.01	Yes
75202	4	0.25	0.19	0.03 ± 0.08	< 0.01	Yes
80702	3	0.00	0.04	0.01 ± 0.06	-	No
85202	2	0.00	0.00	0.00 ± 0.00	-	No
93402	2	0.00	1.62	0.17 ± 0.24	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.00	0.00 ± 0.00	-	No
95202	4	0.00	0.15	0.03 ± 0.08	-	No
96002	4	0.75	4.58	0.56 ± 0.23	< 0.01	Yes
98102	2	0.00	0.24	0.10 ± 0.20	-	No
98202	4	0.25	1.24	0.08 ± 0.12	< 0.01	Yes
101702	2	1.00	4.93	0.58 ± 0.34	< 0.01	Yes
102202	4	0.50	1.66	0.28 ± 0.24	< 0.01	Yes
104602	2	0.00	0.00	0.00 ± 0.00	-	No
109502	2	0.00	0.26	0.00 ± 0.00	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	0.00	0.74	0.11 ± 0.16	-	No
113902	4	0.75	0.46	0.16 ± 0.14	< 0.01	Yes
114702	6	0.17	0.06	0.03 ± 0.06	< 0.01	Yes
114902	4	0.00	0.10	0.03 ± 0.08	-	No
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Table B.8: Seizure prediction performance and statistical validation results of Approach B with 20-min SOP.

Approach B with 20-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.50	0.44	0.25 ± 0.25	< 0.01	Yes
8902	2	0.50	0.36	0.05 ± 0.15	< 0.01	Yes
11002	2	0.50	0.29	0.08 ± 0.23	< 0.01	Yes
16202	4	0.00	0.00	0.00 ± 0.00	-	No
23902	2	1.00	11.47	0.65 ± 0.35	< 0.01	Yes
30802	5	0.20	5.26	0.35 ± 0.17	-	No
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.00	0.00	0.00 ± 0.00	-	No
50802	2	0.50	5.48	0.67 ± 0.30	-	No
53402	2	0.00	0.83	0.23 ± 0.28	-	No
55202	5	0.40	3.38	0.40 ± 0.23	-	No
56402	3	0.33	3.56	0.29 ± 0.19	0.21	No
58602	4	0.00	0.00	0.00 ± 0.00	-	No
59102	2	0.50	6.66	0.32 ± 0.24	< 0.01	Yes
60002	3	0.33	7.63	0.59 ± 0.27	-	No
64702	2	0.50	3.19	0.33 ± 0.24	< 0.01	Yes
75202	4	0.25	0.69	0.15 ± 0.17	< 0.01	Yes
80702	3	0.33	2.27	0.21 ± 0.16	< 0.01	Yes
85202	2	1.00	2.17	0.55 ± 0.33	< 0.01	Yes
93402	2	0.00	3.34	0.18 ± 0.24	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.00	0.00 ± 0.00	-	No
95202	4	0.25	3.72	0.48 ± 0.20	-	No
96002	4	1.00	6.14	0.48 ± 0.20	< 0.01	Yes
98102	2	0.00	0.79	0.20 ± 0.28	-	No
98202	4	0.25	1.37	0.14 ± 0.12	< 0.01	Yes
101702	2	0.50	6.95	0.58 ± 0.37	-	No
102202	4	0.75	1.66	0.36 ± 0.21	< 0.01	Yes
104602	2	0.50	1.35	0.37 ± 0.22	< 0.01	Yes
109502	2	0.00	2.31	0.22 ± 0.25	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	1.00	22.79	0.63 ± 0.26	< 0.01	Yes
113902	4	0.75	0.82	0.21 ± 0.15	< 0.01	Yes
114702	6	0.17	0.11	0.03 ± 0.09	< 0.01	Yes
114902	4	0.75	2.08	0.30 ± 0.18	< 0.01	Yes
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Table B.9: Seizure prediction performance and statistical validation results of Approach C with 20-min SOP.

