UNIVERSIDADE D
COIMBRA

Edson David dos Santos Sousa Pontes

Concept-Drifts Adaptation for EEG Epilepsy Seizure Prediction

Thesis submitted to the Faculty of Sciences 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 Prof. Dr. Mauro Pinto

January 2023

Faculty of Sciences and Technology

University of Coimbra

Concept-Drifts Adaptation for EEG Epilepsy Seizure Prediction

Edson David dos Santos Sousa Pontes

Supervisor:
César Alexandre Domingues Teixeira

Co-supervisor:
Mauro Filipe da Silva Pinto

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

Coimbra, 2023

This work was developed in collaboration with:

CISUC - Center for Informatics and Systems of the University of Coimbra

This work was supported by the Portuguese Foundation for Science and Technology (FCT), along with projects CISUC (UID/CEC/00326/2020) and project RECoD (PTDC/EEI-EEE/5788/2020), both financed by national funds through the FCT – Foundation for Science and Technology, I.P..

Great things may come to those who wait,
but only the things left by those who hustle.
Abraham Lincoln

Acknowledgements

I thank my supervisors, Professor César Teixeira and Professor Mauro Pinto, for all the guidance, support and patience throughout these months. I could not have done this without them. I would also like to thank Dr Adriana Leal for all the help and for what I have learned from her.

I also thank the remaining colleagues and friends from the laboratory for all the discussions and sharing of ideas.

A heartfelt thank you to Miguel Carvalho, José Sousa, João Loureiro, Diogo Cruz, Diogo Oliveira, Henrique Tavares, Valéria Lopes, and André Silva, with whom I experienced these five years of despair and adventures. It was my pleasure to have you by my side.

To Biomedical Engineering and my colleagues, with whom, every day, I learned something new. To all the people in Coimbra and Enschede that made my academic life so unique and worthy of a story book, thank you for showing me that 1 year can be worth 100 of experiences.

A special thank you to Renato Barroso, Inês Cebola and Francisco Dias for their lifelong friendship, which began when we were four years old. It was my absolute privilege to have grown with you.

I want to express my genuine gratitude to my family, who made me who I am today. To my mother for being my role model, supporting me and encouraging me to go further. To my sister, thank you for always being there, especially during this challenging year.

To Coimbra, the city where I was born, gave my first steps and uttered my first words. I never thought I would return here after 15 years, but now, I cannot imagine any other way how.

Last but not least, I want to thank me. I want to thank me for believing in me. I want to thank me for doing all this hard work. I want to thank me for never quitting.

Abstract

The administration of antiepileptic drugs or surgical interventions fails to control seizures in about 30% of epileptic patients. Seizure prediction is a viable strategy for enhancing their quality of life because it can be used in intervention or warning systems. These systems would try to stop seizures from happening or, at the very least, lessen their adverse effects. Identifying the preictal interval, which marks the change from regular brain activity to a seizure, is a critical part of this research field.

Even though several predictive studies applied various Electroencephalogram (EEG) based methodologies, very few have been applied in medical devices, and none have been clinically applicable. Recent studies have shown that tracking and handling concept drifts is highly relevant in seizure prediction; therefore, it is important to develop methods able to automatically detect and handle changes in context without human intervention.

The present work aimed to evaluate the impact of automatic concept drift adapting methods in seizure prediction. For this reason, two analyses were conducted. The first analysis aimed to deal with concept drifts by simply retraining the model; for that, two different iterative retraining strategies were developed (Add-One-Forget-One and Chronological). Those two methods were compared to the most common partition method, the Control partitioning. The three strategies were tested on the Control approach, developed to predict seizure without incorporating intrinsic concept drift adaptation. The method with the best performance was used in the second analysis.

The second analysis tested approaches to predicting seizures while intrinsically adapting to concept drifts during the model's learning process; for that, three patient-specific seizure prediction approaches (Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble) with a 10-minute seizure prediction horizon were proposed and compared to the Control. The proposed method-

ology combines a set of univariate linear features with a classifier based on Support Vector Machines (SVMs) and the Firing Power as a post-processing technique to generate alarms before seizures.

Considering a group of 37 patients with Temporal Lobe Epilepsy (TLE) from the EPILEPSIAE database, the best-performing approach (Backwards-Landmark Window with the Add-One-Forget-One method) aimed to select data from the concept closest to the preictal period of the last training seizure; this led to results of 0.75 $\pm$ 0.33 for sensitivity and 1.03 $\pm$ 1.00 for false positive rate per hour. Even though the best performing approach statistically validated 89% of the patients, it is necessary to determine the maximum false positive rate appropriate for each intervention system.

Keywords: Epilepsy, Seizure Prediction, Electroencephalogram, Concept Drifts, Machine Learning

Resumo

A administração de medicamentos antiepilépticos ou intervenções cirúrgicas falha em controlar crises em cerca de 30% dos doentes com epilepsia. A previsão de crises é uma estratégia viável para melhorar a qualidade de vida desses doentes, pois pode ser usada em sistemas de intervenção ou alerta. Esses sistemas tentariam evitar a ocorrência de crises ou, pelo menos, diminuir os seus efeitos adversos. Identificar o intervalo pré-ictal, que marca a mudança da atividade cerebral regular para uma crise, é uma parte crítica deste campo de investigação.

Embora vários estudos de previsão estejam a usar vários métodos baseados em eletroencefalograma, apenas alguns foram aplicados em dispositivos médicos e nenhum foi clinicamente viável. Estudos recentes mostram que a adaptação a mudanças de conceito é altamente relevante na previsão de crises; portanto, é importante desenvolver métodos capazes de detetar e lidar automaticamente com mudanças de contexto sem intervenção humana.

O presente trabalho teve como objetivo avaliar o impacto dos métodos de adaptação automática a mudanças de conceito na predição de crises. Por este motivo, foram realizadas duas análises. A primeira análise visava lidar com os desvios de conceito simplesmente re-treinando o modelo; para tal, foram desenvolvidas duas estratégias para re-treinar (*Add-One-Forget-One* e *Chronological*). Estes dois métodos foram comparados com o método de partição mais comum, o *Control partitioning*. As três estratégias foram testadas na abordagem *Control*, desenvolvida para prever crises sem incluir mecanismos para adaptação automática a mudanças de conceito. O método com a melhor performance foi usado na segunda análise.

A segunda análise testou abordagens para previsão de crises que se adaptam intrinsecamente a variações de conceito durante o processo de aprendizagem do modelo; para isso, foram propostas três abordagens de previsão de crises específicas

para cada doente que foram comparadas ao *Control* (*Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble*) com um horizonte de previsão de crise de 10 minutos. A metodologia proposta combina um conjunto de *features* lineares univariadas com um classificador baseado em *Support Vector Machines* e o *Firing Power* como técnica de pós-processamento para gerar alarmes antes das crises.

Considerando um grupo de 37 doentes com Epilépsia do Lobo Temporal da base de dados EPILEPSIAE, a abordagem com melhor performance (*Backwards-Landmark Window* com *Add-One-Forget-One*) visou selecionar dados do conceito mais próximo do período pré-ictal da última crise de treino; isso levou a resultados de 0,75 $\pm$ 0,33 para sensibilidade e 1,03 $\pm$ 1,00 para taxa de falsos positivos por hora. Embora a abordagem com melhor performance tenha validado estatisticamente 89% dos pacientes, é necessário determinar a taxa máxima de falsos positivos apropriada para cada sistema de intervenção.

Palavras-chave: Epilepsia, Previsão de Crises, Eletroencefalograma, Aprendizado de Máquina, Mudanças de Conceito

Contents

List of Figures

List of Tables

List of Acronyms

AED Anti-Epileptic Drug 1, 8–10

ANFIS Adaptive Neuro-Fuzzy Inference Systems 40

CD Concept drift 45, 52

CNN Convolutional Neural Network 36, 40–43, 57

DBS Deep Brain Stimulation 10

DL Deep Learning xvi, 31, 36, 38, 40, 41, 43, 51, 52

DRE Drug-Resistant Epilepsy 1, 9, 27, 55

EEG Electroencephalogram vii, xi, xv, xvi, xix, 1–5, 7, 8, 12–19, 23–25, 27, 29–37, 43, 45, 51, 55, 57, 58, 67, 87, 96, 119–122, 124

FFT Fast Fourier Transform 42, 121, 122

FIFO First-In-First-Out 127, 132

FIR Finite Impulse Response 35

FOA Focal Onset Aware 6, 7, 87, 89

FOIA Focal Onset Impaired Awareness 7, 87, 89

FPR/h False Positive Rate per Hour xvii, 20, 21, 23, 27, 36, 44, 45, 52, 66–68, 73, 77, 82, 86–88, 96, 99, 133

GRU Gated Recurrent Unit 40

IBE International Bureau for Epilepsy 5

iEEG Invasive Electroencephalogram 14–16, 27, 32

IIR Infinite Impulse Response 35

ILAE International League Against Epilepsy 5, 6, 8, 9

IT Intervention Time 18

kNN k-nearest neighbors 49

LOOCV Leave-One-Out Cross-Validation xvii, 61, 62

LSTM Long Short-Term Memory 40, 43, 52

mDAD maximum Difference Amplitude Distribution of histogram 40

MPC Mean Phase Coherence 124

mRMR minimum Redundance Maximum Relevance 40

PCA Principal Component Analysis 40

PSD Power Spectral Density 121

RBF Radial Basis Function 41, 61

RNN Recurrent Neural Network 43

RNS Responsive Neurostimulation System 10, 45

SEF Spectral edge frequency 122

SOP Seizure Occurrence Period xv, 2, 18–21, 23, 27, 34, 36, 37, 44, 58, 59, 61, 66, 69, 70, 78, 89, 96, 97

SPH Seizure Prediction Horizon xv, 2, 3, 18–21, 24, 27, 34, 36, 37, 44, 58, 66, 70, 97

SVM Support Vector Machine viii, 3, 40, 41, 54, 61, 63–65, 67, 69, 89, 90, 95, 131

TiW Time in Warning 45

TLE Temporal Lobe Epilepsy viii, 9, 15, 27

VNS Vagus Nerve Stimulation 10

Chapter 1

Introduction

Seizure prediction using the Electroencephalogram (EEG) signal has made tremendous improvements during the past forty years. Prediction models must deal with the inherent complexities of epilepsy, its seizures and brain dynamics, where their decisions could substantially impact the patient's life. Furthermore, machine learning's primary goal has traditionally been to learn from data assumed to be sufficient and representative of the underlying fixed, yet unknown, distribution. Real-world problems, such as the case of EEG seizure prediction, rarely fit those assumptions. For example, class distributions are often skewed, which leads to a class imbalance issue. Often, the data used is acquired from a non-stationary distribution where changes in the hidden context or data distribution may occur, leading to concept drift. When a learner does not account for these changes, a decrease in performance usually occurs [9]. Therefore, dealing with concept drifts and their influence in the EEG may provide important information for seizure prediction. This chapter explains the present work's motivation, objectives, main goals and limitations.

1.1 Motivation

Over 50 million people worldwide have epilepsy, one of the most prevalent neurological diseases. This illness is characterised by abnormal brain activity, which can cause seizures or unusual behaviours, sensations, and possibly loss of awareness. There are a variety of neurological, cognitive, psychological, and social effects brought on by this unusual activity [14].

Anti-Epileptic Drugs (AEDs) are the first-line treatment for epilepsy. Nevertheless, nearly one-third of patients—those with Drug-Resistant Epilepsy (DRE)—cannot control their seizures with medicine alone [15]. These individuals are more likely to experience a range of psychological issues, including psychosis, sadness, anxiety, and, in the worst case, early death [15, 16]. Even though epilepsy surgery is a tried-and-true treatment for DRE patients, very few people qualify for it [16].

When complete seizure remission cannot be achieved with medication alone, seizure prediction plays an essential role in clinical management and treatment. The quality of life for patients vulnerable to the sudden occurrence of seizures can be improved by seizure prediction.

1.2 Goals

The seizure prediction field seeks to create an algorithm to anticipate epileptic seizures and release alarms before the seizure onset. In this field, the assumption that a transitional period exists is made, the preictal, which is a period that comes before the seizure and forms the basis of the seizure prediction domain. The EEG signals' ability to record the preictal period has allowed this field to advance. [7].

Ultimately, for warning devices, the goal is to build an online system to process the data to inform the patient of an impending seizure in a well-defined time window, the Seizure Occurrence Period (SOP) after a predefined Seizure Prediction Horizon (SPH). The SPH duration must be enough to allow the patient to take action. For closed-loop intervention devices, the goal is also to develop an online system that can control the seizure by administering anti-convulsive medications or inducing electrical stimulation. These systems may offer new therapeutic alternatives that help the patient avoid unsafe conditions [7, 17].

1.3 Seizure prediction challenges

There are several factors that difficult seizure prediction. They are: the heterogeneity of seizures and epilepsy syndromes, a significant class imbalance brought on by the infrequent occurrence of seizures, and concept drifts. And when applicable, also, the complexity of the EEG signal [17–19].

Due to its complexity, the scientific community still needs to understand the EEG signal completely. Furthermore, most of the available EEG databases comprise presurgical monitoring conditions, which does not accurately reflect everyday seizure dynamics. Thus, ultra long-term recordings made over several months or years and collected during daily life strengthen the therapeutic applicability of the developed techniques [4, 12, 17, 20].

It is important to note that the preictal period is the most challenging to identify and manually annotate by specialists because it is associated with significant diversity. As a result, no guideline or ideal value has been established for its duration. There is proof that it can change between patients, even between their seizures. As a result, seizure prediction is significantly impacted by this state's complexity [12, 19, 20].

The relative rarity of seizures results in a significantly longer interictal interval than the preictal; this leads to class imbalance, which is a severe issue. This problem

can lead to a specialisation of the classifier on the interictal class [12].

Concept drifts are yet another problematic issue. They show up as misleading aspects in the EEG signal and may have a negative effect on how well the seizure prediction models work. Such concept drifts are associated with oscillations of the patient seizure propensity, which can be strongly influenced by variables such as circadian, ultradian and infradian cycles, vigilance states and sleep quality, and changes in medication intake [17–22]. Therefore it is important to have methods that automatically assess data quality over time and that are capable of retraining the classifier.

1.4 Objective

This thesis aims to ascertain if it is possible to improve seizure prediction by dealing with the existence of concept drifts. For this purpose, two analyses were done. The first analysis aimed to deal with concept drifts by simply retraining the model; for that, two different iterative retraining strategies were developed (Add-One-Forget-One and Chronological). Those two methods were compared to the most common partition method, the Control partitioning. The three strategies were tested on the Control approach, developed to predict seizure without incorporating intrinsic concept drift adaptation. The method with the best performance was used in the second analysis. Each method is characterised as follows: i) the Control partitioning is a common partition method where the first three chronological seizures are used to train the model; ii) the Add-One-Forget-One is a retraining method where only the last three chronological labelled seizures are used to train the model; iii) the Chronological is a retraining method where all the past seizures are used to train the model.

The second analysis tested approaches to predicting seizures while intrinsically adapting to concept drifts during the model's learning process; for that, three patient-specific seizure prediction approaches (Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble) with a 10-minute SPH were proposed and compared to the Control. The approaches are described as follows: i) the Control is a common seizure prediction algorithm; ii) the Backwards-Landmark Window is a seizure prediction algorithm incorporating a window adjustment method by optimising performance with Support Vector Machines (SVMs) [**?**, 23]; iii) the Seizure-batch Regression is a seizure prediction algorithm incorporating a data-batch (seizures) selection method using a logistic regression [24]; iv) the Dynamic Weighted Ensemble is a seizure prediction algorithm with a dynamic integration of classifiers [21].

Using long-term EEG data and machine learning algorithms, the expected contributions of this thesis are the following:

- Development of a patient-specific methodology for seizure prediction using scalp EEG signals from the European Epilepsy Database (EPILEPSIAE).
- Integrating different concept drift adaptation techniques in the prediction algorithms to evaluate its impact.

1.5 Structure

Beyond the introduction, this thesis contains more six chapters structured as follows.

Chapter 2 provides background information related to epilepsy, EEG, Seizure prediction and Concept drifts.

Chapter 3 presents a literature overview on EEG seizure prediction and Concept drift adaptation.

Chapter 4 describes all the steps of the adopted methodology.

Chapter 5 reports the results obtained in this work and their interpretative analysis.

Chapter 6 provides a detailed discussion about the obtained findings and limitations.

Chapter 7 presents the conclusions and addresses future work.

Chapter 2

Background Concepts

This chapter is dedicated to the introduction of the main concepts required to understand this document. Section 2.1 presents the definitions related to epilepsy. Section 2.2 includes an overview of Electroencephalogram (EEG) signal, section 2.3 provides theoretical aspects related to seizure prediction, and section 2.4 provides theoretical aspects associated to concept drifts. Finally, section 2.5 provides a summary of the background key concepts.

2.1 Epilepsy and Seizures

Epilepsy is the most common chronic neurological disease. It is characterised by brief and recurrent episodes, known as seizures, affecting more than 0.5% of the world population. They cause profound physical, psychological, and social consequences [25, 26].

Epileptic seizures can vary from the briefest lapses of attention or muscle jerks to severe and prolonged convulsions. Seizures can also vary in frequency, from less than one per year to several per day [14]. They result from uncontrolled electrical discharges [26], that arise from abnormal enhanced synchronisation of neurons. Large populations of neurons become excited, demonstrating high frequencies and amplitudes simultaneously in different areas of the brain leading to electrical bursts of energy [27, 28].

Having a seizure does not automatically mean that the patient suffers from epilepsy disease. As stated by the International League Against Epilepsy (ILAE) and International Bureau for Epilepsy (IBE) in 2005 [26], epilepsy should be defined as "*a disorder of the brain characterised by an enduring predisposition to generate epileptic seizures and by the neurobiological, cognitive, psychological, and social consequences of this condition. The definition o epilepsy requires the occurrence of at least one epileptic seizure*".

Even though epilepsy has traditionally been referred to as a disorder or a family of disorders rather than a disease, in 2014, the ILAE and the IBE agreed that epilepsy

is best considered a disease and updated the previous definition for epilepsy [29]. According to this update, epilepsy is a disease of the brain defined by any of the following conditions:

- at least two unprovoked (or reflex) seizures occurring $\geq$ 24h apart;
- 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 ten years;
- diagnosis of an epilepsy syndrome.

As seen in Figure 2.1, ILAE updated Epilepsy classifications where three levels were presented: firstly, the seizure type, secondly the epilepsy type, and thirdly, epilepsy syndrome. The new classification addresses etiology at every stage, emphasising the importance of considering it at each step since it often carries significant treatment implications. In the interest of simplicity, etiology influence will not be developed in this document [1, 29].

Furthermore, the main goal of this update was to enhance the clarity and clinical relevance associated with the diagnostic process. Hence, more emphasis was granted to the first unprovoked epileptic seizure in individuals who possess other factors coupled with the high likelihood of a perpetual lowered seizure threshold. Consequently, a high recurrence risk [29]. The term "unprovoked" indicates the absence of a temporary or reversible factor that reduces the threshold causing a seizure at any time.

Epilepsy is not necessarily a life-long condition as it is considered to be resolved for individuals with an age-dependent epilepsy syndrome but are now past the applicable age or for those who have remained seizure-free for the last ten years, with at least five years without anti-epileptic drugs [29].

2.1.1 Seizure types

The physician's first task is to establish that an event has the essential qualities of a seizure. Upon identifying the event as an epileptic seizure, the physician must define its type. The classification of a seizure type begins by establishing whether the initial manifestations (onset) are focal, generalised, or unknown, as seen in Figure 2.2.

Focal seizures have their onset in just one cerebral hemisphere, while generalised ones have their onset in both hemispheres. In technical terms, focal seizures are formed within a neural network in a single cerebral hemisphere, whereas the engagement of bilateral networks characterises generalised seizures which do not necessarily have to be symmetrical. Furthermore, seizure onset could also be defined as motor or non-motor, depending on the symptoms during a seizure. [1]

Also, focal seizures can be classified regarding the state of awareness and environment during the event. If the patient is aware, the seizure is classified as Focal Onset

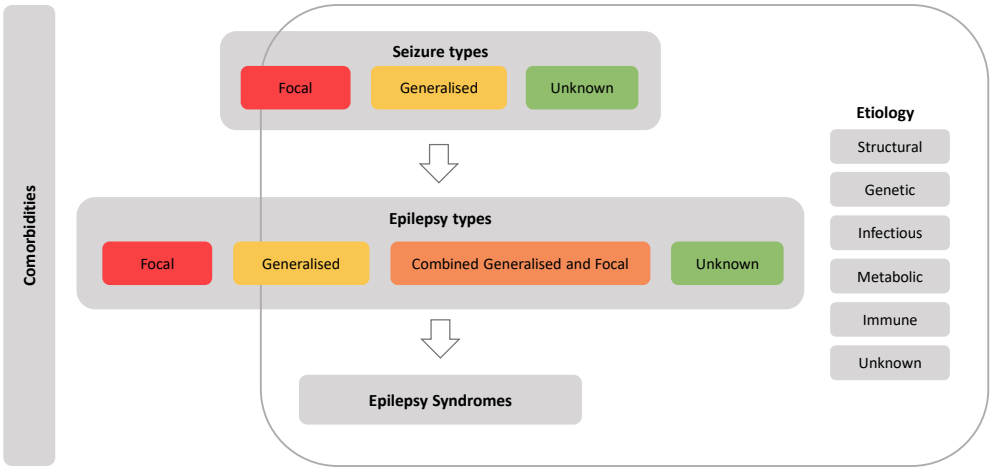

Figure 2.1: ILAE 2017 framework for classification of epilepsy. *Denotes onset of seizure. Source: Sheffer et al. 2017 [1]

Aware (FOA), otherwise is Focal Onset Impaired Awareness (FOIA). In addition, Focal to bilateral tonic-clonic is a particular type o seizure that rapidly propagates to the opposite cerebral hemisphere despite having a focal onset. These often lead to tonic (body stiffness) and clonic (jerking movements) symptoms. [1]

Finally, the seizure is classified as unknown when the onset location is not found. When there is no significant degree of confidence whether or not a seizure has a focal or generalised onset, the seizure is classified as unclassified [1].