Approach C with 20-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.50	0.23	0.02 ± 0.09	< 0.01	Yes
8902	2	0.50	0.62	0.07 ± 0.17	< 0.01	Yes
11002	2	0.00	0.38	0.07 ± 0.17	-	No
16202	4	0.25	0.61	0.07 ± 0.11	< 0.01	Yes
23902	2	0.00	7.31	0.33 ± 0.34	-	No
30802	5	0.60	4.51	0.39 ± 0.14	< 0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.50	0.18	0.07 ± 0.17	< 0.01	Yes
50802	2	0.50	0.84	0.35 ± 0.23	< 0.01	Yes
53402	2	0.50	1.19	0.42 ± 0.26	0.09	No
55202	5	0.80	2.61	0.37 ± 0.16	< 0.01	Yes
56402	3	0.00	0.00	0.00 ± 0.00	-	No
58602	4	0.25	0.00	0.00 ± 0.00	< 0.01	Yes
59102	2	0.50	6.84	0.28 ± 0.25	< 0.01	Yes
60002	3	0.33	276.56	0.81 ± 0.29	-	No
64702	2	0.50	3.02	0.20 ± 0.24	< 0.01	Yes
75202	4	0.25	0.69	0.12 ± 0.14	< 0.01	Yes
80702	3	0.33	2.27	0.31 ± 0.08	0.15	No
85202	2	1.00	2.17	0.42 ± 0.29	< 0.01	Yes
93402	2	0.00	1.68	0.20 ± 0.24	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.18	0.06 ± 0.11	-	No
95202	4	0.25	3.72	0.47 ± 0.19	-	No
96002	4	1.00	5.65	0.53 ± 0.24	< 0.01	Yes
98102	2	0.00	0.79	0.28 ± 0.25	-	No
98202	4	0.25	1.30	0.14 ± 0.12	< 0.01	Yes
101702	2	1.00	5.82	0.67 ± 0.30	< 0.01	Yes
102202	4	0.75	1.60	0.27 ± 0.17	< 0.01	Yes
104602	2	0.50	1.35	0.43 ± 0.17	0.04	Yes
109502	2	0.00	1.49	0.22 ± 0.25	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	1.00	22.79	0.51 ± 0.21	< 0.01	Yes
113902	4	0.75	0.57	0.12 ± 0.12	< 0.01	Yes
114702	6	0.17	0.11	0.03 ± 0.07	< 0.01	Yes
114902	4	0.50	1.76	0.27 ± 0.18	< 0.01	Yes
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Table B.10: Seizure prediction performance and statistical validation results of Approach A with 30-min SOP.

Approach A with 30-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.50	0.30	0.25 ± 0.25	< 0.01	Yes
8902	2	0.00	0.26	0.03 ± 0.12	-	No
11002	2	0.00	0.00	0.00 ± 0.00	-	No
16202	4	0.25	0.13	0.06 ± 0.11	< 0.01	Yes
23902	2	0.50	0.10	0.23 ± 0.25	< 0.01	Yes
30802	5	0.60	3.01	0.41 ± 0.17	< 0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.00	0.00	0.00 ± 0.00	-	No
50802	2	0.50	0.52	0.40 ± 0.20	0.01	Yes
53402	2	0.00	0.34	0.18 ± 0.24	-	No
55202	5	0.60	0.48	0.16 ± 0.12	< 0.01	Yes
56402	3	0.00	0.00	0.00 ± 0.00	-	No
58602	4	0.00	0.00	0.00 ± 0.00	-	No
59102	2	0.50	4.52	0.37 ± 0.22	< 0.01	Yes
60002	3	0.33	1.60	0.42 ± 0.26	-	No
64702	2	0.50	4.27	0.37 ± 0.22	< 0.01	Yes
75202	4	0.25	0.12	0.07 ± 0.11	< 0.01	Yes
80702	3	0.00	0.00	0.00 ± 0.00	-	No
85202	2	0.50	1.52	0.38 ± 0.21	0.01	Yes
93402	2	0.00	1.17	0.12 ± 0.21	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.00	0.00 ± 0.00	-	No
95202	4	0.50	2.27	0.43 ± 0.21	0.10	No
96002	4	1.00	3.60	0.54 ± 0.17	< 0.01	Yes
98102	2	0.50	0.09	0.05 ± 0.15	< 0.01	Yes
98202	4	0.00	0.00	0.00 ± 0.00	-	No
101702	2	1.00	3.62	0.65 ± 0.29	< 0.01	Yes
102202	4	0.75	0.42	0.23 ± 0.17	< 0.01	Yes
104602	2	0.00	0.00	0.00 ± 0.00	-	No
109502	2	0.00	0.47	0.20 ± 0.24	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	0.00	0.48	0.13 ± 0.16	-	No
113902	4	0.50	0.45	0.16 ± 0.20	< 0.01	Yes
114702	6	0.00	0.00	0.00 ± 0.00	-	No
114902	4	0.00	0.56	0.19 ± 0.14	-	No
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Table B.11: Seizure prediction performance and statistical validation results of Approach B with 30-min SOP.