2.1.2 Epilepsy type

The second level of epilepsy diagnosis concerns the type. As seen in Figure 2.1 epilepsy disease can be classified into four types:

- Focal epilepsy includes unifocal and multifocal disorders and seizures involving one hemisphere. Several seizure types can be included in this group, such as FOA, FOIA, Focal motor, Focal non-motor, and Focal to bilateral tonic-clonic. Focal epileptiform discharges characterise the interictal EEG.

- Generalised epilepsy, the patient typically shows generalised spike-wave activity on EEG and en-globe the following seizure types: absence, tonic-clonic, tonic, atonic, and myoclonic.

- In Combined Generalised and Focal, the interictal EEG may show both generalised spike-wave and focal epileptiform discharges.

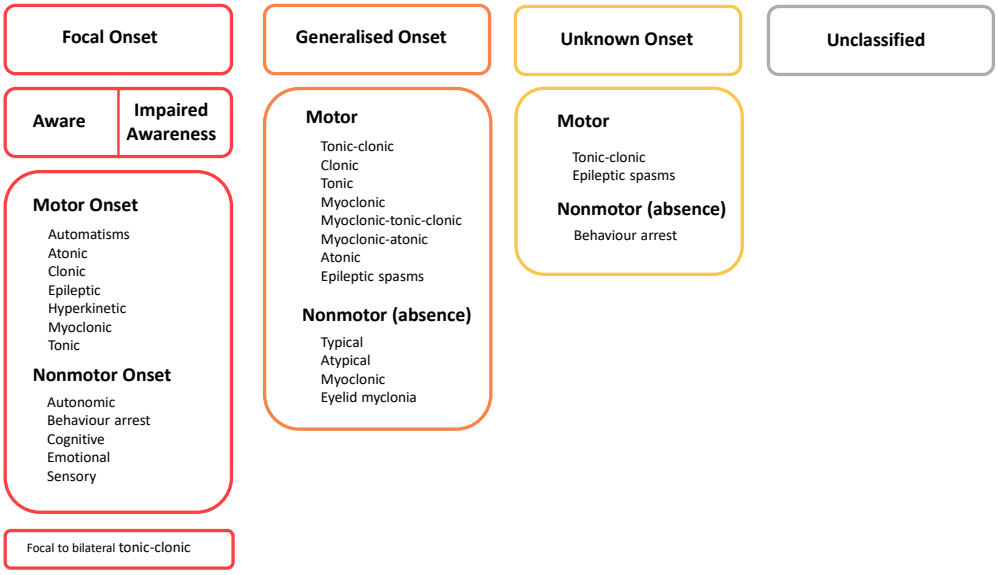

Figure 2.2: The expanded ILAE 2017 operational classification of seizure types. Adapted from Fisher et al. 2017 [2].

- Epilepsy is usually classified as unknown whenever the physician does not possess enough information for proper classification. Also, it is worth mentioning that when the seizure type is unknown, epilepsy is commonly classified as unknown.

Each epilepsy type is coupled with a high degree of complexity, as each category contains multiple types of seizures. The Epilepsy type may also be the final level of diagnosis achievable when the physician cannot proceed to an Epilepsy Syndrome diagnosis [1].

2.1.3 Epilepsy syndrome

The third level in epilepsy diagnosis is identifying the Epilepsy Syndrome. An epilepsy syndrome refers to a cluster of features incorporating seizure types, specific findings on EEG, and brain imaging studies which tend to occur together. Moreover, others can also be considered, such as the age of onset and remission, diurnal variation, and seizure triggers. Even though numerous well-known syndromes exist, no formal classification has been established by the ILAE [1, 30].

Identifying an epilepsy syndrome is beneficial as it provides information on the underlying etiologies to be considered and which anti-seizure drugs might be the most advantageous. The syndrome identification is particularly relevant, given that several epilepsy syndromes demonstrate seizure aggravation with certain Anti-Epileptic Drugs (AEDs), which can be circumvented through the appropriate early diagnosis [30].

Temporal Lobe Epilepsy (TLE)

Most patients submitted to presurgical monitoring and resistant to AEDs suffer from TLE. Most of the patients studied in this thesis have TLE syndrome. The most common syndrome in adults is the TLE. It affects approximately 60% of the entire epilepsy community. TLE starts typically in late childhood and adolescence and is distinguished by seizures involving the temporal lobes [31, 32]. Most patients suffer from focal seizures (aware and impaired awareness), and some suffer from focal to bilateral tonic-clonic seizures. It is complicated for a patient suffering from TLE to become ultimately seizure-free with anti-seizure drugs alone, even though these drugs may lower the number of seizures. In such cases, surgery is considered an option to control seizure occurrence, and when it does not work, an alternative to increasing quality of life is seizure prediction [32].

2.1.4 Treatment and therapeutics

Treatment goals for epilepsy are complete seizure remission and no side effects, but these goals are often not achieved. Medication with AED is the initial treatment plan for patients with epilepsy, and when it is ineffective in preventing seizures, other treatments are available. These treatment options include resective surgery, neurostimulation, dietary therapy, and warning devices. Epilepsy treatment offers the opportunity to live free from stigma and discrimination in all parts of the world [33], and the chance to avoid irreversible psychological and social problems, a lifetime of disability, and premature death [15].

2.1.4.1 Antiepileptic drugs and drug-resistant epilepsy

The purpose of AEDs is to alter the balance of excitation and inhibition that characterises epilepsy, which is caused by hyper excitatory or hypersynchronous neural activity. Approximately 30 AEDs are currently in use to regulate seizures by restraining excitatory pathways, modifying voltage-gated ion channels, or enhancing inhibitory mechanisms [34].

Nowadays, in approximately 30% of patients, medication is not effective. These individuals are considered patients with Drug-Resistant Epilepsy (DRE) [35]. DRE according to the ILAE in 2009, is it defined as "*a failure of adequate trials of two tolerated, appropriately chosen and used AED schedules to achieve sustained seizure freedom*" [35]. Patients suffering from this condition should be further evaluated in epilepsy centres by a team of specialised multidisciplinary experts.

2.1.4.2 Surgery

In these centres, it is assessed whether the patient is a candidate for resective surgery in case of medication failure. Surgery is the most effective strategy to manage

seizures in drug-resistant focal epilepsies by resecting the epileptogenic zone, which is the area of the brain responsible for causing seizures. This surgery is not appropriate for all patients. Its success depends on identifying the epileptogenic zone, which in turn, must be kept small. Patients are evaluated before surgery to determine whether or not the procedure can be performed [15, 36, 37].

The presurgical monitoring has two primary goals, the precise localization, and delineation of the epileptogenic zone coverage area and assessing whether its removal can be achieved safely without significant functional impairments [15]. Patients are subjected to sleep deprivation and AED withdrawal to increase seizure occurrence and shorten hospital stay [37, 38]. In those conditions clinicians presume seeing seizures with similar onset traits. To localise and delineate, clinicians need to perform a multimodality approach where each contributes with unique and complementary information. When the data from all modalities are combined, it is possible to generate a hypothesis concerning the epileptogenic zone. Clinical history, long-term video-EEG recordings, high-resolution MRI, and neuropsychological evaluation are the primary modalities for presurgical monitoring. Since these tests are already needed for surgery, their data is used in several studies and in this thesis [16].

2.1.4.3 Neurostimulation

After establishing a patient's ineligibility for surgery, neurostimulation can be suggested. It entails implanting a device that sends electrical pulses to peripheral nerves or specific brain parts of the central nervous system to prevent seizures. This strategy is palliative because only a small percentage of patients achieve seizure-free status for more than a year [3, 39–42].

Current neurostimulation treatments are classified as invasive or noninvasive based on the need for surgery for placement. They can also be classified as open and closed-loop when a scheduled or responsive intervention is pondered accordingly. The most common invasive techniques are Vagus Nerve Stimulation (VNS), Deep Brain Stimulation (DBS), and Responsive Neurostimulation System (RNS). Transcutaneous vagus nerve stimulation, trigeminal nerve stimulation, and transcranial magnetic stimulation are noninvasive methods [3, 39, 40, 42]. These, however, have not been clinically validated. Figure 2.3, which provides an intuitive view of the neurostimulation brain targets and key anatomical pathways, can help better understand howVagus Nerve Stimulation, Deep Brain Stimulation, and RNS systems function.

2.1.4.4 Rescue medication

Rescue medication is essential in epilepsy since it can: i) offer seizure freedom when paired with AEDs; ii) reduce the constant dosage of AEDs, hence reducing their long-term side effects; and iii) terminate seizure clusters and extended seizures [43–

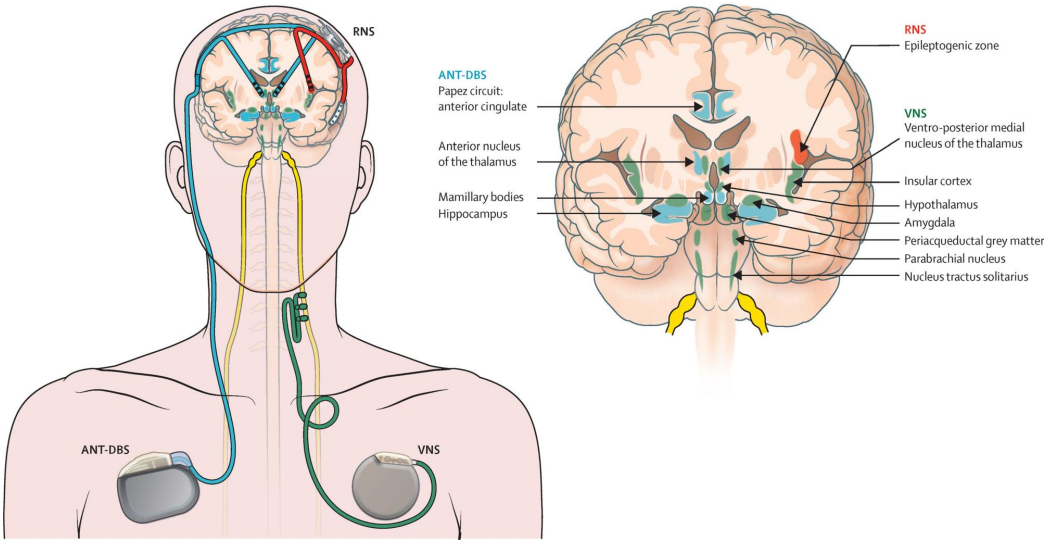

Figure 2.3: Approved neurostimulation therapies in epilepsy, also showing the brain targets for each neuromodulation approach according to sites of stimulation and known primary anatomical pathways. Source: Ryvlin et al. [3].

45]. Because of their quick action, *benzodiazepines* are commonly used as epilepsy rescue therapy. Due to their long-term adverse side effects, intense addiction, and habituation, they should only be used as an emergency treatment [46]. Table 2.1 lists the various emergency drug options, their route, peak effect level, and time to take effect.

Table 2.1: Outside-of-the-hospital treatment options for rescue drugs.

Drug	Route	Time to take effect	Peak level	Approval
Diazepam	Rectal	5-10 minutes	10-45 minutes	FDA in 1997 for seizure clusters
Midazolam	Buccal	<5 minutes	20-30 minutes	European Union in 2011 for prolonged seizures in non-adults
Midazolam	Intranasal	<10 minutes	15-120 minutes	FDA in 2019 for seizure clusters in patients older than 11 years
Diazepam	Intranasal	<5 minutes	>60 minutes	FDA in 2020 for seizure clusters in patients older than 5 years

2.1.4.5 Warning devices

Warning devices (see Figure 2.4) for epilepsy control, particularly seizure detection, have been studied. The ability to constantly monitor a biosignal and detect or predict a seizure promptly, followed by issuing an alarm, may give the patient or caretaker sufficient time to mitigate seizure repercussions or allow the administration of rescue medication [43, 47]. These deviceS are intended to incorporate algorithms that analyse long-term signals and raise alarms while rejecting data segments with artefacts. UNEEG SubQ, EpiMinder Subscalp, and Byteflies Sensor

Dots are among the latest EEG acquisition technologies presently available [48]. In addition to the EEG signal, researchers consider other noninvasive signals such as accelerometry, electrodermal activity, photoplethysmography, electromyography, body temperature, and electrocardiography for patient comfort [47,49,50]. This is achievable because of gadgets like the Empatica E4, Fitbit Charge HR, and Fitbit Inspire, among others [51]. The NeuroVista Seizure Advisory System (NCT01043406) is one of the most important seizure prediction warning devices. This device continuously monitors the brain using intracranial electrodes to record the EEG and gives the patients a likelihood of having a seizure [4].

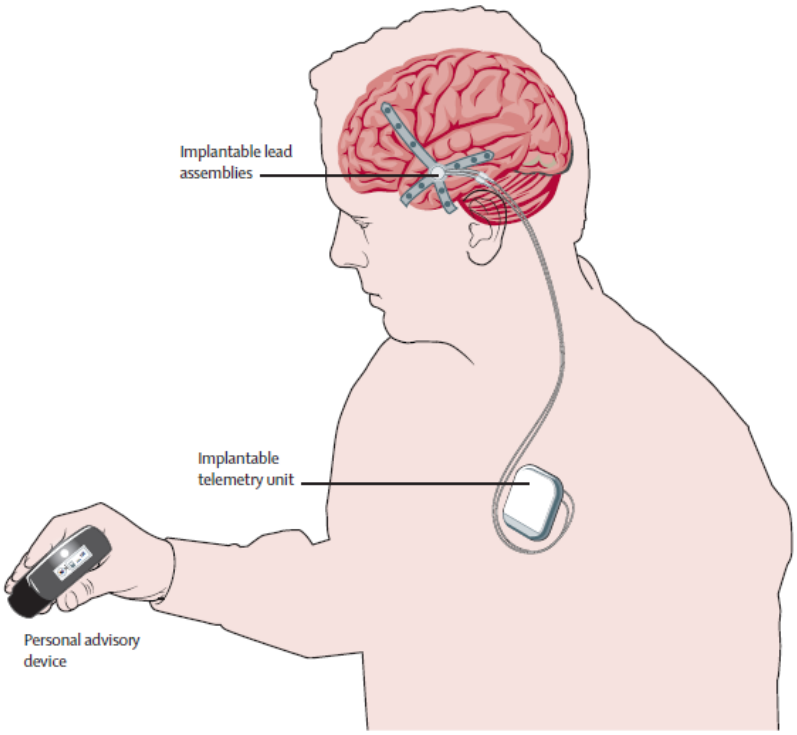

Figure 2.4: Major components of seizure advisory system. Source: Cook et al. [4].

2.2 EEG

The brain is a complex system composed of billions of interconnected neurons. The neurons work jointly to process and transmit information through electrical impulses [5,52].

The EEG provides a window into the brain by recording the time-evolving voltages generated by brain activity [5]. The activity recorded is a consequence of the electrical potentials arising from the total sum of excitatory and inhibitory postsynaptic potentials that are generated mainly by cortical pyramidal cells [53,54]. A synchronous neural activity involving thousands of neurons is considered necessary for detection with scalp EEG [5,53,54].

The EEG is the most efficient medical imaging tool to study and explain the

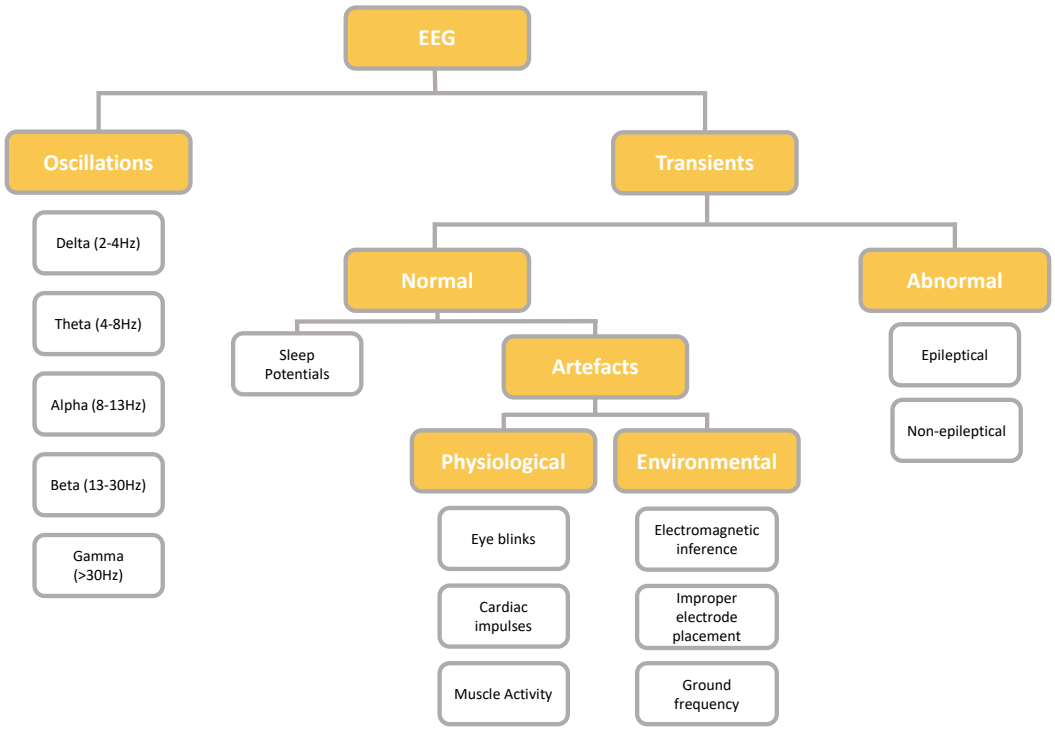

Figure 2.5: The EEG activity categorisation

essential qualities of a brain disorder, which assists the physician in making a proper diagnosis [52]. Usually, diagnosis is accomplished by a visual inspection of the EEG by the physician [52, 55]. The majority who have epilepsy disease possess EEG abnormalities (spike waves, sharp waves, spike-and-slow complex waves, sharp-and-slow complex waves, and other epileptiform EEG signals). These abnormalities are not only visible during a seizure. They are present throughout the patient's everyday life [55].

The EEG presents an oscillatory behaviour, although non-rhythmic shorter patterns can also be observed. Its potentials can be categorised into oscillations and transients (see Figure 2.5). The first are rhythmic patterns divided into different frequency sub-bands: delta (2-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30), gamma ($\geq$ 30Hz) [52, 53]. It is worth mentioning that a consensus among authors regarding these sub-bands cannot be found. These bands are connected to many human endeavours. Profound sleep is linked with delta oscillations, whereas drowsiness, inspiration, and deep concentration are associated with theta rhythms. The most prominent brain rhythm, alpha waves, are frequently seen across the occipital lobe. Beta oscillations often manifest during alert and agitated states, particularly in the frontal and central brain areas, in addition to mental and cognitive activities. Gamma oscillations are uncommon and typically obscured by muscular artefacts, especially in scalp EEG [52, 53].

EEG transients are sharp and are divided into normal and abnormal. Normal

transients include sleep potentials and artefacts [52, 53]. When changing from one state of awareness to another, sleep potentials can be found in everyone. They help perform the vigilance shift easily and successfully [52].

Artefacts are non-cerebral electrical potentials detected with EEG. Physiological artefacts are considered noise, such as eye blinks, cardiac impulses, and muscle activity [53]. Environmental artefacts, also known as non-physiological, can be electromagnetic interference from the surrounding environment, improper electrode placement, and the 50 or 60 Hz ground frequency [53, 56].

Regarding the abnormal EEG transients, these form the basis of epilepsy diagnosis and can be split into epileptical and non-epileptical [52, 53]. The non-epileptical can be indicators of various encephalopathies. [52, 53].

2.2.1 Acquisition

The EEG signal is acquired by placing several electrodes on the scalp (Scalp EEG) or inside the skull (Invasive Electroencephalogram (iEEG)). The number of electrodes and their location determines the signal spatial resolution spread, while the sampling frequency determines the time resolution.

Scalp EEG

Scalp EEG is an accessible technique of low cost and non-invasive (see Figure 2.6). The scalp EEG signals are acquired through electrodes placed on the subject's scalp. Given that the electrodes in this technique are placed on the scalp, cerebral activity potentials are only measured after going through cerebrospinal fluid, the skull, and the scalp. Commonly, the EEG signals are collected with the help of an electroconductive gel to reduce the exiting impedance. [5, 57, 58]

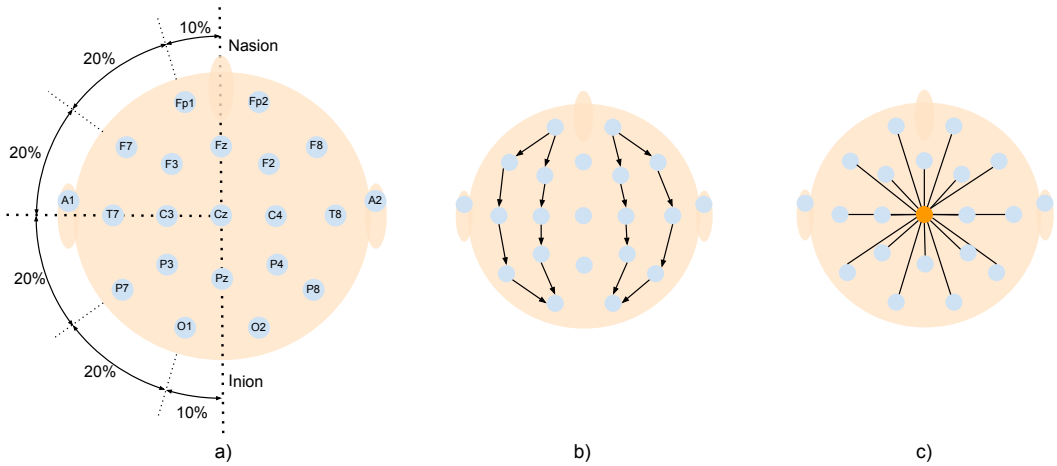

Figure 2.6: International 10-20 system for placement of scalp EEG electrodes. In (a) is shown the standard positions and names of the electrodes. In (b) is represented a bipolar montage and in (c) a referential montage. Adapted from: Varsavsky et al. [5]

Commonly, electrodes are placed on the patient's scalp according to the interna-

tional 10-20 system. In a referential montage, 19 recording electrodes, in addition to a ground (often placed on the forehead) and system reference (usually placed at the ear), are used [57, 59, 60]. Referential montages measure the voltage difference between the electrode itself and a reference. After the acquisition, the original signals from the referential montage can be kept, or the bipolar montage can be created. The bipolar montage is simply the difference between the voltages measured from adjacent electrodes [59, 60].

iEEG

For some patients, the seizure onset zone cannot be properly localised by using scalp electrodes. When this occurs, a more invasive alternative to measuring the cerebral activity potentials is used, the iEEG. It is performed using i) intracranial electrodes that record information directly from the brain or ii) subscalp electrodes, which are subcutaneously implanted between the scalp and the bone. There are three types of intracranial electrodes: subdural strips, subdural grids, and depth electrodes. These can only be placed after a craniotomy, where the subdural grids or strips are placed on the exposed surface of the brain (see in Figure 2.7.) and depth electrodes are stereotactically implanted (also known as stereotactic EEG), this way, allowing the recording of the electrocorticography [5, 53, 61, 62].