Approach B with 30-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	0.50	1.23	0.38 ± 0.33	0.07	No
8902	2	0.50	0.46	0.13 ± 0.26	< 0.01	Yes
11002	2	0.00	0.32	0.05 ± 0.15	-	No
16202	4	0.00	0.00	0.00 ± 0.00	-	No
23902	2	0.00	10.12	0.73 ± 0.31	-	No
30802	5	0.60	3.91	0.45 ± 0.18	< 0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.00	0.00	0.00 ± 0.00	-	No
50802	2	0.50	3.42	0.57 ± 0.28	-	No
53402	2	0.00	0.36	0.20 ± 0.28	-	No
55202	5	0.80	5.61	0.65 ± 0.26	< 0.01	Yes
56402	3	0.67	2.49	0.33 ± 0.19	< 0.01	Yes
58602	4	0.00	0.00	0.00 ± 0.00	-	No
59102	2	0.50	4.86	0.35 ± 0.23	< 0.01	Yes
60002	3	1.00	5.86	0.66 ± 0.29	< 0.01	Yes
64702	2	0.00	1.76	0.27 ± 0.25	-	No
75202	4	0.50	0.44	0.14 ± 0.15	< 0.01	Yes
80702	3	0.33	1.64	0.29 ± 0.11	0.04	Yes
85202	2	1.00	1.52	0.47 ± 0.31	< 0.01	Yes
93402	2	0.00	1.22	0.25 ± 0.25	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.00	0.00 ± 0.00	-	No
95202	4	0.50	2.46	0.42 ± 0.17	0.01	Yes
96002	4	0.75	4.18	0.60 ± 0.20	< 0.01	Yes
98102	2	0.50	0.52	0.22 ± 0.25	< 0.01	Yes
98202	4	0.00	1.07	0.14 ± 0.12	-	No
101702	2	1.00	4.61	0.67 ± 0.32	< 0.01	Yes
102202	4	0.75	1.29	0.31 ± 0.27	< 0.01	Yes
104602	2	0.00	0.00	0.00 ± 0.00	-	No
109502	2	0.00	1.65	0.27 ± 0.25	-	No
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	1.00	16.79	0.59 ± 0.21	< 0.01	Yes
113902	4	0.75	0.65	0.22 ± 0.20	< 0.01	Yes
114702	6	0.00	0.12	0.01 ± 0.03	-	No
114902	4	0.75	4.13	0.49 ± 0.15	< 0.01	Yes
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Table B.12: Seizure prediction performance and statistical validation results of Approach C with 30-min SOP.