Similarly to scalp recordings, in the intracranial, more than one option for system reference can be used, another intracranial electrode, the average of all electrodes in the grid or strip, or an external electrode. Depth electrodes have shown promising results in detecting epileptiform activity in patients with TLE and, compared with subdural grids or strips, have minimal risks of complications. In contrast, brain bleeds and infection are possible complications with subdural grids or strips but do not frequently happen. [5, 53, 61, 62].

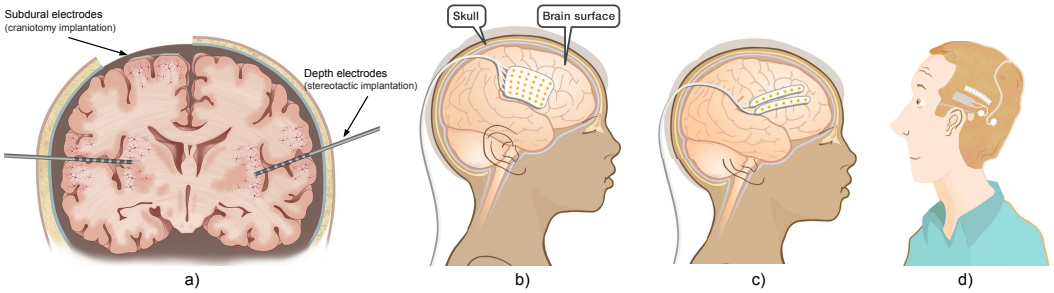

Figure 2.7: Several invasive EEG electrodes: examples of electrode placement for subdural and depth electrodes (a), a subdural electrode grid (b) and a strip one (c), and subscalp electrodes (d). Adapted from [63–65].

Scalp EEG and iEEG comparison

Intuitively, scalp EEG recordings have more artefacts, and a smaller signal-to-noise ratio as the electrical signals are more attenuated than iEEG recordings. Even

though the iEEG provides a clearer signal, one must consider the ethical reasons for this technique. Furthermore, with the risk of brain bleeding and infection, placing electrodes on a broader brain region than strictly necessary may not be the better option.

Invasive recordings have considerable limitations regarding coverage area, shown by the difficulty of capturing large-scale synchronous activity (e.g., alpha band). Therefore, even though the EEG has a higher spatial resolution, less synchronous activity over large regions (e.g., beta and gamma bands) could be compromised in the scalp EEG. On the other hand, scalp recordings could, in principle, be used in an ambulatory setting to monitor a patient's seizure situation without changes to everyday brain activity. The primary setback with this approach is the need for a high degree of compliance on the patient accepting the inconvenience of constantly wearing an EEG device. Anyhow, it is of particular interest that regardless of the acquisition type used, the EEG is a complex signal that envisions recording by approximating the existing linear and non-linear interactions among neurons [5, 20, 53, 57–62].

2.2.2 Seizure period division

With the EEG, it is possible to not only evaluate whether a patient has epilepsy but also possible to divide and characterise the different periods associated with ictogenesis and epileptogenesis, as seen in Figure 2.8: preictal, the period that precedes the seizure; ictal, the period corresponding to the seizure; postictal, the period that succeeds the seizure; and ultimately the interictal, the period found between the postictal and the preictal periods of consecutive seizures [17, 66].

The preictal is the most challenging period to annotate as there are no effective bio-markers for all patients. The lack of guidelines comes from the fact that a clinician cannot timely predict a seizure by visually inspecting the EEG, as well as the fact that heterogeneity can be found, not only from one patient to another but also between seizures within the same patient. On top of that, abnormal non-epileptiform activity increases the task's difficulty, and in some cases, a seizure can occur without, apparently, a specific preceding event caught in the EEG. Overcoming all of these challenges and performing a proper detection of the preictal period is the essence of epilepsy seizure prediction [19].

2.3 Seizure Prediction

The main goal associated with seizure prediction is to build an algorithm to predict seizures, not the exact time, but a well-defined window where the seizure is expected to occur. Algorithms should then work in a device that receives the EEG signal in real-time and releases alarms, giving the patient a minimal preparation time [12, 53].

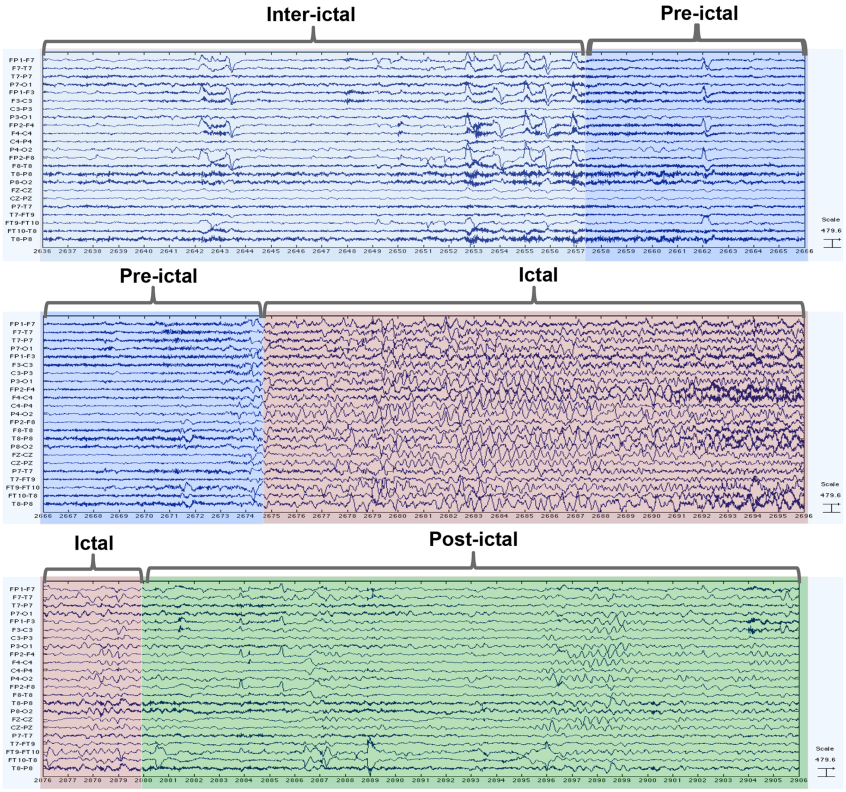

Figure 2.8: The different periods of an seizure episode annotated on the EEG signal. Source: Moghim et al. [6].

There are two types of seizure onsets, electrographic and clinical. The former is attributed to the moment where significant changes in the EEG are observed, while the latter is identified when the first symptoms start. If one considers the chronology of events, the electrographic onset precedes the clinical one. Therefore for seizure prediction, the electrographic is selected. Moreover, seizure symptoms are not always visible, notably in non-motor seizures [12, 20, 53].

Another vital aspect are lead seizures. Since seizures are usually clustered in time (e.g., 1-5 seizures occur in close succession), the built algorithms assume seizures as independent events. Consequently, all studies aim to predict the first seizure in a cluster (lead or leading seizure), as those are much harder to predict since, for the others, it is thought that the brain never truly left a state of excitability [67]. A 'lead seizure' is any seizure preceded by a seizure-free period of length T (or longer). As there is no consensus regarding this period's duration, studies have considered different values of T, of 1 hour [68], 1.5 hours [69], 4 hours [70], 4.5 hours [71], 5 hours [18], and 8 hours [4].

Seizure detection: a parallel research field

First and foremost, seizure prediction and detection are two distinct concepts. While the former focuses on locating the preictal period, the latter aims to identify the electrographic onset by finding the ictal period and detecting it before the first

symptoms arise. Even though seizure detection has better performance results, prediction is preferable, as it offers more preparation time than the 5 to 12 seconds that are accomplished with detection. Detection helps to find the seizure focus. It can also be used in a closed-loop system to trigger neurostimulation [6, 20].

Seizure prediction vs forecasting

Another concept often confused with seizure prediction is forecasting. In this context, they have different meanings. Seizure prediction anchors on locating the preictal period. Forecasting aims to combine the information of the fluctuating rates of interictal epileptiform discharges with circadian and patient-specific multidien (multi-day) cycles to find periods of higher seizure occurrence (proictal period) [72].

2.3.1 Characterisation

Several EEG seizure prediction methodologies had already been developed by the early 2000s after this scientific area sparked interest in the 1970s. However, the lack of a recognised criterion made proper evaluation and comparison difficult. Additionally, it was not easy to assess whether an algorithm's performance was adequate for clinical use [7, 20].

As a result, in 2003, Winterhalder et al. [7] proposed a general framework to assess and compare seizure prediction methods, known as the "seizure prediction characteristic", by accounting for clinical, behavioural, and statistical considerations. Given that the built algorithms, as previously said, cannot predict an exact point in time where the seizure is deemed to occur, two concepts to face this uncertainty were proposed: Seizure Prediction Horizon (SPH) and Seizure Occurrence Period (SOP). The SPH, also known as Intervention Time (IT), is the interval following the alarm that renders a possible intervention. The SOP is defined as the period where the seizure is expected to occur (see Figure 2.9).

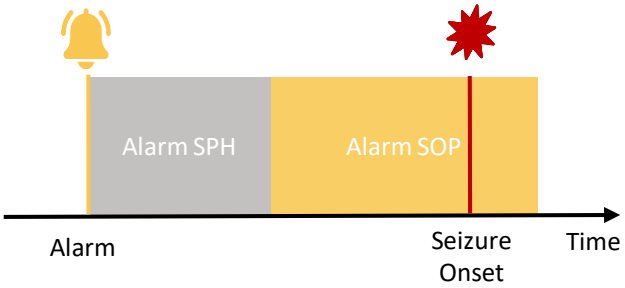

Figure 2.9: Practical definition of SPH and SOP. Adapted from [7].

Since preictal duration differs between patients and seizures, the optimal SOP and SPH have not yet been determined. If the chosen SOP is enormously long, the device might not be helpful. For example, suppose a patient has three seizures in a day and a SOP of eight hours is used. In that case, the algorithm will not

be helpful regardless of accurately anticipating every seizure. Therefore, the SPH and SOP must have a reasonable duration associated with the chosen intervention system. High SOP values may lead to anxiety when the goal is to warn patients about seizures. Continuous intervention is also difficult with long SOPs such as those for electrostimulation systems [7, 73].

2.3.2 Performance assessment

In this subsection, it is outlined how both types of performance in this study are assessed. Regardless, performance must be evaluated using previously unknown data to the trained models.

2.3.2.1 Sample performance

The performance of a seizure prediction system is achieved by working on the decisions of a binary machine learning model over a time series. Therefore, the temporal data chunks are classified into interictal (0) and preictal (1). The trained model is applied to the testing data, and after, the performance of the prediction system is characterised as described in 2.3.2.2.

Figure 2.10 shows the relation between the EEG clinical label and the algorithm's output.

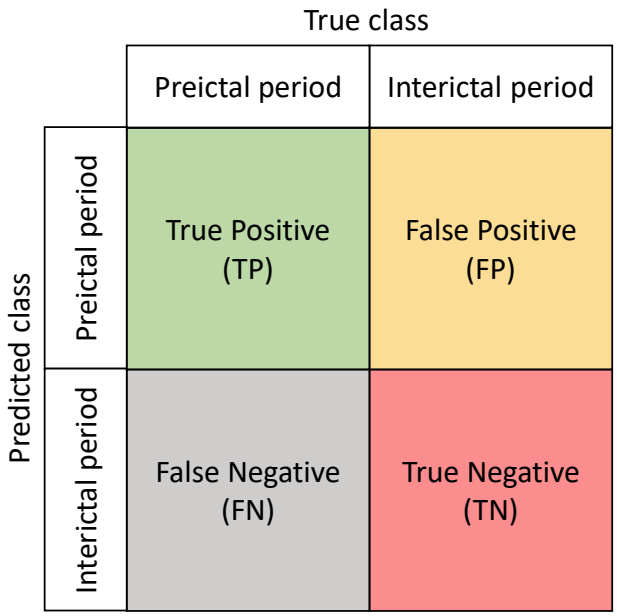

Figure 2.10: Confusion matrix for assessing sample seizure prediction performance.

In classical Machine Learning, classification is measured using the confusion matrix. For its evaluation, the following measures (2.1 and 2.1) can be used.

$$Sensitivity = \frac{TP}{TP + FN} \tag{2.1}$$

.

$$Specificity = \frac{TN}{TN + FP} \tag{2.2}$$

.

2.3.2.2 Performance of a seizure prediction system

The performance of predictive methods is evaluated by the sensitivity, and False Positive Rate per Hour (FPR/h) [7], to calculate them, it is necessary to distinguish a correctly generated alarm from an incorrect one [12].

As represented in Figure 2.11, an alarm is considered correct (true alarm) if the seizure occurs during the proposed SOP. Therefore, the alarm needs to be raised during the preictal period. An alarm is considered incorrect (false alarm) when raised during the interictal period or in the SPH. Consequently, the predicted seizure occurs outside the SOP [7, 53].

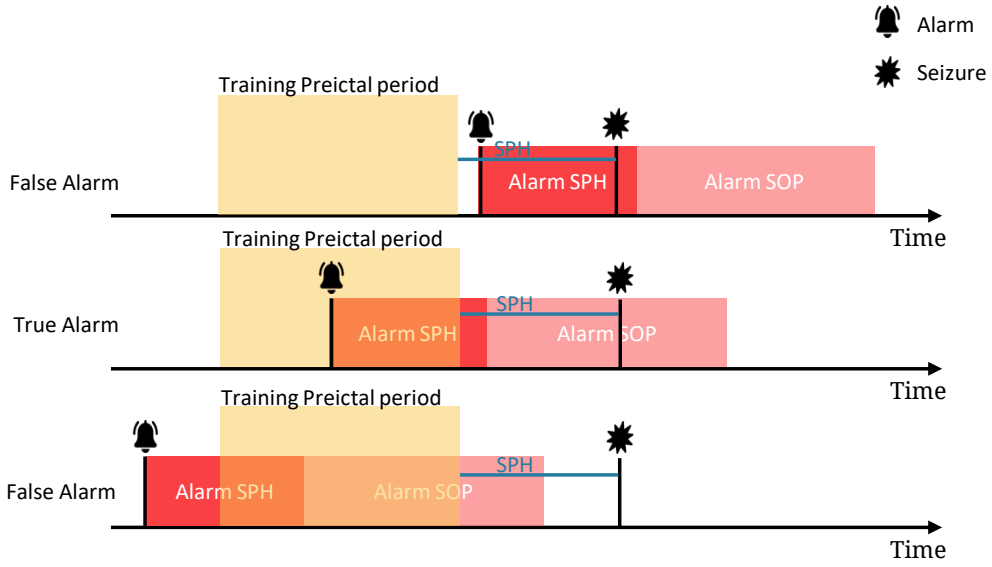

Figure 2.11: The definition of a true and false alarms, SPH and SOP in a practical view, adapted from [7].

It is important to note that authors often select an equal duration for SOP+SPH as the preictal period in supervised learning techniques, as shown in Figure 2.12.

Sensitivity measures the fraction of correctly predicted seizures, as described in 2.3:

$$Sensitivity = \frac{Predicted\ seizures}{All\ seizures} \tag{2.3}$$

One cannot measure specificity in seizure prediction as in a traditional machine

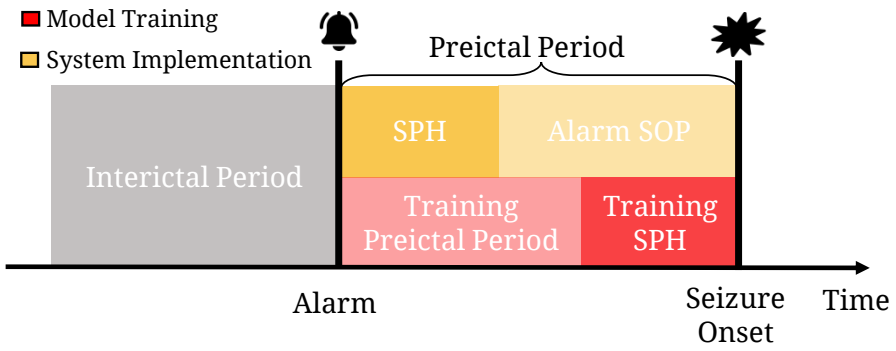

Figure 2.12: Visual representation of the relation between the preictal period, SOP, SPH. It is a true alarm and is raised in the first preictal sample.

learning problem. Therefore, a more appropriate measure, the FPR/h is used. Usually, FPR/h is depicted as the number of false alarms during the interictal period. Also, refractory periods (SPH+SOP) are often considered where consecutive alarms cannot be raised. The FPR/h should be the number of false alarms during the period when it is possible to raise an alarm [7, 12, 53]. Consequently, the refractory period of each alarm period is removed when calculating the FPR/h, as described in 2.4:

$$FPR/h = \frac{N_{False\ Alarms}}{Inter_{ictal}\ duration - N_{False\ Alarms} \times (SPH + SOP)} \tag{2.4}$$

Optimal metric values

If the algorithm is adjusted to increase the sensitivity, the FPR/h will also be affected, as represented in Figure 2.13. Therefore, the two parameters must be evaluated together due to their relationship [7]. If the algorithm releases too many false alarms, it makes patients ignore the device. Patients who take all warnings seriously can suffer from substantial psychological stress or experience additional neuropsychological impairments due to the unnecessary use of anti-convulsive drugs or electrical stimulation devices [7, 73]. The average seizure incidence may indicate a reasonable range.

According to Winterhalder et al. [7], in presurgical monitoring, there is an artificially high seizure frequency, as there is a reduction of anti-convulsive medication. In those situations, a maximum FPR/h of 0.15 (3.6 seizures per day) was reported.

Under normal conditions, patients with pharmacorefractory focal epilepsy have a mean seizure frequency of about three seizures per month, meaning 0.0042 seizures per hour. Therefore, if one maintained a maximum FPR/h of 0.15 in a month analysis while correctly anticipating every seizure, 97% of the alarms would be false in contrast to the 50% during presurgical monitoring. Thus, a maximum FPR/h of 0.15 is considered a reasonable value only during presurgical monitoring.

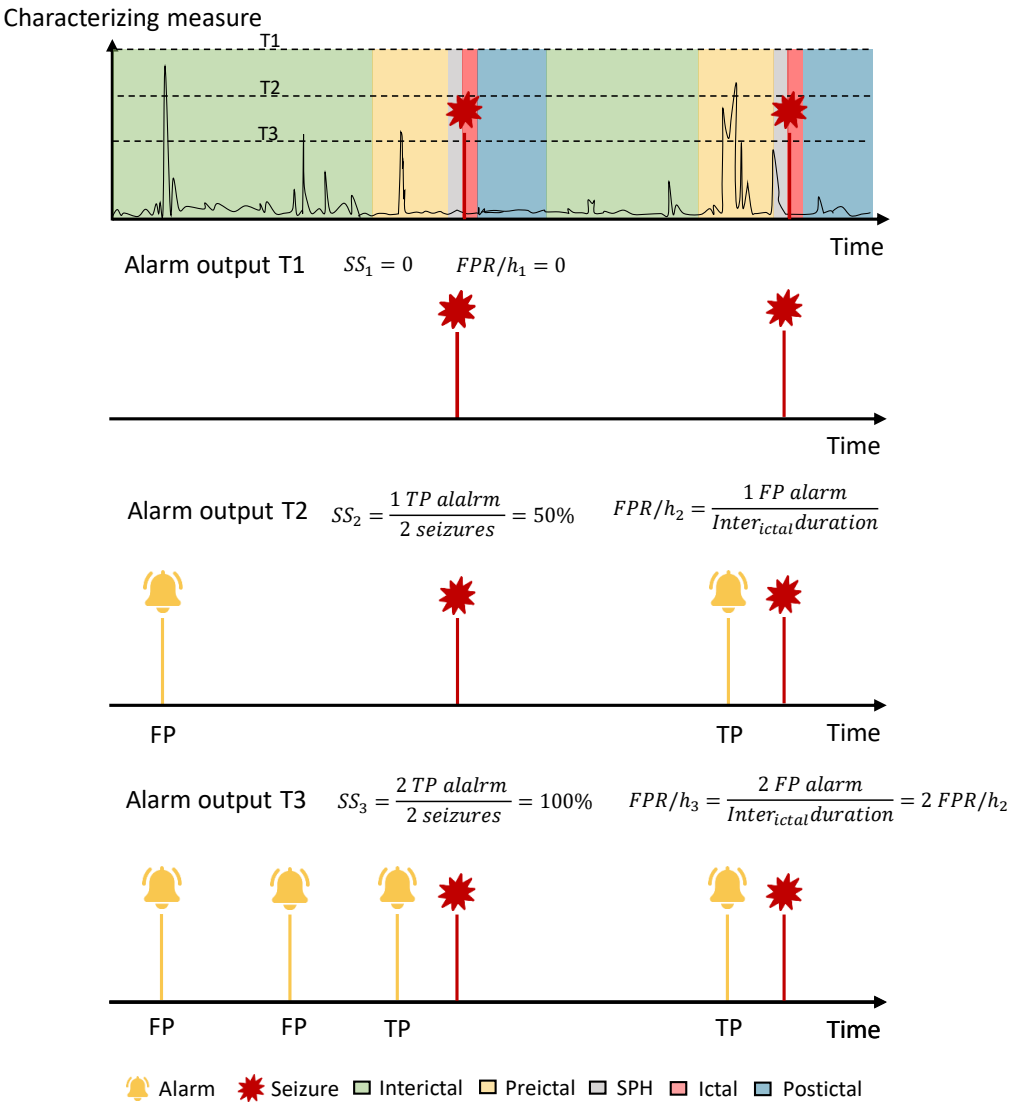

Figure 2.13: Upward crossing of a threshold triggers an alarm. The dependence between sensitivity and false prediction rate is shown by three distinct thresholds (dashed lines): For T1, there is no preictal or interictal alarm; hence there is also no sensitivity and no false predictions. At the expense of one inaccurate prediction made during the interictal epoch, threshold T2 accurately predicts the second seizure. Another false alarm is produced when the threshold is lowered to T3. Therefore, while adjusting the seizure prediction method's parameters, disregarding the false prediction rate may result in high sensitivity. To evaluate a prediction approach, the sensitivity and false prediction rate must be calculated together. Adapted from [7].

2.3.3 Statistical validation

Another important aspect of seizure prediction is statistical validation to determine whether a proposed algorithm is better than the chance level; statistical validation must be implemented [12,20,74]. The most used techniques are the random predictor [8,74] and the surrogate time series analysis [8,75,76].

2.3.3.1 Random predictors

Unspecific prediction methods assume that alarms are released without using any information from the EEG [7, 8, 74].

Based on a homogeneous Poisson process for false predictions, Schelter et al. [7, 8, 74] suggested an analytic random predictor. The probability of releasing an alarm at any single sampling point of the time series with N samples is described by (2.5):

$$P_{Poisson} = \frac{N_{false\ alarms}}{N_{samples}} \tag{2.5}$$

Considering a time interval equal to SOP, the probability of raising at least an alarm in that interval is given by 2.6. This approximation is only valid if the product between FPR/h and SOP is considerably lower than one, which is reasonable to ensure that the patient is not continuously under warming.

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

The probability above represents the sensitivity of the random predictor. Therefore, it forms the basis for testing whether the prediction method's sensitivity is higher than a random predictor's. It constitutes the probability of raising at least one alarm during the SOP.

The probability of randomly predicting k of K seizure follows a binomial distribution, with probability P and several predictors (d). Also, if one considers several electrodes and multiple features, the probability is described by 2.7:

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

Ultimately, for a given significance level α, it is possible to determine the critical value of sensitivity, which the algorithm should outperform 2.8:

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

To summarise, the random predictor has an advantage in that its analytic expression does not require the EEG input and is computationally light. Furthermore, defeating the random predictor may be difficult for a small number of evaluated seizures, which may be a concern because seizures are rare events and the models are patient-specific.

2.3.3.2 Surrogate analysis

Andrzejak et al. [75] suggested the surrogate time series analysis as a statistical validation method. It randomly shuffles the original seizure onset times to generate

artificial ones, as illustrated in Figure: 2.14. Suppose the proposed predictor performance is higher for the original seizure onset times than for the surrogate ones. It can be considered that the proposed algorithm outperforms the random predictor. It can be difficult to generate such a large number of independent surrogates for significant analysis when seizures are frequent, and gaps in EEG records are present [8, 12, 17, 75, 76].

2.3.3.3 Comparison

Despite its ability to calculate critical values almost immediately, the analytical random predictor does not consider the SPH value. Consequently, a 10 seconds or 1 minute SPH does not affect its value, but a 1 minute SPH changes the problem's difficulty. While the surrogate predictor has the advantage of being more flexible and adaptable, offers more confidence, and can account for the non-random occurrence of seizures. However, this methodology is computationally heavier, and to avoid biasing the null hypotheses, its implementation must be performed with care. Some authors consider the surrogate predictor to be more robust [8,17] and for this thesis was the chosen strategy for statistical validation.

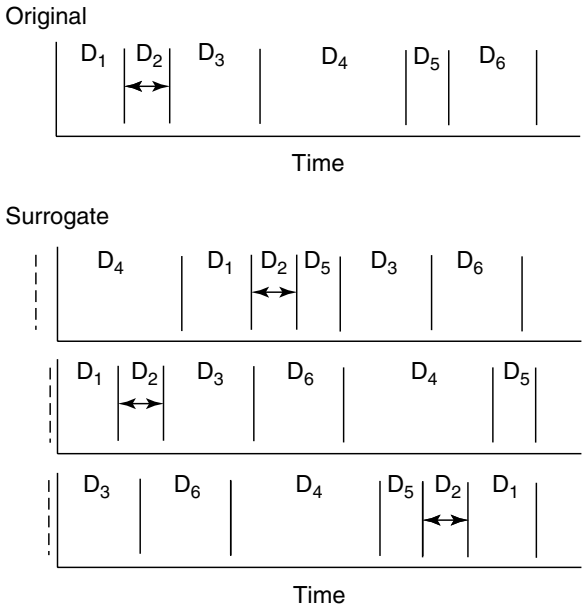

Figure 2.14: Original seizure times and the surrogate times bootstrapped from the inter-seizure intervals. The random onset times for the surrogates are obtained from a uniform distribution and are indicated by the dashed vertical lines. Note that by randomly selecting the offset of the starting point (compared to the original one), the endpoint is different from the original one for all surrogates. Source: Schelter et al. [8].

2.4 Challenges of learning from time series

Machine learning's primary goal has traditionally been to learn from data assumed to be sufficient and representative of the underlying fixed, yet unknown, distribution. Real-world problems, such as the case of EEG seizure prediction, rarely fit those assumptions. For example, class distributions are often skewed, which leads to a class imbalance issue. Often, the data used is acquired from a non-stationary distribution where changes in the hidden context or data distribution may occur, leading to concept drift. When a learner does not account for these changes, a decrease in performance usually occurs [9].

Regarding seizure prediction, the underlying context includes daily-life habits, medication, stress situations, the circadian, ultradian, and infradian rhythms, cognitive states, environmental changes, implanting a neurostimulation device, a sudden brain lesion, and others that can affect the brain dynamics and consequently modify optimal characteristics for anticipating seizures. Additionally, in presurgical monitoring, the alteration of medication and sleep deprivation are also changes expected in this document, as they are used to provoke seizures [17–22].

The number of seizures affecting a patient can vary from 3.6 a day (in presurgical monitoring) to 3 a month, regulated by anti-seizure medication [7]. That being the case, the EEG should be recorded for several weeks up to months. This way, hopefully, all types of concept drifts should be caught.

As previously mentioned, another relevant issue concerning classification algorithms is class imbalance. Traditional classification problems assume that the prevalence of each class will remain equivalent, which does not happen, especially in streaming data applications. In seizure prediction, where seizures are relatively rare events, the interictal period is substantially more extended than the preictal period. Therefore, a learning system must avoid specialisation over the interictal state by dealing with the class imbalance. In addition, it must stay stable to detect irrelevant events (outliers) and remain capable of adapting to changes in the hidden context [9]. At last, possible solutions to build models capable of withstanding these changes are: adaptive learners, modifications to the training set, ensemble techniques, the inclusion of exogenous variables, and many more further explored in sections 3.2 and 3.3 [9].

More about Concept Drifts

A concept drift can be classified as either real or virtual (as seen in Figure: 2.15). A real concept drift is defined as a change in the class boundary. Alternatively, in a virtual concept drift, the distribution of instances may change (the frequency of each class or the distribution of the classes), but the underlying concept remains the same. Even though only changes that occur in a real concept drift indicate a

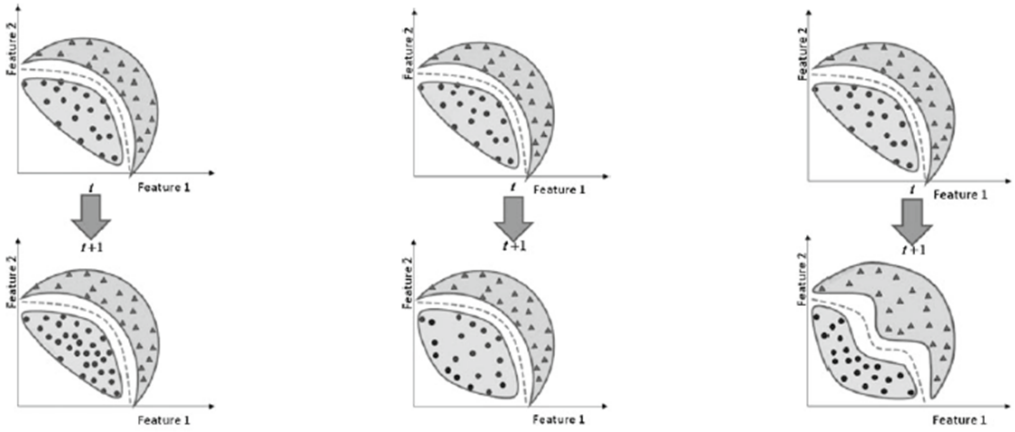

Figure 2.15: Graphical representation of the three different drift types. In (a), the circle class occurs more frequently after the drift, and in (b), the distribution of the circle class changes. While in (c), the boundary separating both classes changes. (a) and (b) are virtual concept drifts, while (c) is a real concept drift. Source: Hoens et al. [9].

change in the function generated by the underlying context, a distinction between both types of drifts, in practice, is not made because both require a change or re-fit of the model [9].

One can further classify a concept drift based on its speed, as seen in Figure 2.16. Dealing with sudden or reoccurring (if the transition is also sudden) concepts is deemed easier, given that a clear boundary, a definite point in time where the change in context occurs, exists [9, 10].

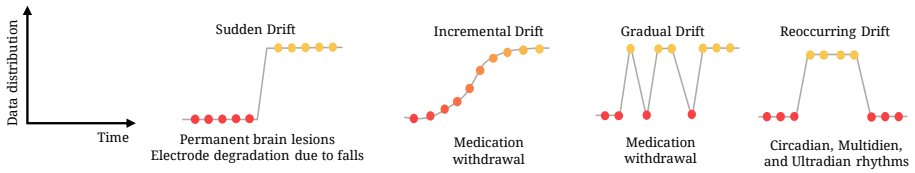

Figure 2.16: Graphical representation of the various speeds of concept drifts and their possible translation for the specific case of epilepsy seizure prediction, including presurgical monitoring conditions. Adapted from: Zliobaite [10] and Gama et al. [11].

Ultimately, when the factors mentioned earlier are not adequately accounted for, the concept drifts negatively influence the learner as confounding variables in the data stream. Their proper inclusion during the learning process could lead to better results [20].

2.5 Summary

A substantial heterogeneity characterises epilepsy as for seizures, types of epilepsy, and epilepsy syndromes. Several aspects can describe a seizure: its initial manifestations/symptoms, awareness, vigilance state at the onset moment, and epileptic focus

localisation in lobes and/or hemispheres. The most common syndrome in adults is the TLE, defined by seizure with a focus on the temporal lobe. The emphasis of seizure prediction is DRE patients, who cannot achieve seizure freedom through medication, as they are exposed to the physical and social implications of the unpredictability of seizures. This group of patients is frequently monitored for weeks or months to assess their condition before undergoing surgical interventions, hence why the majority of databases are made up of data collected during this period.

The EEG provides a window into the brain by recording its electrical activity, thus representing the physician's most efficient tool, even though its morphology is not entirely understood. Its potentials can be categorised into two types of phenomena: oscillations and transients. The first are rhythmic patterns divided into different frequency bands, while the second are sharp and can be categorised into normal and abnormal. Normal transients englobe eye blinks, cardiac impulses, and other physiological body functions, while the abnormal may or may not be related to epilepsy. There are two acquisition techniques: scalp EEG and iEEG. Although non-invasive EEG recordings have a higher spatial resolution and can capture low-frequency activity over more significant regions, iEEG recordings have fewer artefacts and a higher signal-to-noise ratio. However, iEEG has a considerable risk of brain bleeds and infection. The EEG is a complex signal that envisions recording by approximating the existing interactions among neurons. A thorough examination of all EEG activity types is necessary to make reliable predictions, as not all epileptiform activity can predict seizures.

Authors categorise the EEG signal related to seizures into interictal, preictal, ictal, and postictal in seizure prediction. Seizure prediction seeks to identify the preictal period and anticipate seizures by promptly sounding an alarm. It can be difficult to locate the preictal period because it is a transitional stage that differs between patients and seizures. Each alarm has an occurrence period SOP and an intervention time SPH, their values must allow for proper intervention in real-life. An alarm system's evaluation should consider FPR/h and seizure sensitivity. Also, statistical validation should be conducted, where outperforming a random predictor is a minimum requirement.

The presented methodologies must deal with data imbalance and concept drifts, just like any rare event prediction task within a time series. Sleep deprivation and medication tapering are the most frequent presurgical monitoring concept drifts.

Chapter 3

State of the art

This chapter briefly explains the current state of the art on Electroencephalogram (EEG) seizure prediction and concept drift adaptation. Section 3.1 presents the common framework along with the traditionally used techniques and features. Section 3.2 presents current strategies to deal with concept drifts in seizure prediction. Section 3.3 presents approaches to learn from streaming data with concept drift. Finally, section 3.4 provides a summary of the state-of-the-art key concepts.

3.1 Common Framework

3.1.1 Overview

One of the most incapacitating aspects of epilepsy disease is the rapid and seemingly random nature of seizures [12, 17, 20]. Since a method capable of predicting the occurrence of seizures would open new therapeutic possibilities, investigations on the predictability of seizures have advanced since the 1970s [17, 20]. Early works were carried out with linear approaches, such as autoregressive models. Then, studies suggesting the possibility of preictal phenomena started emerging. The latter generally based on nonlinear dynamics [12,20]. Later, proof-of-principle studies compared preictal changes in dynamics to interictal control recordings [20]. Around 2003, although those early findings were optimistic, the absence of statistical validation and reproducibility started to be discussed [12]. That led to a phase that Mormann et al. [20] called "the rise of scepticism", during which research with large databases showed worse performance than earlier ones.

Framework

As shown in Figure 3.1, current seizure prediction algorithms follow a common framework that includes signal acquisition, signal processing, feature extraction, feature selection, classification, regularisation, and performance evaluation. Follow-

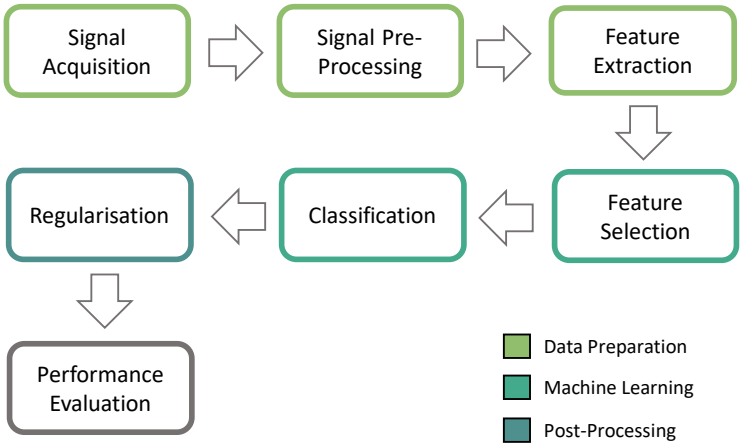

Figure 3.1: Flowchart of a typical seizure prediction pipeline. Adapted from Bou Assi et al. 2018 [12]

ing the collection of the EEG recordings, the following processes can be outlined for each:

- Signal Pre-processing: to enhance the quality of the EEG and/or to extract signal information through sliding time-window analysis.
- Feature extraction and selection: assembling features from the time series data and choosing those that can better distinguish between the various epileptic states.
- Classification: training machine learning models using the previously selected features to identify periods as either interictal or preictal.
- Regularisation: to smooth the classification output with post-processing methods.
- Performance evaluation: assessment of the performance and significance of the results according to appropriate metrics.

The variability of current approaches can be explained by the fact that there are a significant number of choices, despite the existence of a common framework. That also results from the lack of a gold-standard algorithm.

Briefly, the authors select a preictal temporal threshold and divide the intervals into distinct, independent inter/preictal blocks. They carry out a typical machine learning approach through feature extraction, feature selection, dimensionality reduction, and classification. Then, to emulate the real-time seizure prediction job, the generated model is applied to the continuous EEG, where post-processing can be carried out by including a moving average filter. The framework used in this thesis will follow the pipeline presented in Figure 3.1.

Deep Learning approaches

More sophisticated machine learning models, known as "Deep Learning", have recently emerged due to the rise in computational power and the amount of data available. Seizure prediction is just one of the many fields where these have advanced to the state of the art [77].

These models are representation learning techniques that may automatically identify the representations required for detection or classification tasks given a set of raw data. Complex functions can be learned using many layers of representation [77].

In the context of the general seizure prediction framework, the feature extraction, feature selection, dimensionality reduction, and classification stages are where these techniques primarily introduce modifications.

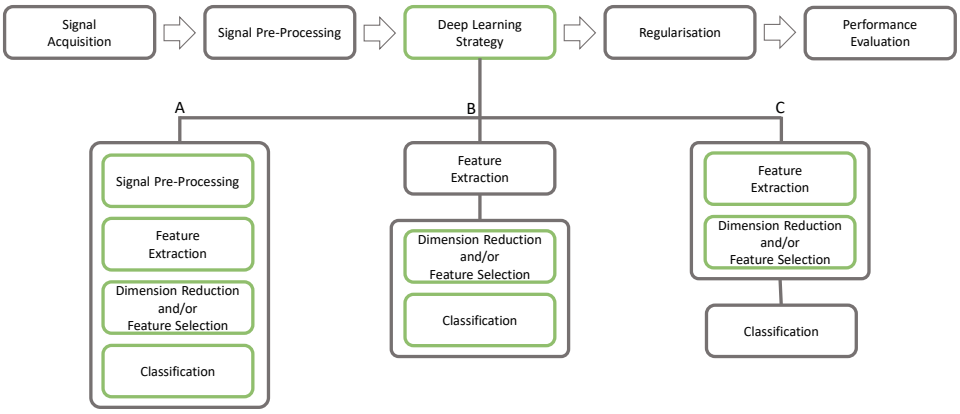

Figure 3.2: Flowchart of current Deep Learning (DL) pipelines for predicting EEG seizures. While a DL algorithm is capable of performing signal processing, feature engineering, and classification automatically (option A), some authors work with features as input (B) or even employing DL as a technique for feature engineering (C). The Green boxes denote the ones fulfilled by a DL model.

The simplest option (A in Figure 3.2) is to give the selected DL model the raw input (which may have a low processing level) and to extract the classification output using sequence analysis. In this instance, the model is in charge of feature engineering, classification, and optional signal processing procedures like artefact removal or noise reduction.

Some authors perform feature extraction beforehand and feed the derived measures to the models rather than using raw data as input (B, in Figure 3.2). The model will likely prioritise feature selection, dimensionality reduction, and classification in this instance. Last but not least, some authors use these models as feature engineering by extracting the acquired coefficients and feeding them to another classifier (C, in Figure 3.2).

3.1.2 Signal Acquisition

As the EEG is a remarkably complex signal, the choice of the database can significantly impact the findings of various studies [12]. Early research relied on data gathered from patients undergoing evaluation for epilepsy surgery from local databases [12]. These recordings, however, were short and interrupted, which reduced the number of ictal periods. For this reason, it was often difficult, if not impossible, to provide reasonable estimates of specificity for interictal periods. In recent studies, this limitation was addressed with the increase in the total number of hours of EEG and the number of recorded seizures. However, many studies with longer recordings are still with data from patients undergoing pre-surgical monitoring [12, 25].

There are now a number of databases available, with the University of Freiburg [6, 78], Children's Hospital Boston (CHB-MIT) [79–81] and European Database on Epilepsy (Epilepsiae) [69, 82–88] being the most commonly used.

As for the length of recordings per patient, the NeuroVista database curated by Cook et al. [4] is the greatest. Only data from 15 patients were available; however, they were all monitored for up to two years outside of monitoring units, representing data in real-life conditions. However, this kind of long-term monitoring is challenging to conduct because of logistical and ethical concerns.

Table 3.1 contains the signal acquisition parameters for some of the studies in the last 10 years. Notably, this does not necessarily imply that it can represent reality for the entire collection of studies on EEG seizure prediction. With the exception of Teixeira et al. [86], Direito et al. [83] and Pinto et al. [89] that performed studies with more than 90 patients, no other study included more than 53 patients. Also, not all databases are exclusively comprised of epileptic human patients, as some include dogs [90, 91].

Signal type

Concerning the EEG type, both the scalp and Invasive Electroencephalogram (iEEG) have been widely used. As previously mentioned in section 2.2.1 each signal as is advantages over the other, but regarding seizure prediction, various studies [84, 86, 87] show that the two types of EEG have no significant differences in performance [12].

In 2017, Assi et al. [12] deemed the iEEG more suitable for use in clinical prediction devices. However, recently, new studies using Subcutaneous EEG [92, 93], as well as Blood Volume Pulse, Accelerometry, Electrodermal Activity, and Temperature signals acquired from smartwatches or wrist bands [68, 70], have emerged. These signals may reduce the use risks and the discomfort from the iEEG and scalp EEG.

Table 3.1: A signal acquisition characteristics overview of EEG seizure prediction studies in the past 10 years.