Approach C with 30-min SOP						
Patient	**Evaluated seizures**	**SS$_{sensitivity}$**	**FPR/h**	**Surrogate SS$_{sensitivity}$**	**p-value**	**Above chance**
402	2	1.00	1.00	0.45 ± 0.30	< 0.01	Yes
8902	2	0.50	0.26	0.08 ± 0.19	< 0.01	Yes
11002	2	0.50	0.23	0.13 ± 0.26	< 0.01	Yes
16202	4	0.25	0.13	0.04 ± 0.09	< 0.01	Yes
23902	2	1.00	6.98	0.60 ± 0.35	< 0.01	Yes
30802	5	0.60	3.42	0.42 ± 0.17	< 0.01	Yes
32702	2	0.00	0.00	0.00 ± 0.00	-	No
46702	2	0.50	0.09	0.05 ± 0.15	< 0.01	Yes
50802	2	1.00	4.33	0.70 ± 0.31	< 0.01	Yes
53402	2	0.00	0.36	0.03 ± 0.12	-	No
55202	5	0.80	1.92	0.41 ± 0.17	< 0.01	Yes
56402	3	0.00	0.00	0.00 ± 0.00	-	No
58602	4	0.00	0.00	0.00 ± 0.00	-	No
59102	2	0.00	6.72	0.42 ± 0.19	-	No
60002	3	1.00	6.29	0.72 ± 0.26	< 0.01	Yes
64702	2	0.00	5.17	0.38 ± 0.21	-	No
75202	4	0.50	0.41	0.14 ± 0.14	< 0.01	Yes
80702	3	0.33	1.64	0.28 ± 0.12	0.02	Yes
85202	2	1.00	1.35	0.47 ± 0.29	< 0.01	Yes
93402	2	0.00	1.22	0.17 ± 0.24	-	No
93902	2	0.00	0.00	0.00 ± 0.00	-	No
94402	4	0.00	0.11	0.03 ± 0.08	-	No
95202	4	0.50	2.39	0.47 ± 0.21	0.51	No
96002	4	1.00	4.18	0.62 ± 0.23	< 0.01	Yes
98102	2	0.50	0.56	0.20 ± 0.24	< 0.01	Yes
98202	4	0.25	1.00	0.16 ± 0.14	< 0.01	Yes
101702	2	1.00	4.61	0.62 ± 0.38	< 0.01	Yes
102202	4	0.75	1.17	0.39 ± 0.25	< 0.01	Yes
104602	2	0.50	0.95	0.42 ± 0.19	0.02	Yes
109502	2	0.50	1.04	0.23 ± 0.25	< 0.01	Yes
110602	2	0.00	0.00	0.00 ± 0.00	-	No
112802	3	1.00	16.79	0.61 ± 0.21	< 0.01	Yes
113902	4	0.50	0.58	0.15 ± 0.18	< 0.01	Yes
114702	6	0.00	0.09	0.03 ± 0.06	-	No
114902	4	0.25	0.84	0.23 ± 0.21	0.67	No
123902	2	0.00	0.00	0.00 ± 0.00	-	No

SS$_{seizure}$: seizure sensitivity; Surrogate SS$_{seizure}$: seizure sensitivity from surrogate predictor

Table B.13: Constitution of the patient groups for each stratification criteria.

Patient	FOA/FOIA	Rhythmic patterns	Awake	Day	Night
402		x	x	x	
8902	x		x		x
11002	x	x		x	
16202			x		x
23902	x	x	x	x	
30802	x	x			x
32702	x		x	x	
46702		x			x
50802		x			x
53402	x	x		x	
55202	x		x	x	
56402		x	x	x	x
58602	x			x	x
59102	x	x	x	x	
60002	x				x
64702				x	x
75202	x	x		x	x
80702			x	x	x
85202	x			x	x
93402		x			x
93902		x		x	x
94402	x	x		x	
95202				x	x
96002	x	x	x	x	
98102			x		x
98202		x	x	x	x
101702	x			x	x
102202	x	x		x	x
104602		x		x	x
109502	x	x		x	x
110602	x	x	x	x	
112802	x	x	x	x	x
113902	x	x		x	x
114702	x	x		x	x
114902				x	
123902		x			

Note: For the formation of FOA/FOIA and rhythmic patterns groups, upon the presence of unknown classification ("UC") and unclear patterns ("?") respectively, it was assumed that they complied with the annotations of the remaining seizures.