Study	Database	Patients	Aggregated analysed time	No. of Seizures	Signal	Electrodes
Viana et al. [92] (2022)	ZUH KCL's clinical trial	6	594 days	82	Subcutaneous EEG	-
Pal Attia et al. [93] (2022)	ZUH KCL's clinical trial	6	409 days	N.A.	Subcutaneous EEG	-
Pinto et al. [89] (2022)	EPILEPSIAE	93	3687t h	238t	Scalp EEG	All
Stirling et al. [68] (2021)	Personal	11	13.5 years	1493	BVP Sleep stages	Smartwatch
Nasseri et al. [70] (2021)	NeuroPace	6	4 years	278	ACC BVP, EDA TEMP	Wrist-worn band
Tamanna et al. [79] (2021)	CHB-MIT	6	N.A.	40	Scalp EEG	All
Usman et al. [80] (2021)	CHB-MIT	23	27 days	198	Scalp EEG	All
Pinto et al. [82] (2021)	EPILEPSIAE	19	710t h	49t	Scalp EEG	All
Tsiouris et al. [81] (2018)	CHB-MIT	12	40 days	185	Scalp EEG	All
Agarwal et al. [94] (2018)	Kaggle (U. Pensylvania and Mayo Clinic)	12	N.A.	N.A.	Scalp EEG	All
Chamseddine et al. [90] (2018)	Kaggle (American Epilepsy Society)	1 dog	85 h	N.A.	Scalp EEG	16
Kiral-Kornek et al. [95] (2018)	NeuroVista	15	16.29 years	2817	iEEG	16
Karoly et al. [18] (2017)	NeuroVista	9	10.35 years	1458	iEEG	16
Khan et al. [96] (2017)	MSSM, CHB-MIT	50	N.A.	N.A.	Scalp EEG	All
Aarabi et al. [78] (2017)	FSPEEG	10	242 h	38	iEEG	3 in focal region and 3 outside focal region
Direito et al. [83] (2017)	EPILEPSIAE	216	697t days	1206t	Scalp EEG, iEEG	3 methods: F7, FZ, F8, T5, PZ, T6 6 random 3 in focal region and 3 outside focal region
Assi et al. [91] (2015)	Kaggle (AES)	5 dogs	N.A.	44	iEEG	16
Bandarabadi et al. [84] (2015)	EPILEPSIAE	24	150t days	183t	Scalp EEG, iEEG	3 in focal region and 3 outside focal region
Rasekhi et al. [85] (2015)	EPILEPSIAE	10	58 days	86	Scalp EEG, iEEG	3 in focal region and 3 outside focal region
Moghim et al. [6] (2014)	FSPEEG	21	24 days	N.A.	Scalp EEG, iEEG	3 in focal region and 3 outside focal region
Teixeira et al. [86] (2014)	EPILEPSIAE	278	2031 days	2702	Scalp EEG, iEEG	3 methods: 6 random F7, Fz, F8, T5, Pz, T6 3 in focal region and 3 outside focal region
Alvarado-Rojas et al. [69] (2014)	EPILEPSIAE	53	531 days	558	iEEG	All
Rasekhi et al. [87] (2013)	EPILEPSIAE	10	31t days	46t	Scalp EEG, iEEG	3 in focal region and 3 outside focal region
Rabbi et al. [88] (2013)	EPILEPSIAE	1	1.5 days	7	iEEG	2 in focal region

AES satnds for American Epilepsy Society, FSPEEG stands for Freiburg Seizure Prediction EEG, MSSM stands for Mount Sinai Epilepsy Centre at the Mount Sinai Hospital and CHB for Children's Hospital Boston. ZUH for Zealand University Hospital, and KCL for King's College London. "t" stands for testing data. BVP, ACC, EDA, and TEMP stand for blood volume pulse, accelerometry, electrodermal activity, and temperature.

Electrode selection

Electrode selection is another detail pointed in Table 3.1 where authors have implemented different methods. While some focus only on electrodes placed on the focal region [83, 86, 88], others choose six electrodes from which three are placed on the focal region and the remaining far from it [6, 78, 83–87]. Others use all available electrodes [18, 69, 79–82, 89–91, 94–96] or try to maximise scalp coverage with a smaller number of electrodes, selecting them (e.g. F7, FZ, F8, T5, PZ, T6) [83, 86]. These options arise from different assumptions that are worth being looked into.

For example, the idea is that by using only electrodes from focal regions, the seizure-generation process can be sufficiently captured by the focal region's activity. By choosing three electrodes placed on the focal area and three far from it, the authors assume that it is necessary to relate information from the focal to other brain regions without the need to use all available electrodes. When all or a random choice of electrodes are used, the assumption is that the seizure-generating process can be captured in any brain location. At the same time, it is logical to use all available electrodes since it provides more information; however, the computational cost increases. Until now, no assumption has proven to be more effective [87].

3.1.3 Signal Pre-processing

The recommended solutions must consider their viability in the actual world, given that the primary objective is to design a tool to receive online data and analyse it in real time. Therefore, the first step is usually data segmentation by a sliding window. Later, further optional steps like denoising, filtering, artefact removal and/or decomposition may be used to improve signal quality. Finally, the definitions of the preictal period, Seizure Occurrence Period (SOP), and Seizure Prediction Horizon (SPH) are given to avoid being influenced by performance outcomes. Their range should be stated before the machine learning stage, even if they may be handled later during classification [12].

A general pipeline for the signal processing stage is shown in Figure 3.3.

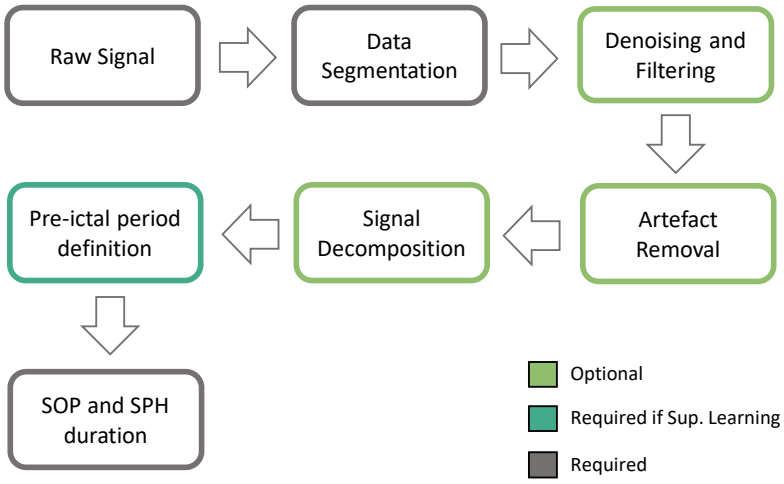

Figure 3.3: Flowchart of the typical signal processing pipeline in seizure prediction. Defining the preictal period is required when a supervised learning approach is used.

Overview

Table 3.2 provides a broad overview of the author's signal-processing choices. In summary, because the EEG is a complicated signal and challenging to comprehend, most researchers do not devote much attention to the denoising, filtering, and artefact removal steps. With sliding window analysis, it is feasible to identify similarities between studies over time and the percentage of overlap. Lastly, authors tend to differ regarding how long the preictal stage lasts. In most studies, SPH duration is omitted. As a result, in these situations, SPH is thought to be minimal, 5 to 10 seconds, which is an unrealistic scenario for warning deices, except for for closed-loop systems.

Table 3.2: A signal pre-processing characteristics overview of EEG seizure prediction studies in the past 10 years.

Study	Filters	Sliding Window Length	Preictal	SPH
Viana et al. [92] (2022)	0.5-48Hz band-pass and 25Hz low-pass filters 40dB attenuation filter	60s (0% overlap)	60 min	N.A.
Pal Attia et al. [93] (2022)	0.5-48Hz band-pass filter 40dB attenuation filter	60s (0% overlap)	60 min	5 min
Pinto et al. [89] (2022)	50Hz notch 0.5Hz highpass	5s (0% overlap)	30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90 min	N.A.
Stirling et al. [68] (2021)	Butterworth band-pass filter Hilbert transform	5s and 60s (0% overlap)	1 hour and 24 hours	N.A.
Nasseri et al. [70] (2021)	N.A.	20s (0% overlap)	1 hour	15 min
Tamanna et al. [79] (2021)	N.A.	10s (N.A overlap)	30	N.A.
Usman et al. [80] (2021)	Empirical Mode Decomposition	29s (0% overlap)	32 min	N.A.
Pinto et al. [82] (2021)	0.1 - 120Hz bandpass 50Hz notch	5s (0% overlap)	40, 50, 60 min	10 min
Tsiouris et al. [81] (2018)	N.A.	5s (0% overlap)	15, 30, 60, 120 min	N.A.
Agarwal et al. [94] (2018)	50Hz notch	1s (0% overlap)	10 min	N.A.
Chamseddine et al. [90] (2018)	0 - 190Hz bandpass and 60Hz notch	5s (0% overlap)	60 min	N.A.
Kiral-Kornek et al. [95] (2018)	Octave-wide digital notch filters 8Hz-128Hz	5s (0% overlap)	15 min	N.A.
Karoly et al. [18] (2017)	1 - 140Hz bandpass	60s (0% overlap)	30 min	1 min
Khan et al. [96] (2017)	0 - 128Hz bandpass	1s (0% overlap)	10 min	N.A.
Aarabi et al. [78] (2017)	50Hz notch 0.5 - 100Hz bandpass	10s (0% overlap)	30, 50 min	10s
Direito et al. [83] (2017)	8–52Hz bandpass	5s (0% overlap)	10, 20, 30, 40 min	10s
Assi et al. [91] (2015)	50Hz notch 0.5 - 180Hz band-pass	5s (0% overlap)	60 min	5s
Bandarabadi et al. [84] (2015)	50Hz notch	5s (0% overlap)	10, 20, 30, 40 min	N.A.
Rasekhi et al. [85] (2015)	50Hz notch	5s (0% overlap)	10, 20, 30, 40 min	N.A.
Moghim et al. [6] (2014)	Artefact removal	5s, 9s, 180s (0% overlap)	5 min	N.A.
Teixeira et al. [86] (2014)	50Hz notch	5s (0% overlap)	10, 20, 30, 40 min	N.A.
Alvarado-Rojas et al. [69] (2014)	N.A.	60s (0% overlap)	10, 30, 60 min	N.A.
Rasekhi et al. [87] (2013)	50Hz notch	5s (0% overlap)	10, 20, 30, 40 min	N.A.
Rabbi et al. [88] (2013)	0.5 - 100Hz bandpass and 60Hz notch	10s (50% overlap)	15, 30, 45 min	N.A.

Data segmentation

The EEG data is split and examined in brief windows to extract information chronologically to simulate an online time series situation. The selected studies' window lengths range from 1 to 180 seconds, while the commonest overlap percentage is 0 or 50%.

The trade-off between computational cost and execution speed largely determines the window length and overlap level. Many authors have decided for a 5-second window with no overlap because it is seen as a balance between detecting specific patterns and signal stationarity assumptions, while considering the number of electrodes, sampling frequency, and recording time [5, 12].

Denoising, filtering and artefact removal

Generally, this stage entails eliminating powerline interference 50Hz [78, 82, 84–87, 89, 91, 94] or 60Hz [88, 90], band-pass filtering, and abnormal transients regarded as artefacts. As other activity patterns outside EEG oscillations are eliminated, frequency decomposition into the frequency bands of interest or wavelet decomposition can also be considered filtering and artefact removal [6, 97].

Time-domain filters with Finite Impulse Response (FIR) and Infinite Impulse Response (IIR) have been employed extensively. IIR filters have demonstrated to not produce substantial ripple within EEG frequencies of relevance, in contrast to FIR ones that create a linear phase response and permit zero-phase distortion [12].

The band-pass filter cut-off frequencies used by authors vary. As a result,

they often eliminate high-frequency components that are considered noise and low-frequency components below 0.5Hz [78,88,89,91–93], considered breathing artefacts. The upper limit for high frequencies varies since their discriminative ability was investigated by numerous researchers.

Recently, Lopes et al. [71] proposed a DL strategy based on Convolutional Neural Networks (CNNs), developed with EPILEPSIAE database, for EEG pre-processing. The algorithm can reconstruct the EEG signal filtered, without noise and artefacts, by imitating experts' manual behaviour.

Definition of preictal period duration, SOP and SPH

A standard or optimal preictal time threshold and duration have not been defined. Authors have adopted fixed periods ranging from 5 [6] to 60 [90, 91] minutes in seizure prediction and up to 24 hours [68] in seizure forecasting, or experimented with different periods varying between 10 and 120 minutes [69, 78, 81–89].

As in the cases of Teixeira et al. [86], Bandarabadi et al. [84] and Pinto et. al [82,89], several studies have attempted to determine the ideal preictal value. For the first, the authors investigated four different preictal intervals and found a drop in False Positive Rate per Hour (FPR/h) with no discernible variations in sensitivity for longer periods. However, the regularisation approach could have had a role in this. The preictal period for the second study varied between 5 and 180 minutes, and the researchers concluded that the ideal preictal value varied between seizures and from patient to patient. For the third, a multi-objective evolutionary approach was used to perform high-level feature extraction as well as feature and parameter selection (namely, the chosen minimum preictal period).

The definition of the preictal duration is important when supervised learning methods are used, which is not the case for Leal et al. [98]. Using unsupervised techniques, the authors searched for preictal patterns on the EEG signal. It was shown that it is possible to identify seizure-specific preictal signatures for some patients and some seizures within the same patient.

Several authors only state the chosen training preictal period P_t used for supervised learning, which, as illustrated in figure 2.12, can be related to the duration of SOP, SPH and preictal period P:

$$SOP = P_t = P - SPH \tag{3.1}$$

There is also no standard value for SPH. Depending on the final application, this value may change. An SPH of a few minutes is sufficient to allow patient action in the case of a warning system. For an intervention system, a shorter time frame might be adequate. In the literature, SPH varies from a few seconds [78, 83, 91] to a few minutes [70, 82, 93].

Furthermore, in many studies [6, 68, 69, 79–81, 84–88, 90, 92, 94–96], the SPH

duration is omitted; thus, a standard SPH value must be assumed. The most secure approach is to consider the smallest value corresponding to a sliding window length since it theoretically optimises sensitivity, and 3.1 may be roughly translated to 3.2:

$$SOP \approx P_t \approx P \ for \ SPH \approx 0 \tag{3.2}$$

3.1.4 Feature Extraction

Feature extraction represents the step with the greatest degree of heterogeneity due to the wide range of methodologies that have been proposed. A significant range of features is divided into four major categories based on the linearity and number of channels employed (see Figure 3.4). The extracted features generally aim to capture three aspects that best describe seizure activity [12, 20]: i) An increase in energy (resulting from electrical discharges in the brain); ii) A shift in spectral content from low to high frequencies; iii) A rise in synchronisation in neural activity.

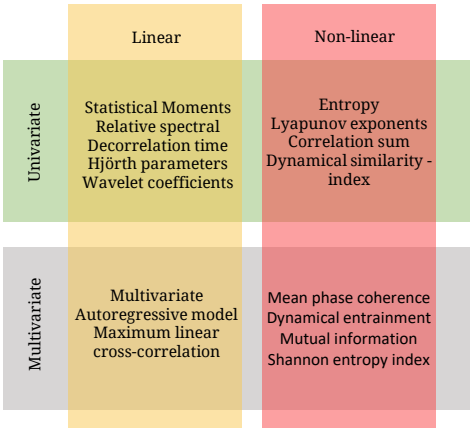

Figure 3.4: Categorisation of common features used in seizure Prediction.

Features are classified as either univariate (one single EEG channel) or multivariate depending on the number of used channels (several EEG channels). In the literature, the term ”bivariate”, part of the multivariate, is frequently used to describe features extracted as a combination of two channels. Multivariate measures provide more information and describe how the channels relate. Concerning linearity, features are classified as linear or non-linear.

With a glance at Table 3.3, it is clear that there is a predominance in the choice of univariate features, as well as a predominance of linear over the non-linear ones. That probably occurs as multivariate features need greater computational power, and linear measures are better understood and computed more quickly since they are lighter.

Not all studies compare the various types of features extracted for prediction purposes. Rasekhi et al. [85] pointed out that multivariate features produce fewer

incorrect predictions than univariate ones.

Additionally, it is worth noting that many recent studies that use DL techniques have opted to use time series data for automatic feature engineering [70, 81, 92–95]. Nevertheless, some still use DL classification models following traditional features extraction, typically with features based on frequency band or wavelet decomposition [68, 80, 86, 87, 90, 94–96]. Concerning interpretability, conventional features are preferable over automatically extracted ones. A more detailed characterisation of the most common features is depicted in Appendix A.

Table 3.3: Overview of the features adopted for some of the studies in the past 10 years.

		Univariate Linear features							Univariate Non-linear features								Multivariate Linear features			Multivariate Non-linear features				
Study	**Other**	**Statistical Moments**	**Energy related**	**Hjörth parameters**	**Decorrelation time**	**Autoregressive modelling**	**Power related**	**Wavelet coefficients**	**Correlation dimension and correlation sum**	**Noise Level**	**Lyapunov exponent**	**Lempel-Ziv complexity**	**Entropy**	**Dynamical similarity index**	**Mean coupling phases**	**Line length**	**Max. Linear cross-correlation**	**Ratio and Differences**	**Wavelet related**	**Non-linear interdependence**	**Correlation entropy**	**Mean phase coherence**	**Dynamical entrainment**	**Synchrony related**
Viana et al. [92] (2022)	Raw data and FFT data						x																	
Pal Attia et al. [93] (2022)	Raw data, FFT data, and TOD						x																	
Pinto et al. [89] (2022)		x																						
Stirling et al. [68] (2021)	HR features Time of the day Sleep features																							
Nasseri et al. [70] (2021)	Raw data HR Time of the day																							
Tamanna et al. [79] (2021)		x	x					x					x											
Usman et al. [80] (2021)	From raw data to STFT						x																	
Pinto et al. [82] (2021)			x																					
Tsiouris et al. [81] (2018)	From raw data																							
Agarwal et al. [94] (2018)	CNN Feature extraction																							
Chamseddine et al. [90] (2018)							x																	
Kiral-Kornek et al. [95] (2018)	From raw data to Spectograms Time of the day						x																	
Karoly et al. [18] (2017)			x													x								
Khan et al. [96] (2017)								x																
Aarabi et al. [78] (2017)									x		x	x								x	x			
Direito et al. [83] (2017)		x	x	x	x	x	x	x																
Assi et al. [91] (2015)				x	x		x																	
Bandarabadi et al. [84] (2015)							x																	
Rasekhi et al. [85] (2015)		x	x	x	x	x	x	x										x						
Moghim et al. [?] (2014)		x	x				x	x	x		x													
Teixeira et al. [86] (2014)		x	x	x	x	x	x	x																
Alvarado-Rojas et al. [69] (2014)															x									
Rasekhi et al. [87] (2013)		x	x	x	x	x	x	x																
Rabbi et al. [88] (2013)		x	x	x	x	x	x	x						x						x		x		x

FFT stands for Fast-Fourier Transform, STFT for Short-Time Fourier Transform, TOD for Time Of the Day, HR for Heart Rate, and CNN for Convolutional Neural Network.

3.1.5 Feature selection and reduction

Prediction algorithms typically incorporate numerous features to account for brain dynamics, as the transition from the interictal to the ictal state involves complex mechanisms. That results in high dimensional feature spaces; some features might be redundant while others might be confounding, thus damaging the classifier's performance. Therefore, choosing the features that help identify the preictal period most effectively while minimising information loss is essential. [12].

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 [91]. maximum Difference Amplitude Distribution of histogram (mDAD) [84] is based on amplitude distributions histograms. The most discriminatory variables are those that contributed with the least superposition of the histograms for each class. minimum Redundance Maximum Relevance (mRMR) ranks features by maximising their relevance while minimising the redundancy among them [84, 85]. Genetic Algorithms use natural selection as their inspiration. They are based on biological principles: from a starting population, the most resilient individuals will survive and mate to adapt to environmental changes [82, 89, 91].

Another way to deal with the curse of dimensionality is feature reduction, where the number of features is reduced by combining the original ones into new ones while aiming not to lose important information. Such as Principal Component Analysis (PCA) this method projects the data onto an orthogonal space and select the projections with higher variance values [99].

Finally, in DL approaches, reduction is performed either by convolutional layers [80, 94] or through autoencoders [100].

3.1.6 Classification

It is assumed that with the remaining features, the trained model can distinguish between interictal and preictal periods. Consequently, a classification algorithm will use the information given by that feature space to determine which class it belongs to. The algorithm must first be trained on making decisions, and only then can it be used with unobserved data.

From simpler to more complex, there is a considerable heterogeneity of algorithms that can be used. The transition from Support Vector Machines (SVMs) [6, 79, 83–87, 91, 94] to Long Short-Term Memory (LSTMs) [68, 70, 80, 81, 90, 92, 93] and CNNs [80, 90, 95, 96] has been seen. Other classifiers, like logistic regressions [18, 68, 82, 89], Gated Recurrent Units (GRUs) [90], and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [88, 91] have also been used. Table 3.4 presents the classification and performance evaluation decisions taken in studies in recent years.

Since the interictal time is substantially longer than the preictal period, as was

mentioned in section 2.4, data imbalance poses a severe problem for seizure prediction. While several authors have addressed the issue by undersampling (deleting interictal samples) [83,84,86,87,91], others have used cost-sensitive classification algorithms [12,101] or even created new preictal samples artificially using techniques like Generative Adversarial Networks [102].

Partition strategies

Several data partition methods have been adopted, as there is no standard procedure. Authors should not use for testing data used in training and not use segments from the same ictal event for training and testing. Also, should not select random segments from the entire data, as this may lead to a biased performance since the model may be trained with segments close to the ones used for testing.

The different approaches arise from different assumptions concerning the seizure-generating process. Some consider seizure generation a patient-independent process, so they pool all data from all patients and then select a certain number of seizures for training and the remaining for testing [93, 103, 104]. Most studies (see Table 3.4) assume a patient-specific approach, in which the model is trained and tested for each patient [6, 83, 88], and others go further and also take into account the concept drift of seizures, thus considering the chronology of seizures to split the data [18,69,70,82,84–87,89,92,95].

Studies performed with ultra long-term recordings (lasting at least months for each patient) assume and aim to capture all types of concept drifts [18,70,95] or go as far as periodically retraining their classifiers [68].

SVMs

SVMs are a widely adopted classifier for seizure prediction. It is a supervised learning model that define a linear separation hyperplane in an N-dimensional space to maximise the distance between nearby points of various classes [5,12,85].

Using kernel functions like the Radial Basis Function (RBF), SVMs can carry out non-linear classifications. It is possible to convert the initial feature space into a higher-dimensional one, enabling the creation of a hyperplane with a greater margin of separation between the two classes [5,12,85].

The SVM is appealing from the perspective of interpretability due to its ability to linearise the feature space and analyse the produced support vectors. However, it is essential to keep in mind that (as with any other classifier) its interpretability may be lost if the number of features becomes excessive.

CNNs

CNNs are DL algorithms that can learn the best features from the input and are designed to process data with more than one dimension or several arrays (e.g. im-

Table 3.4: Overview of the classifiers, regularisation, performance, and statistical validation in the past 10 years.

Study	Partition Strategy	Classifier	Regularisation	Performance	Statistical Validation
Viana et al. [92] (2022)	Training: initial 1/3 of data Testing: last 2/3 of data	LSTM	1h smooth	SS=0.73 TiW=0.34	Surrogate analysis
Pal Attia et al. [93] (2022)	k-fold cross validation with patients	LSTM	1h smooth	SS=0.54 TiW=0.33	Surrogate analysis
Pinto et al. [89] (2022)	Training: first 3 seizures Testing: the remaining	Logistic Regression	Firing Power	SS=0.16 FPR/h=0.34	Surrogate analysis
Stirling et al. [68] (2021)	Retraining and testing chronologically and iteratively	LSTM+Random Forest+Log Reg	Kalman filter	AUC=0.74	Random Forecast
Nasseri et al. [70] (2021)	Training: first 2/3 of data Testing: last 1/3 of data	LSTM	Kalman filter	AUC=0.80	Random Predictor
Tamanna et al. [79] (2021)	Training: 80% sample Testing: the remaining	SVM	K-of-N analysis	SS=0.96 FPR/h=0.19	N.A.
Usman et al. [80] (2021)	k-fold cross validation with seizures	CNN+LSTM	N.A.	SS=0.93 SP=0.92	No
Pinto et al. [82] (2021)	Training: first 60% chronological seizures Testing: last 40% chronological seizures	Logistic Regression	Firing Power	SS=0.37 FPR/h=0.79	Surrogate analysis
Tsiouris et al. [81] (2018)	K-fold with recordings	LSTM	N.A.	SS=0.99 FPR/h=0.002	No
Agarwal et al. [94] (2018)	Training: 50% samples Testing: remaining	SVM	N.A.	SS=0.96 SP=0.98	No
Chamseddine et al. [90] (2018)	Training: 80% samples Testing: remaining	LSTM, GRU, CNN	Kalman filter, Firing Power	SS=0.88 SP=0.99	No
Kiral-Kornek et al. [95] (2018)	Training: first two months Testing: remaining	CNN	N.A.	SS=0.69 FPR/h=0.00	Random predictor
Karoly et al. [18] (2017)	Training: Day 100-200 Testing: Day 200 onwards	Logistic Regression	Bin width of 1h	SS=0.60 TiW=0.23	Time-matched predictor
Khan et al. [96] (2017)	10-fold cross validation	CNN	N.A.	SS=0.87 FPR/h=0.14	Random predictor
Aarabi et al. [78] (2017)	Training: 1 seizure Testing: the remaining	Thresholding	N.A.	SS=0.89 FPR/h=0.11	Random predictor
Direito et al. [83] (2017)	Training: 2-3 seizures Testing: the remaining	SVM	Firing Power	SS=0.38 FPR/h=0.20	Random predictor
Assi et al. [91] (2015)	Training: 80% segemnts Testing: the remaining	SVM, ANFIS	N.A.	SS=0.85 SP=0.80	No
Bandarabadi et al. [84] (2015)	Training: first 3 seizures Testing: the remaining	SVM	Firing Power	SS=0.75 FPR/h=0.10	Random predictor
Rasekhi et al. [85] (2015)	Training: first 3 seizures Testing: the remaining	SVM	Firing Power	SS=0.60 FPR/h=0.11	Random predictor
Moghim et al. [6] (2014)	10-fold cross validation Training: 70% data samples (random) Testing: 30% data samples	SVM	N.A.	SS=0.91 SP=1.01	Unspecific predictors
Teixeira et al. [86] (2014)	Training: first 2-3 seizures Testing: the remaining	SVM, ANN	Firing Power	SS=0.70 FPR/h=0.34	No
Alvarado-Rojas et al. [69] (2014)	Training: first 2-4 seizures Testing: the remaining	Thresholding	Kalman filter	SS=0.68 FPR/h=0.33	Random predictor
Rasekhi et al. [87] (2013)	Training: first 3 seizures Testing: the remaining	SVM, ANN	Firing Power	SS=0.73 FPR/h=0.15	Random predictor
Rabbi et al. [88] (2013)	Training: 1 seizure Testing: the remaining	ANFIS	N.A.	SS=0.80 FPR/h=0.46	No

LSTM stands for Long Short-Term Memory, CNN for Convolutional Neural Network, GRU for Gated Recurrent Unit, SVM for Support Vector Machine, ANFIS for Adaptive Neuro-Fuzzy Inference Systems, ANN for Artificial Neural Network, SS for Sensitivity, FPR/h for False Positive Rate per Hour, TiW for Time in Warning, and AUC for Area Under the Curve

ages). In other words, by identifying patterns, neural networks can automatically train features that deal with the temporal property directly [77, 105]. In the case of seizure prediction, the time series data is transformed to a compatible format (either raw, through Fast Fourier Transform (FFT), or wavelet decomposition) to serve as input. A CNN is a powerful tool but requires an enormous amount of labelled data points for training [96, 102].

Regarding architecture, CNNs typically stack various convolutional layers, which build feature maps with filtering operations using kernels. These are then followed by pooling layers that learn features from the resulting maps of the previous layers, which classification layers can then use. Lastly, dropout layers are often employed to prevent overfitting by setting the output of arbitrary units to zero during training [96].

LSTMs

LSTMs networks are another type of DL model, based on Recurrent Neural Networks (RNNs). They incorporate special units, named gates, responsible for controlling which information should be stored in memory and which should be forgotten (by learning the respective weights) [77].

Since they do not rely on a fixed window, LSTMs can learn temporal features of brain activity during different states while maintaining long-time dependencies, which in seizure prediction is an advantage over CNNs. Like the CNNs, they require vast amounts of data and are susceptible to overfitting [77, 100].

3.1.7 Regularisation

It is necessary to handle the temporal relationships between each classifier output as the generated classifiers are taught to make classifications on independent EEG segments. Since it is improbable that all samples are accurately classified and because the noise in online data is frequently present, especially during long-term recording [12, 86].

Regularisation methods are used to reduce the number of false alarms raised by the classification algorithm, such as the Kalman filtering [68–70, 90] and the Firing Power method [82–87,89,90]. These often include functions that consider the signal's temporal dynamics and smooth the classifier's output accordingly.

Kalman filter

The Kalman filter is based on the state estimation of a linear dynamic system at instant k where s_k is the system's state, and y_k is the classifier's predicted output. τ the prediction interval, and w_k and z_k are zero-mean white noise vectors (3.3). An alarm is only raised when the Kalman filter output is classified as a preictal sample. Moreover, new alarms can only be released when the output crosses the zero-threshold in a rising manner [106].

$$\begin{cases} s_{k+1} = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix} s_k + w_k \\ y_k = \begin{bmatrix} 1 & 0 \end{bmatrix} s_k + z_k \end{cases} . \tag{3.3}$$

Firing Power

The Firing Power method was proposed by Teixeira et al. [107] and implemented in several studies. Considering the binary classifier output $O[k]$ and τ, the number of samples in a sliding window (with equal size as the preictal period). Every time $O[k] = 1$, the sample k was classified as preictal, while when $O[k] = 0$, it was

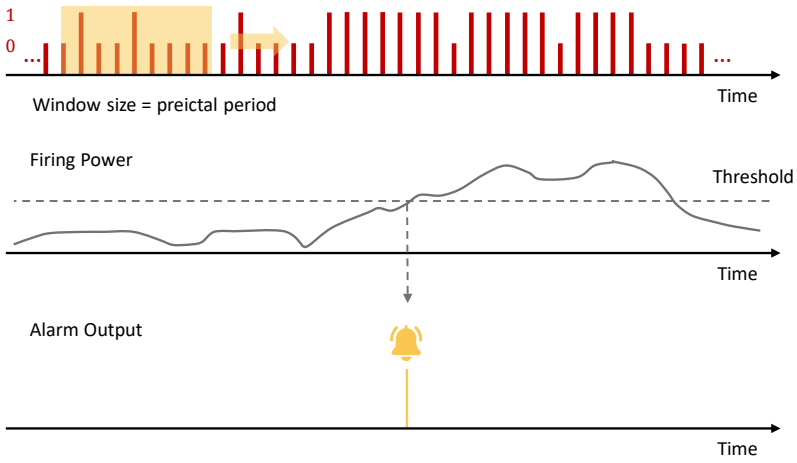

Figure 3.5: Visual representation of the Firing Power. An alarm is raised when a certain threshold is passed.

classified as interictal. The Firing Power output $fp[n]$ at instant n is shown in Equation 3.4:

$$fp\,[n] = \frac{\sum_{k=n-\tau}^{n} O\,[k]}{\tau}. \tag{3.4}$$

This method quantifies the relative number of samples classified as preictal and raises the alarm when its value is above a determined threshold (3.5). Despite the number of studies using Firing Power, no optimal threshold has been identified [12]. The higher the threshold is, the fewer false alarms are released. It is worth mentioning that since the window has the same duration as the preictal period, a single alarm should be released. To enforce that a refractory period of SOP+SPH duration is used after each alarm. When compared with the Kalman filter, the Firing Power (illustrated in Figure 3.5) produces fewer false alarms [12, 106].

$$\begin{cases} alarm\ if\ \ fp[n] \geq threshold, \\ no\ alarm\ if\ otherwise. \end{cases} \tag{3.5}$$

3.1.8 Performance Evaluation

Finally, a constructed approach has to be assessed using a set of metrics. Even though the seizure prediction characteristic [7] recommends employing seizure sensitivity, FPR/h, and statistical validation, including surrogate analysis or unspecific random predictors, not all studies adopt this approach. The Area Under the Curve [68, 70], sample sensitivity, and sample specificity [6, 80, 90, 91, 94] may all be valuable indicators of classifier performance. Still, they might not sufficiently comprehend how the system would function in a real-life scenario. As in recent years, a shift from seizure prediction to forecasting [18, 92, 93] was noticed, where in this

case, it is used Time in Warning (TiW) metric rather than FPR/h.

3.2 Concept Drifts in Epilepsy

Over the years, the effect of Concept drifts (CDs) in seizure prediction has been investigated. As mentioned in section 2.4, stress situations, the circadian, ultradian and infradian rhythms, cognitive states, environmental changes, implanting a neurostimulation device, a sudden brain lesion, and medication withdrawal in presurgical monitoring can affect brain dynamics. For each type of CD, different approaches can be taken. Training multiple classifiers, one for each concept, may prevent a gradual or recurrent drift. One may gradually retrain a classifier or give weights to instances to account for incremental drifts. Last but not least, concept states (such sleep-wake state, time of day, and others) might be used as extra features [108].

Identifying cyclic seizure patterns on several temporal scales, including circadian, multidien, and yearly, has been a focus in seizure prediction research. First, the frequency of seizures found in patient seizure diaries was evaluated to understand these cycles. Depending on the condition, there are particular times of the day, month, and year when seizures are more likely to occur. These patterns have been found by analysing ultra long-term recordings [109,110]. The findings from the most significant seizure cycle investigations are presented in Table 3.5.

Table 3.5: Studies on seizure occurrence cycles.

Study	Patient data	Seizure cycle prevalence
Leguia et al. [111] (2021)	Analysing only EEG Seizures: 85 patients using the RNS System; Analysing both EEG Seizures and diaries: 186 patients using the RNS System; Analysing Seizure diaries only: 194 patients	Circadian: 89% of patients; Multidien: 60% of patients; Circannual: 12% of patients
Baud et al. [72] (2018)	EEG seizures: 37 patients using the RNS System	Circadian in: 86% of patients; Multidien in: 93% of patients; Seizures often occur during the rising phase of multidien interictal epileptiform activity rhythms
Karoly et al. [112] (2018)	EEG Seizures: 12 patients from NeuroVista study (during 2 years); Seizure diaries 1118 patients from SeizureTracker (during 9 years)	Circadian in at least: 80% of SeizureTracker patients and 92% of Neurovista patients; Circaseptan: between 7% and 21% of patients
Ferastraoaru et al. [113] (2018)	Seizure diaries: 10186 patients (up to during 8 years) from SeizureTracker	Circadian pattern: higher seizure frequency between 7am and 10am, and lower overnight. Multidien pattern: higher seizure frequency during the work days comparing to the weekend

Responsive Neurostimulation System (RNS) is a closed-loop system that stimulates the cortex directly in up to two epileptogenic regions. SeizureTracker is an online seizure diaries, these diaries are used captures patterns of seizure cycles.

Authors confirmed a circadian cycle influence varying between 80% and 92% of patients using long-term EEG recordings [72,111,112]. As there are peaks associated with sleep-wake transitions, the sleep-wake cycle has also been shown to reflect seizure cycles. The circadian cycle and the sleep-wake states may be influenced by one another [109]. It is important to remember that not all circadian cycles are

associated with seizures. While seizures are not always the result of these cycles, they certainly raise their occurrence likelihood [110].

Seizures may also be influenced by interictal epileptiform activity. Seizures occur during the rising phase of this patient-specific cycle, which has a multidien periodicity. Circadian rhythms and sleep-wake cycles may, in part, influence this activity [22, 72, 110].

An electrocorticography-based logistic regression model, a circadian probability model, and an electrocorticography and circadian model combination were all compared by Karoly et al. [18]. Performance was maximised throughout several metrics with the addition of circadian information (the combined model).

3.3 Concept Drifts Adaptation

To learn in the presence of concept drift, algorithm designers must deal with two main problems. The first is detecting the current concept drift in the stream of data. Having identified concept drift, one needs to determine how to make appropriate predictions based on the new data. The following (see Figure 3.6) are the usual questions that are asked when building a concept drift handling technique:

- How is data processed in the application?
- How is the learning task processed?
- How is the concept drift monitored?
- How is the concept drift handled?

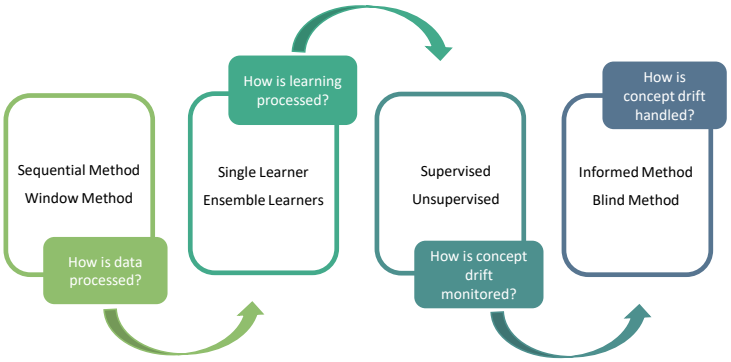

Figure 3.6: General scheme for drift handling methods. Adapted from Khamassi et al. 2018 [13].

3.3.1 Data process

Concept drift can also be addressed by modifying the training set seen by the algorithm (see Figure 3.7). In contrast to single or ensemble learners 3.3.2, modification

approaches are often classifier agnostic and, therefore, more flexible. The more common techniques used are sequential and windowing. These techniques consider that the most recent observations are the most informative and process data either sequentially, a single instance at a time or through a data window, multiple instances at once. Summarily, these strategies are not about developing new classifiers but are a set of rules to select or weigh the instances seen by the classifier [9, 13].

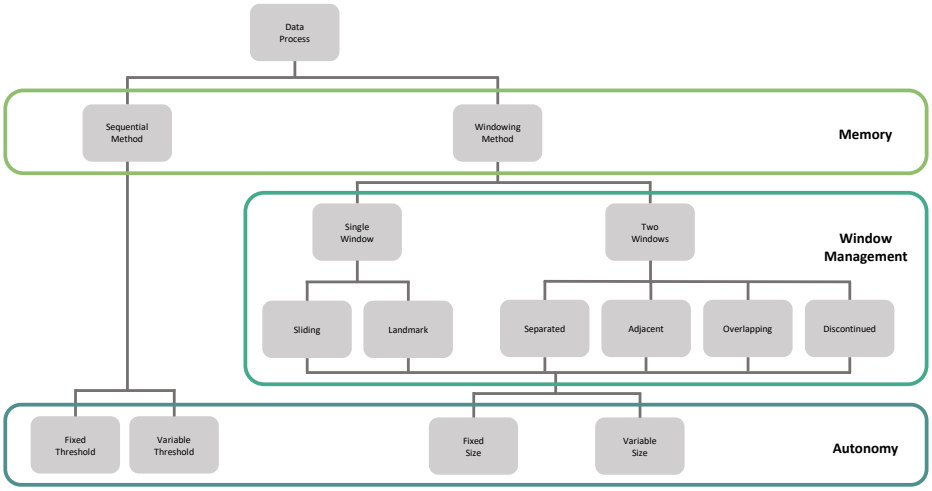

Figure 3.7: Methods taxonomy according to how data are processed in the application. Adapted from Khamassi et al. 2018 [13].

3.3.1.1 Sequential methods - Change detection

Online statistical analysis for anomaly/change detection is the source of inspiration for the sequential methods. Many researchers have adapted the well-established theories in statistics to learn in dynamic environments due to the similarity between the change detection problem in statistics and concept drift detection in machine learning [13]. These methods detect changes by evaluating how similar data is at certain point in time with a *Dissimilarity Measure*. Once it crosses a threshold (*Change threshold*), a concept drift is detected. Table 3.6 shows some studies that use sequential methods. Also, a more detailed explanation of these methods is depicted in Appendix B.1.

3.3.1.2 Windowing techniques

As a result of windowing, the naive algorithm keeps a fixed number of the most recent instances when changing the training set, as they are considered the most informative. The major disadvantage of this technique is that one cannot know *a priori* the proper window size [9, 13]. Table 3.7 presents an overview of studies exploring windowing techniques and a more detailed explanation of these methods is in Appendix B.1.

Table 3.6: Sequential methods overview for drift detection in the recent years.

Study	Distance Measures					Change threshold	
	Enclidean	Heterogeneous	Euclidean-Overlap	Mahalanobis	Hellinger	Fixed	Variable
Tran [114] (2019)	x						
Martfnez-Rego et al. [115] (2015)							x
Toubakh and Sayed-Mouchaweh [116] (2015)				x			
Goncalves et al. [117] (2014)				x			
Mejri et al. [118] (2013)							x
Ross and Adams [119] (2012)						x	
Ditzler et al. [120] (2011)					x		
Sobhani et al. [121] (2011)			x				
Lichtenwalter and Chawla [122] (2010)					x		
Luo et al. [123] (2009)							x
Dries and Ruckert [124] (2009)							x
Cieslak et al. [125] (2009)					x		
Tsymbal et al. [21] (2008)			x				
Muthukrishnan et al. [126] (2007)						x	
Nishida and Yamauchi [127] (2007)						x	

Table 3.7: Windowing techniques overview for handling drifts in the recent years.

Study	Specificity		Nature		Size		Positioning strategy						
							Single window		Two windows				Multiple windows
	Inside	Independent	Data-based	Time-based	Fixed	Variable	Sliding	Landmark	Separated	Adjacent	Overlapping	Discontinued	
Tran [114] (2019)											x		
Khamassi et al. [128] (2015)		x				x						x	
Khamassi et al. [129] (2014)						x				x			
Mejri et al. [118] (2013)						x				x			
Ditzler et al. [120] (2011)		x											
Cieslak et al. [125] (2009)		x											
Bifet et al. [130] (2009)	x												
Bach et al. [131] (2008)									x				
Pfahringer et al. [132] (2007)	x												
Bifet et al. [133] (2007)		x				x			x				
Hoeglinger et al. [134] (2007)	x												
Gama et al. [135] (2006)		x						x					
Baena et al. [136] (2006)		x						x					
Lazarescu et al. [137] (2004)													x
Kifer et al. [138] (2004)					x					x	x		
Black et al. [139] (2003)	x												
Babcock et al. [140] (2002)			x	x									
Hulyen et al. [141] (2001)	x						x						
Domingos et al. [142] (2000)	x												
Klinkenberg et al. [23] (2000)		x				x							

3.3.2 Learning process

When developing a drift handling technique, choosing between single and ensemble learners is frequently riddled with complications (see Figure 3.8). As a result, the principles, difficulties, and possibilities relating to individual and ensemble learners are examined in this section [13].

Table 3.8 presents an overview of studies exploring different techniques exploring the learning process.

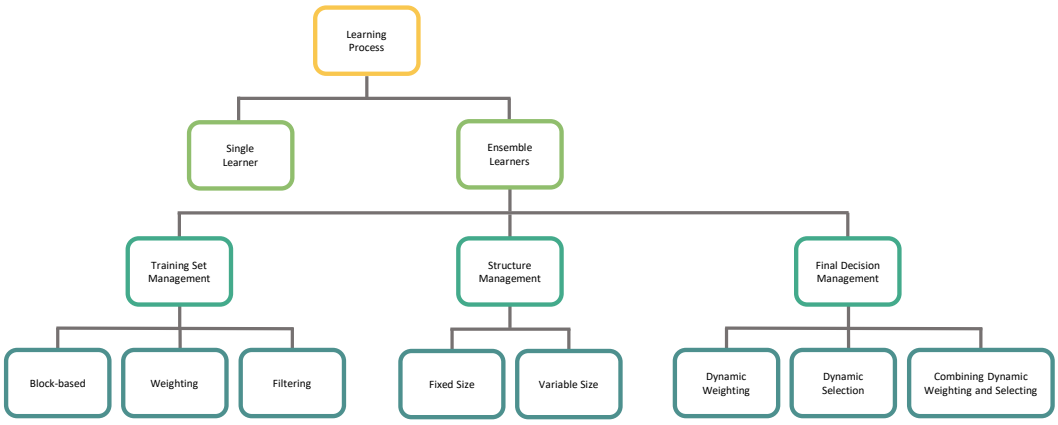

Figure 3.8: Methods taxonomy according to how the learning task is processed in the application. Adapted from Khamassi et al. 2018 [13].

3.3.2.1 Single learner

Single or adaptive base learners are deemed to be the simplest way to deal with concept drift. These learners can dynamically adapt to a new training data batch that contradicts the previous concept. This adaptation can take numerous forms, depending on the base learner implemented, but usually hinges on restricting or expanding the data used to make decisions [9].

Nonetheless, they are not recommended for handling recurrent drifts. As they are employed online, they constantly adapt to the current concept. Therefore, when a previous concept reappears, these methods relearn it from the beginning without taking advantage of its previous existence [13]. A few examples of single learner strategies are k-nearest neighbors (kNN) based methods [143, 144] or Decision tree-based methods [130, 132, 134, 139, 141, 142, 145].

3.3.2.2 Ensemble learners

Ensemble methods are popular in the data mining community due to their ability to deal with reoccurring concepts. This is an essential advantage over the previous techniques, as other approaches often discard historical data to learn new concepts. Accordingly, the success of the ensemble methods for handling concept drift relies on two primordial points: *diversity* and *adaptability* [13]. Those two points can be achieved through three levels:

- Training set management: how is data managed for training and adapting the ensemble learners?
- Structure management: how are base learners managed among the ensemble?
- Final decision management: how is the final decision processed and updated?

More details on Ensemble learners methods is in Appendix B.2.

Table 3.8: Learning process studies overview for handling drifts in the recent years.

Study	Single learners	Ensemble learnners							
		Training set management			Structure management		Final decision management		
		Block-based	Weighting	Filtering	Fixed size	Variable size	Dynamic weighting	Dynamic selection	Combining dynamic weighting and selection
Tran [114] (2019)				x					
Song et al. [146] (2016)							x		x
Brzezinski et al. [147] (2014)					x		x		
Jackowski et al. [148] (2014)						x			x
Sobolewski et al. [149] (2013)								x	
Khamassi et al. [150] (2013)			x		x		x		
Brzezinski et al. [151] (2013)		x			x		x		
Mejri et al. [152] (2013)							x		
Woźniak et al. [153] (2012)				x					
Elwell et al. [154] (2011)			x						
Bifet et al. [145] (2009)	x		x						
Alippi et al. [144] (2008)	x								
Tsymbal et al. [21] (2008)									x
Kolter et al. [155] (2007)		x			x		x		
Chu et al. [156] (2004)			x						
Polikar et al. [157] (2001)		x	x						

3.3.3 Monitoring process

The availability of predictive feedback may also impact the approach used to deal with drifting data (see Figure 3.9). Supervised indicator-based algorithms may be used if the true labels are instantly accessible following the prediction. However, strategies based on unsupervised indicators are the most appropriate when data is only partially labelled and the prediction feedback is delayed [13]. Table 3.9 presents an overview of studies exploring different techniques exploring the monitoring process. More details on methods that monitor concept drifts is depicted in Appendix B.3.

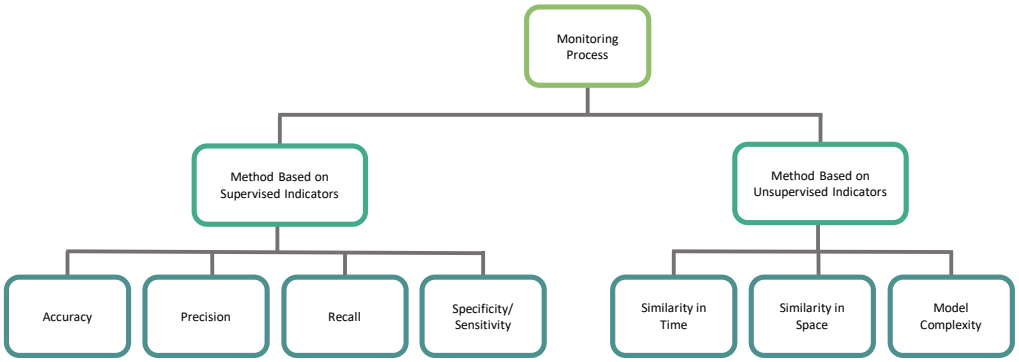

Figure 3.9: Methods taxonomy according to how concept drift is monitored. Adapted from Khamassi et al. 2018 [13].

Table 3.9: Monitoring process studies overview for handling drifts in the recent years.

Study	Supervised indicators	Unsupervised indicators		
	Accuracy	Similarity in Time	Similarity in Space	Model Complexity Measure
Tran [114] (2019)			x	
Khamassi et al. [128] (2015)	x			
Toubakh and Sayed-Mouchaweh [116] (2015)			x	
Goncalves et al. [117] (2014)			x	
Khamassi et al. [129] (2014)	x			
Shaker et al. [158] (2014)		x		
Ditzler [120] (2011)			x	
Lichtenwalter [122] (2010)			x	
Cieslak [125] (2009)			x	
Kuncheva et al. [159] (2009)		x		
Alippi et al. [144] (2008)		x		
Tsymbal et al. [21] (2008)			x	
Bifet et al. [133] (2007)	x			
Gama et al. [135] (2006)	x			
Baena et al. [136] (2006)	x			
Cauwenberghs et al. [160] (2000)				x

3.3.4 Adapting process

Determining how the learner will be adapted is the key to dealing with concept drift. There are two types of methods for this. In contrast to the second group, which includes blind systems that implicitly adapt to changes without any drift detection, the first category refers to informed methods that deliberately detect drift through triggering mechanisms [9, 13]. Table 3.10 presents an overview of studies exploring

different techniques exploring the adapting process. More details on methods that monitor concept drifts is depicted in Appendix B.4.

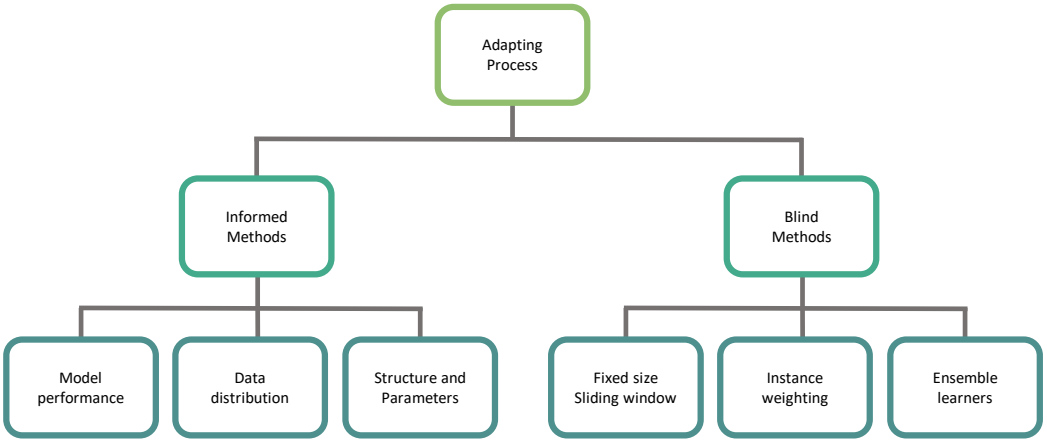

Figure 3.10: Methods taxonomy according to how concept drift is handled. Adapted from Khamassi et al. 2018 [13].

Table 3.10: Adapting process studies overview for handling drifts in the recent years.

Study	Informed methods			Blind methods		
	Model performance	Data distributions	Structure and Parameters	Fixed size sliding window	Instance weighting	Ensemble learners
Krawczyk et al. [161] (2015)					x	
Goncalves et al. [117] (2014)	x					
Khamassi et al. [129] (2014)	x					
Ditzler [120] (2011)		x				
Lichtenwalter [122] (2010)		x				
Cieslak [125] (2009)		x				
Pinto et al. [162] (2007)					x	
Muthukrishnan et al. [126] (2007)				x		
Gama et al. [135] (2006)	x					
Baena et al. [136] (2006)	x					
Kuncheva et al. [163] (2004)						x
Cauwenberghs et al. [160] (2000)			x			

3.4 Summary

Most seizure prediction studies use a similar framework consisting of several essential parts. Data of patients in presurgical monitoring are used in most existing databases. This is a barrier for devices used in real-world applications. In recent years, more studies have been done using ultra long-term recordings (acquired in a usual day-to-day scenario).

Many authors decide to filter the EEG signal after it has been acquired to remove power line interference, frequencies associated with noise, and other physiological and non-physiological artefacts. The most complex stage comes next: feature extraction, which typically involves a sliding window method, where univariate linear features are more often used. In recent years, DL has also been used for automatic feature extraction.

Additionally, a classification method is then applied to find preictal alterations. The classifier's output must go through a post-processing stage to lower the frequency of false alarms and give temporal meaning to successive classifications. Several authors have employed DL techniques, including LSTMs, that deal directly with signal temporality. However, the therapeutic implementation of these techniques is compromised by their difficulty in interpretation and computational power. Post-processing procedures must be developed, with clinical interpretability being given priority.

Performance must be assessed using sensitivity, FPR/h or another appropriate metric; also, statistical validation should be performed.

Recently, a change from seizure prediction to forecasting has been noted. Additionally, this perspective can better account for current discoveries on seizure occurrence cycles and the impact of circannual, multidien, and circadian CDs. It is also important to underline that the concept drift problem is extensive, and new ideas are being continuously developed over time.

Chapter 4

Methodology

This chapter describes the methodology developed for this thesis. First, Section 4.1 makes a brief explanation of the proposed methods, followed by a short description of the used data in Section 4.2. Sections 4.3 to 4.10 characterise, respectively, the pre-processing strategy, algorithm design and performance evaluation. Section 4.11 describes the experimental setup and used python libraries. Finally, a brief summary is given in Section 4.12.

4.1 Pipeline overview

The present work aims to develop EEG-based patient-specific algorithms for epilepsy seizure prediction, considering concept drifts. In a real-life scenario, the idea would be to have a prediction algorithm able to retrain itself, recall, or forget information to improve its performance. Motivated by this goal, two independent analyses were conducted (see Figure 4.1).

The first analysis was to partition the data, where *hard-rules* were used. They are a straightforward way to deal with concept drifts by simply retraining the model; for that, two different iterative retraining strategies were developed (Add-One-Forget-One and Chronological). Those two methods were compared to the most common partition method, Control partitioning. The details of each will be presented in Section 4.5 of this chapter. The three strategies were tested on the Control approach, developed to predict seizure without incorporating intrinsic concept drift adaptation, and the best was used on the second analysis.

In general, each method is characterised as follows:

- Control partitioning: common partition method where the first three chronological seizures are used to train the model.

- Add-One-Forget-One: retraining method where only the last three chronological labelled seizures are used to train the model.

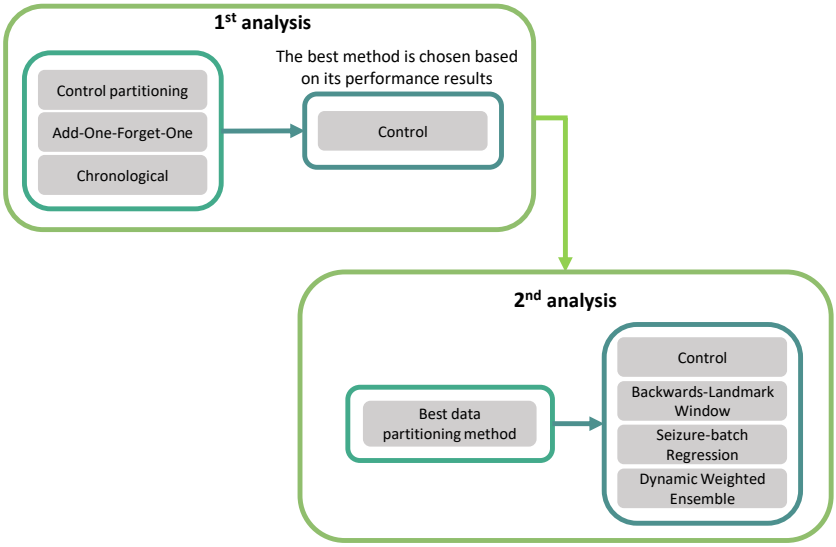

Figure 4.1: General outline of the two analysis preformed in this thesis

- Chronological: retraining method where all the last past seizures are used to train the model.

The second analysis tested approaches to predicting seizures while intrinsically adapting to concept drifts. Three approaches were proposed (Backwards-Landmark Window, Seizure-batch Regression and Dynamic Weighted Ensemble) and compared to the Control. These suggested methods were tested jointly with the two iterative retraining strategies (Add-One-Forget-One and Chronological). The details of each approach are described in Section 4.7.

In general, each approach is characterised as follows:

- Control: common seizure prediction algorithm is the control method.

- Backwards-Landmark Window: seizure prediction algorithm incorporating a window adjustment method by optimising performance with Support Vector Machines (SVMs) [23].

- Seizure-batch Regression: seizure prediction algorithm incorporating a data-batch (seizures) selection method using the angle difference between the logistic regression weights of consecutive seizures [24].

- Dynamic Weighted Ensemble: seizure prediction algorithm with a dynamic integration of classifiers [21].

The general framework of the adopted seizure prediction methodologies is summarised in Figure 4.2

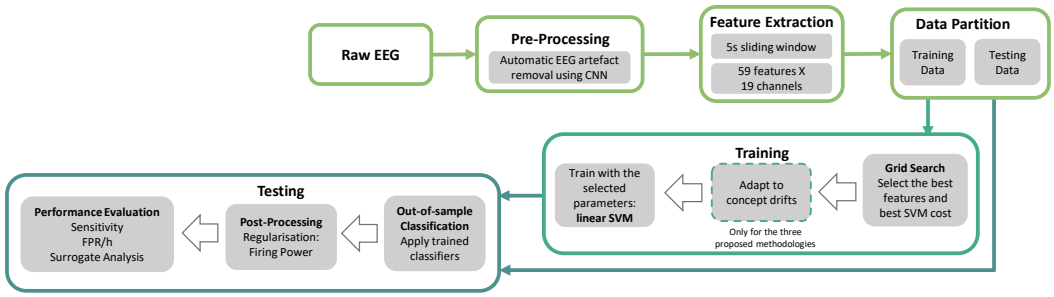

Figure 4.2: General outline of the proposed pipeline for each SOP.

4.2 Data

For the present study, 37 Drug-Resistant Epilepsy (DRE) patients (16 female and 21 male, with a mean age of 40.57 ± 15.76 years) were selected from the EPILEPSIAE. The chosen Electroencephalogram (EEG) data was gathered from patients with seizures in the temporal lobe by the University Medical Centre of Freiburg in Germany. The data comprises presurgical scalp EEG recordings obtained at a sampling rate 256Hz. It covers 19 EEG electrodes placed according to the International 10-20 System with the following channels: FP1, FP2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T7, T8, P7, P8, Fz, Cz, and Pz. Table 4.1 presents information regarding each patient (age and sex) and seizure information (number of seizures, seizure classification, seizure activity patterns, state of vigilance at seizure onset, and recording time). The 37 patients were chosen based on the number of independent seizures. To avoid clustered seizures, only patients with at least four lead seizures spaced by at least 4.5 hours were selected. Consequently, 207 of the 350 seizures were deemed appropriate for analysis.

Because the data is from patients undergoing presurgical monitoring, the current study can only be viewed as a proof-of-concept that, if successful, should be evaluated on more realistic data.

EPILEPSIAE database

The Epilepsiae database [25] has the most number of patients as it provides long-term (165 hours on average per patient) EEG recordings from 275 DRE patients in presurgical monitoring, along with extensive metadata and standardised annotation of the datasets. The stored metadata are about technical and clinical aspects, such as the patient's age, epilepsy characteristics, medication, and type of electrodes. Standardised annotations include seizure dominant pattern and type, state of vigilance, clinical onset, and offset. The recordings were acquired at the Epilepsy Centre of University Medical Centre of Freiburg (Germany), Centro Hospitalar e Universitário de Coimbra (Portugal) and Hôspital de la Pitié Salpêtrière in Paris (France).

Furthermore, due to its noninvasive nature, the scalp EEG recordings are the most popular method of recording. The use of this Data for research purposes has been approved by the Ethical Committee of the three hospitals involved in the development of the database (Ethik-Kommission der Albert-Ludwigs-Universität, Freiburg; Comité consultatif sur le traitement de l'information en matière de recherche dans le domaine de la santé, Pitié- Salpêtrière University Hospital; and Comité de Ética do Centro Hospitalar e Universitário de Coimbra). All methods were performed following the relevant guidelines and regulations. Informed written patient consent from all subjects and/or their legal guardian(s) was also obtained.

Table 4.1: Information for the 37 studied patients.

Patient ID	Age	Sex	Number of seizures	Seizure classification	Seizure activity pattern	Vigilance at seizure onset	Recording duration (h)
402	55	f	5	FOIA, FBTC, FOIA, FBTC, FOIA	t, t, t, t, t	A, A, A, A, A	133.47
8902	67	f	5	UC, FOIA, FOIA, FOIA, FOIA	a, b, a, m, a	A, A, A, A, A	156.41
11002	41	m	4	UC, FOIA, FOIA, FOIA	?, s, a, t	A, R, A, A	108.86
16202	46	f	7	UC, FBTC, UC, FOIA, FOIA, FOIA, FOIA	r, ?, r, r, r, ?, r	A, A, A, A, A, A, A	235.77
21902	47	m	4	UC, FOIA, FOIA, FOIA	t, t, t, b	A, A, A, R	76.84
23902	36	m	5	FOA, FOA, FOA, FOA, FOA	t, t, t, d, t	A, A, A, A, A	104.69
26102	65	m	4	FOIA, FOIA, FOIA, FOIA	m, t, t, t	A, A, A, A	83.23
30802	28	m	8	FOA, FOA, FOA, FOA, FOA, FOA, FOA, FOA	t ,t ,t, t, t, t, t, t	R, A, 2, A, A, R, 2, 2	149.28
32702	62	f	5	FOIA, FOIA, FOIA, FOIA, FOIA	t ,t ,t, r, a	A, A, A, A, A	141.87
45402	41	f	4	FOIA, FOIA, FOA, FOIA	t, t, t, t	A, A, A, A	94.29
46702	15	f	5	FOA, FOIA, FOIA, FBTC, FOIA	a, a, t, b, t	A, 2, A, 2, A	60.06
50802	43	m	5	FOIA, UC, UC, FOIA, FBTC	t, t, t, t, t	A, 2, 2, 2, A	201.53
53402	39	m	4	FOA, FOA, FOA, FOIA	?, ?, ?, T	A, 2, A, A	84.04
55202	17	f	8	FOIA, FOIA, FOA, UC, UC, FOA, UC, FOIA	t, d, t, t, t, t, r, r	A, A, A, A, A, A, A, A	112.42
56402	47	m	4	UC, UC, UC, FBTC	t, ?, ?, A	A, A, A, A	204.47
58602	32	m	6	FOIA, FOIA, FOIA, FOIA, FOIA, FOIA	r, t, t, r, r, t	A, R, A, A, A, 2	120.28
59102	47	m	5	FOA, FOIA, FOIA, FOIA, FOA	?, t, t, t, t	A, A, A, A, A	148.05
60002	55	m	6	FOIA, FOIA, FOIA, UC, FOIA, FOIA	d, c, t, t, d, d	1, A, A, R, R, 1	360.51
64702	51	m	5	FOA, FBTC, FBTC, FBTC, FBTC, FBTC	?, m, t, t, t	A, A, A, A, 2	107.5
75202	13	m	7	FOA, FOA, UC, FOA, FOA, FOA, FOA	t, t, t, t, t, ?, t	2, 2, A, A, A, A, A	153.57
80702	22	f	6	FOIA, FOIA, UC, FOIA, FBTC, FOIA	b, b, ?, c, c, c	A, A, A, A, A, A	78.95
85202	54	f	5	FOIA, FOIA, UC, UC, UC	m, c, m, m, m	2, A, A, A, A	73.91
93402	67	m	5	FBTC, FOIA, FOIA, UC, UC	t, t, t, t, t	2, 2, 2, 2, 2	152.07
93902	50	m	6	FOA, FOIA, FBTC, FOIA, FOIA, UC	t, t, d, d, d, d	A, A, 2, A, 2, A	391.12
94402	37	f	7	FOA, UC, FOIA, UC, FOA, UC, FOA	?, d, b, t, ?, b, ?	A, A, A, 2, A ,2, A	150.6
95202	50	f	7	FBTC, FOIA, FOIA, FOIA, UC, FOIA, UC	b, b, b, m, b, b, t	2, 2, 2, 2, 2, 2, 2	147.13
96002	58	m	7	FOIA, FOIA, FOIA, FOIA, UC, FOIA, FOIA	t, t, t, d ,a ,t ,a	A, A, A, A, A, A, A	130.6
98102	36	m	5	FOA, UC, UC, UC, FBTC	?, ?, ?, ?, ?	A, A, A, A, A	154.29
101702	52	m	5	FOIA, FOIA, FOIA, FOIA, FOIA	t, t, t, r, r	A, A, A, 2, A	52.24

102202	17	m	7	FOA, UC, FOIA, UC, FOA, FOIA, UC	b, ?, t, ?, t, t, t	2, A, 2, A, A, 2, A	108.86
104602	17	f	5	FOIA, FBTC, FBTC, FBTC, UC	t, a, t, t, d	A, 2, 2, 2, 2	103.12
109502	50	m	4	FOIA, FOIA, UC, UC	t, t, t, t	A, A, A, A	115.55
112802	52	m	6	UC, FOIA, UC, FOIA, FOIA, UC	t, t, t, t, t, t	A, A, A, A, A, A	183.08
113902	29	f	6	UC, FOIA, FOIA, FOIA, UC, FOIA	t, d, t, t, t, t	A, A, 2, A, 2, A	84.71
114702	22	f	8	FOIA, FOIA, UC, FOIA, FOIA, FOIA, FOIA, FOIA	t, t, t, t, d, t, d, t	A, A, A, A, A, A, A, A	102.43
114902	16	f	7	FOA, FOIA, FOIA, FBTC, UC, FOIA, FOIA	s, b, s, t, r, a, t	A, A, A, 2, A, A, A	77.21
123902	25	f	5	FBTC, FBTC, FOIA, FOIA, FOA	t, t, t, t, t	2, 2, R, A, A	182.26

Gender: female (f), male (m); Seizure classification: unclassified (UC), Focal Onset Aware (FOA), Focal Onset Impaired (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 interictal 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).

4.3 Pre-processing

An EEG artefact removal model based on Convolutional Neural Networks (CNNs) was used to pre-process the EEG data used in this study. Lopes et al. [71] developed this model to automatically remove artefacts from EEG signals, such as eye blinks, eye movements, muscular activity, heart activity, and electrode connection interference, in a manner comparable to that carried out by experts.

This method was created using EEG segments that had been manually pre-processed and labelled by experts. Those segments were used to train the deep learning model to replicate the actions of the experts during data pre-processing. Its performance was assessed by comparing denoised sections with the target segments (manually pre-processed). According to experimental findings, the proposed model could reduce the impact of EEG signal artefacts without human intervention, making it appropriate for use in long-term real-time scenarios such as warning devices. Additionally, the data used in this work were long-term EEG recordings from the same patients with epilepsy used in this thesis.

4.4 Feature extraction

After pre-processing, a five-second sliding window without overlap was used to extract relevant features from the EEG signals. According to the state of the art in seizure prediction, a five-second window was considered suitable to characterise EEG changes since it is a reasonable window regarding the stationarity, temporal and spectral resolution.

All the available electrodes were used because it was assumed that different brain areas could participate in the seizure-generating process. Also, only univariate linear features were extracted, as they are computationally less expensive. Correspond-

ingly, a sliding window technique was then used to calculate 59 univariate linear features in each window's time and frequency domains for 19 EEG electrodes.

For the frequency domain, the following features were extracted: relative spectral power bands delta (0.5-4Hz), theta (4-8Hz), alpha (8-13Hz), beta (13-30Hz), four gamma sub-bands - gamma band 1 (30-47Hz), gamma band 2 (53-75Hz), gamma band 3 (75-97Hz), and gamma band 4 (103-128 Hz), the ratio between these bands, spectral edge frequency and power, alpha peak frequency, total power, and mean frequency. For the time domain, the following features were extracted: the four statistical moments (mean, variance, skewness, kurtosis), Hjörth parameters (activity, mobility, complexity), and decorrelation time. The energy of the wavelet coefficients (from D1 to D5, using the db4 mother wavelet) was also extracted. More details on the extracted features can be found in Appendix A.

4.5 Data partition and iterative retraining

The feature set was split into two sets for each patient. The first set, containing at least three seizures, was used to train classifier training and parameter optimisation. The second set was constituted by all (Control partitioning) or the first (Add-One-Forget-One and Chronological) of the remaining seizures and used as testing set.

To account for concept drifts, the seizures were chronologically divided, respecting the order in which they occurred, or the classifier was retrained (Add-One-Forget-One or Chronological) after each new seizure. This division aimed to simulate a real seizure prediction scenario. Figure 4.3 illustrates how the data partition and iterative retraining were done.

4.6 Training

4.6.1 Class labelling

The feature set samples for seizure prediction were divided into interictal and preictal classes. The period before a seizure starts referred to as the preictal class, and it corresponds to the whole length of the Seizure Occurrence Period (SOP) and Seizure Prediction Horizon (SPH).

As this work aims to build an algorithm to be used in warning devices, followed by rescue medication after the alarm is released, a SPH of 10 minutes was chosen according to the medication's time to take effect, which varies between 5 and 10 minutes as seen in Table 2.1. The samples from this period were removed from the dataset as they were not necessary to train the model.

Regarding the SOP duration, several values were experimented with. A minimum of 10 minutes was set, as it was commonly used in the seizure prediction field, and a maximum of 50 minutes, as patients tend to prefer preictal periods of less

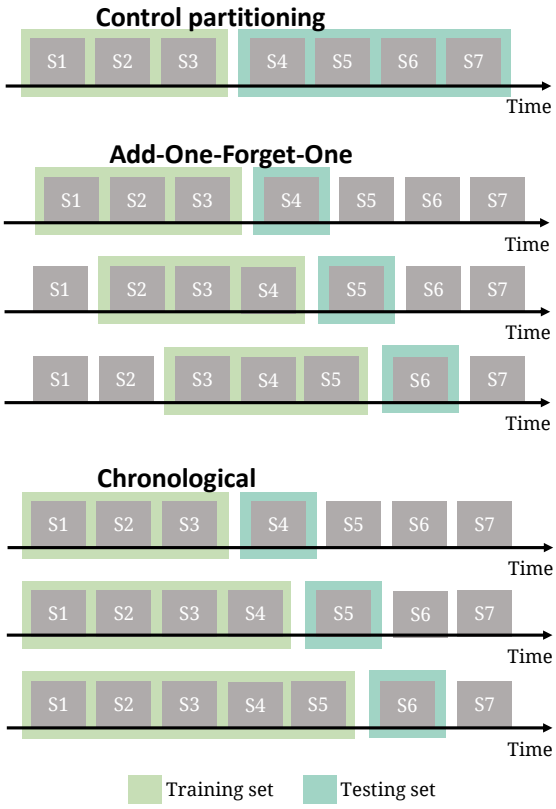

Figure 4.3: An illustrated scheme of the data partition and iterative retraining methods.

than one-hour [164]. Therefore, for each patient, the tested SOP values were: 10, 15, 20, 25, 30, 35, 40, 45, and 50 minutes.

4.6.2 Class balancing

As seizures are relatively rare, there is a considerable imbalance between interictal and preictal classes. During the training phase, a class balancing method was used to prevent bias and specialisation of the classifier over the dominant class. Therefore, a systematic random undersampling was performed for the Control, and weights inversely proportional to the class frequency were given for the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble. Two different methods were used because when a random undersampling is done on a time series, it is more difficult to detect the concept drifts properly. This way, the entire sequence of samples was deemed necessary, and therefore class weights were used to maintain the original evolution of the underlying context.

4.6.2.1 Random undersampling

A systematic random undersampling of the interictal samples was applied in each seizure to obtain an equal number of samples from each class, preserving the sequential chronology of the events. As a result, n groups, n being the number of preictal

samples, were formed from the entire collection of interictal data (see Figure 4.4). Then, a sample was randomly selected from each of these groups. This was done to maintain samples from all interictal intervals ensuring higher representativeness. It is worth noting that in the present work, this method was only used for the Control.

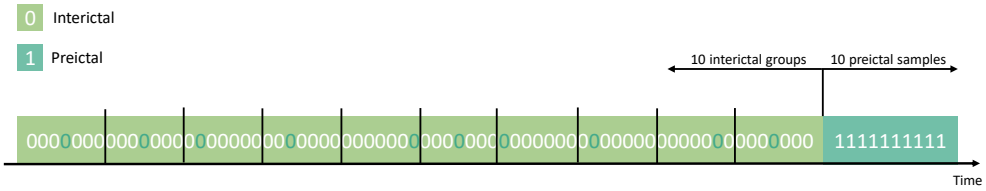

Figure 4.4: Random undersampling of interictal class respecting the sequential chronology of samples. Green coloured samples correspond to interictal samples randomly chosen from each group. Only one hypothetical seizure with 10 preictal samples is illustrated.

4.6.2.2 Class weights

Another strategy to consider the skewed distribution of the classes was to modify the algorithm to penalise the misclassification made by the preictal class. It was setting a higher class weight and simultaneously reducing the weight for the interictal class. To be more precise, the formula used to calculate the weight is represented by Equation 4.1. It is worth noting that in the present work, this method was used for the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble. Since these three were developed with intuit to handle concept drifts, it was deemed essential to have the entire sequence of samples of the interictal class.

$$w(i) = \frac{N_{Total\ \ Samples}}{N_{Class(i)\ \ Samples}}, i = 0, 1; \tag{4.1}$$

4.6.3 Feature standardisation

After class balancing, the range of independent features extracted from the raw data was normalised in a standardisation step. Each value was standardised using the z-score normalisation method, which set the mean of all values to 0 and the standard deviation to 1.

4.6.4 Feature selection

The most discriminative features were selected in this step using a filter-based method. These methods rank features depending on how they relate to the target. Comparing filter methods to other feature selection processes reveals that they are simple, faster, and less computationally expensive.

In this study, the metric used was the ANOVA (Analysis of Variance) f-test, which estimates the degree of linear dependency between each feature and the target. The k most discriminative features are selected according to their ranking to find the

most appropriate number of features (k). A grid-search strategy was implemented to choose k.

4.6.5 Classifier

The classifier used in this work was the SVM as it has been widely used for seizure prediction since this model has shown good results and has few parameters to tune. This classifier can be used with different kernels: linear, quadratic, polynomial, and Radial Basis Function (RBF) are the most well-known [83]. In this study, the kernel selected was the linear as it is simpler and computationally lighter, demonstrating comparable performances to more complex ones. The trade-off between smooth decision boundaries and accurate training point classification is controlled by the parameter C (cost) [83]. When C is high, more complicated decision curves are produced to match all the points, which might result in overfitting. The C parameter was tuned using a grid-search method.

4.6.6 Grid-Search

A grid-search method was used to find the optimal parameters to train the SVM classifier. It involved looking for the best preictal period (SOP), the correct value for the SVM hyperparameter (C), and the appropriate amount of features (k). For parameter C, it was considered 10 different values (2^{-10}, 2^{-8}, 2^{-6}, 2^{-4}, 2^{-2}, 2^{0}, 2^{2}, 2^{4}, 2^{6}, 2^{8}), and for k, four distinct values (10, 20, 30, and 40 features). Resulting in 40 combinations (k, C) evaluated for each SOP value defined in 4.6.1.

As illustrated in Figure 4.5, the Leave-One-Out Cross-Validation (LOOCV) strategy was implemented to find the optimal parameters. The LOOCV had two variations, one did not account for concept drifts, and the other was adapted to account for them during the grid-search procedure.

The first variation was used for the Control; at least two seizures were used as the training set and the remaining one as the validation set to evaluate the classifier. Therefore, for each combination (k, C), all training seizures were used precisely once to validate the model, resulting in at least three iterations of the LOOCV technique.

The second variation was used in the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble. If the validation seizure was the second, only the first was used to train the classifier. When the validation seizure was the third, in one iteration, only the second seizure was used to train the classifier, and in the other, the first and second were used to train it. Therefore, for each combination (k, C), all training seizures, except the first one, were used at least once to validate the model, resulting in at least three iterations of the LOOCV technique. A visual explanation of the different LOOCV combinations for a different number of training seizures is illustrated in Figure 4.6.

This partitioning technique ensures that samples from the preictal and interictal classes are included in the training and validation sets.

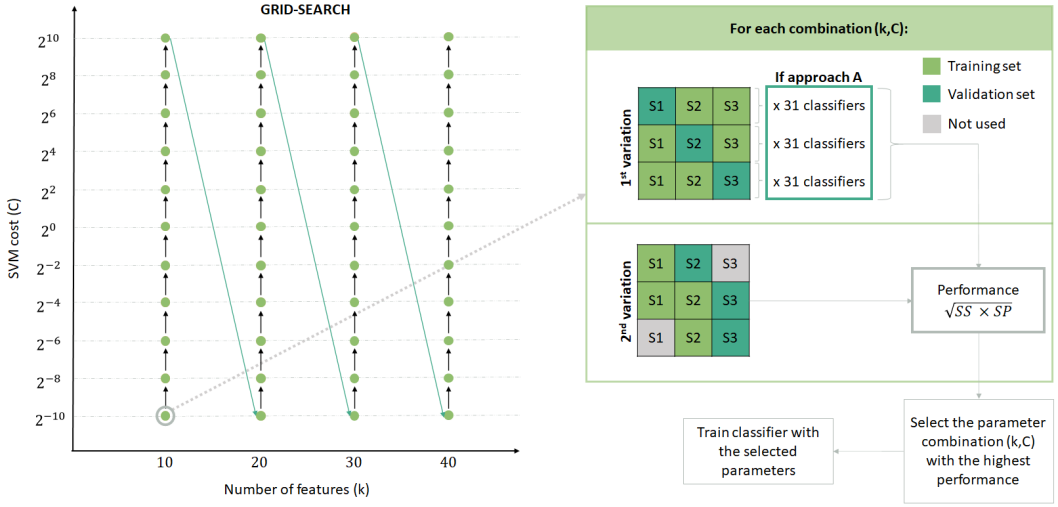

Figure 4.5: Grid-search procedure implemented to select the optimal training parameters for each preictal period.

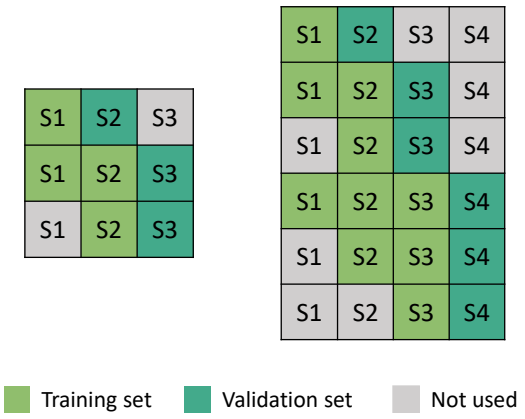

Figure 4.6: Illustration of the LOOCV combinations for grid-search second variation for 3 and 4 training seizures.

4.7 Concept drift adaptation

Regarding concept drift adaptation, three different approaches to predict seizures have been proposed in the present study. Some modifications to the most common seizure prediction methodology were done to implement Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble. Each approach is further detailed in this section and Table 4.2 shows their main differences.

4.7.1 Control

The Control aims to simulate the most common seizure prediction algorithm. Since this is the only approach that does not incorporate concept drift adaptation techniques, it is the control approach.

It is worth noting that due to the stochasticity of the random undersampling performed during the class balancing, an ensemble learning approach was implemented. In this approach, 31 SVM classifiers were trained using various data samples. This number was chosen to reach statistical significance and odd to prevent ties during testing, and also used in previous thesis.

4.7.2 Backwards-Landmark Window

Backwards-Landmark Window arises to create an algorithm that can select the data associated with the concept closest to the preictal period of last labelled seizure. With that in mind, a window adjustment algorithm was employed. The algorithm had to solve the following trade-off. A large window provides the learner with much training data, allowing it to generalise well, assuming that the concept did not change. On the other hand, a large window could contain old data that is no longer applicable (or even confounding) to the concept at hand. Finding the appropriate size means trading off the quality against the number of training examples [23].

To solve this problem, the window adjustment algorithm adapted from Klinkenberg and Joachims [23] was used in this work. At batch t, it essentially tries various window sizes of 1-hour difference, training a SVM for each resulting training set. For each window size, it computes a leave-one-out-estimate. The algorithm selects the window size that minimises the leave-one-out-estimate of the error rate. The algorithm is summarised in Algorithm 1.

Algorithm 1 Pseudocode illustrating the window adjustment method by optimising performance with SVMs.

input $\leftarrow$ $Z = S$ training samples in t batches of n samples (1-hour) each
for $h \in 0, ..., t\text{-}1$ **do**
  train SVM on samples $[Z(t\text{-}h,1),...,Z(t,n)]$
  compute the leave-one-out-estimate on samples $[Z(t\text{-}h,1),...,Z(t,n)]$
  m = estimate.linear_regression(last 12 hours).slope
  if $m \geq 0.05$ **then**
    stop window adjustment
  end if
end for
output $\leftarrow$ W = window size which minimises the leave-one-out-estimate

4.7.3 Seizure-batch Regression

Unlike the previous approach that monitors concept drifts by tracking previous samples, Seizure-batch Regression does it by selecting the data batches (seizures) more relevant to learn the current concept. This approach was adapted from Yeon et al. [24] and aims to discover what is the best combination of past seizure information. The algorithm trains a logistic regression for each chronological combination of train seizures and compares its weights with a logistic regression trained only with the last train seizure. That is done by calculating an angle between the two weight vectors (Equation 4.2) for each combination. The combination selected is the one with the smallest angle. Because when $\theta(w_1, w_2)$ is small it is assumed there is no drift or a gradual concept drift, and when $\theta(w_1, w_2)$ gets larger there is considerable concept drift. The algorithm is summarised in Algorithm 2.

$$\theta(\vec{w_1}, \vec{w_2}) = cos^{-1}\left(\frac{\vec{w_1} \cdot \vec{w_2}}{||\vec{w_1}|| \times ||\vec{w_2}||}\right) \tag{4.2}$$

Algorithm 2 Pseudocode illustrating the concept drift tracker via regression.

input $\leftarrow Z = S$ training seizures
test seizure index $\leftarrow T_i$ = length of Z
validation seizure index $\leftarrow V_i = T_i$ - 1 (last training seizure)
for $h \in 0, ..., V_i$ **do**
  combination of seizures $\leftarrow C = \mathrm{S}(h : V_i)$
  weight vector combination of seizures $\leftarrow \vec{w_1}$ = logistic regression(C)
  weight vector validation seizure $\leftarrow \vec{w_2}$ = logistic regression[$S(V_i)$]
  $\theta_h = \theta(\vec{w_1}, \vec{w_2})$
end for
output $\leftarrow C_{optimal}$ = combination with the smallest angle

4.7.4 Dynamic Weighted Ensemble

This approach was developed because ensembles are among the most popular and effective techniques to handle concept drifts [21]. Here a set of models built over different 1-hour periods were kept, and the models' predictions were combined according to their expertise level regarding the current concept. The assumption that the two hours before seizure onset were the current concept was made. In this work, a dynamic integration of SVM classifiers was employed, where each base classifier is given a weight proportional to its accuracy in the last two hours before the last training seizure. Then the classifiers were integrated using weighted voting. The algorithm is summarised in Algorithm 3.

Algorithm 3 Pseudocode illustrating the dynamic integration of classifiers for handling concept drift.

input ← $Z = S$ training seizures
window size ← W = 1-hour
for seizure index ∈ 0, ..., S-1 **do**
  windows = windowing(S[seizure index)]
  for h ∈ 0, ..., t-1 **do**
    segment = windows(h)
    classifier(seizure index,h).train[segment + S(seizure index).SOP]
    W(seizure index,h) = classifier(seizure index,h).accuracy[$S(t-1)$.last-2-hours]
  end for
end for
output ← W = ensemble weights

Table 4.2: Differences between drift handling methods.

Approach	How is data processed?	How is learning processed?	How is concept drift monitored?	How is concept drift handled?	Window
Backwards-Landmark Window	Window Method: Independent, Time-based, Variable size, Landmark	Single Learner	Supervised: Accuracy	Informed method: Model performance	Incremental 1-hour step
Seizure-batch Regression	Sequential Method: Logistic Regression weights as Dissimilarity Measure	Single Learner	Unsupervised: Similarity in Time and Space	Informed method: Structure and Parameters	N.A.
Dynamic Weighted Ensemble	Window Method: Independent, Time-based, Fixed size, Sliding	Ensemble Learner: Block-based, Variable Size, Dynamic Weighting	Supervised: Accuracy	Blind method: Ensemble Learner	Non-overlapping 1-hour window

4.8 Testing

After training the model, predictions were made using an out-of-sample classification on the testing set. The method used on the testing data was the same as the training set, except for the class balancing, as shown in Figure 4.7. As a result, the training set's z-score parameters were used to standardise the testing set, and the most discriminative features discovered during training were chosen. Finally, the output was decided using the SVM classifier.

This procedure was executed several times, depending on the approach. It was performed for each of the 31 trained classifiers for the Control, resulting in 31 predictions per sample. It was executed once for the Backwards-Landmark Window and the Seizure-batch Regression, resulting in a single prediction per sample. Finally, for the Dynamic Weighted Ensemble, it was conducted for each of the M trained classifiers, resulting in M predictions per sample. M is equal to the number of hours of interictal recordings present in the training set.

Therefore, for the Control and the Dynamic Weighted Ensemble, a voting system was employed. For the Control, the final output was assigned to the class with the most predictions for a specific instance (hard voting). Whereas for the Dynamic Weighted Ensemble, a combination of the weights of each model was done, and the class with the highest total probability was the final output (soft voting).

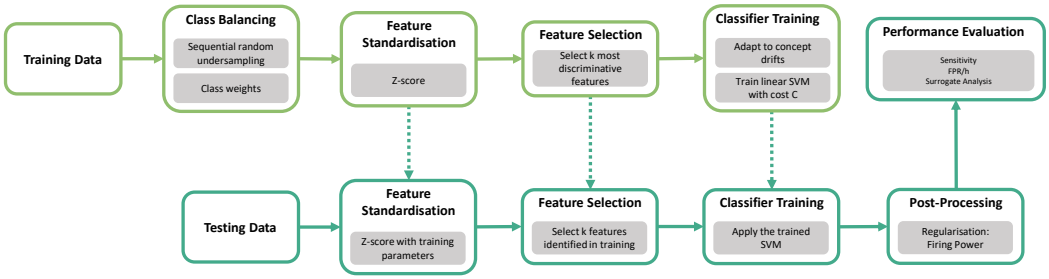

Figure 4.7: Procedure applied to train and test data the seizure prediction model.

4.9 Post-processing

A regularisation step was performed after classification to reduce the number of false alarms by considering the classification's temporal dynamics. The chosen method was the Firing Power, calculated as described in section 3.1.7. After that, an alarm was triggered when the Firing Power value surpassed a predefined threshold and was at least separated by one refractory period from the previously raised alarm (see Figure 4.8). The threshold was set at 0.5 [84], and the refractory period is equal to the whole duration of the preictal period (SOP + SPH). The refractory period was adopted to minimise repeated alerts during a seizure and lessen the patient's stress and anxiety.

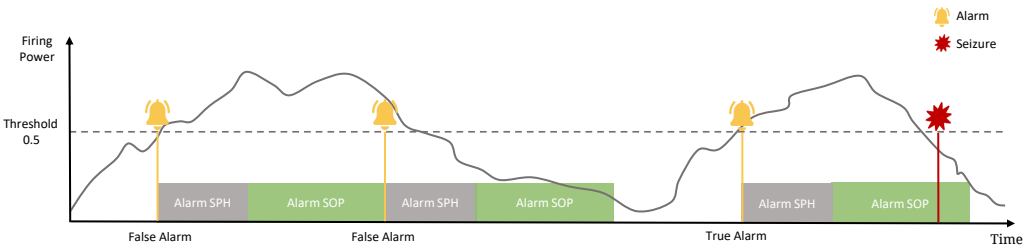

Figure 4.8: An illustration of the firing power technique used. When the firing power reaches a certain threshold (dashed line) and is at least one refractory period apart from the most recent alarm, an alert is triggered. Two false alarms and one true alarm are illustrated.

4.10 Performance evaluation

The performance of the developed seizure prediction models was evaluated using the Sensitivity (Equation 2.3) and False Positive Rate per Hour (FPR/h) (Equation 2.4) metrics, outlined in Section 2.3.2.2. In addition to the performance assessment, the surrogate analysis, described in Section 2.3.3.2, was used as a statistical validation strategy.

In terms of statistical validation, it was determined whether the developed algorithms outperformed the chance level using the surrogate time series analysis. The initial onset times were moved to a random position in the interictal interval. It

was done seizure by seizure to ensure that the simulated seizure times respected the seizure distribution across time. The sensitivity was then calculated using the surrogate times (new labels).

The average sensitivity that resulted from running this process 30 times was compared to the sensitivity estimated using the proposed methods. It outperforms chance in cases where the developed algorithm's sensitivity is larger than the surrogate one and statistically significant. The following null hypothesis was tested using a one-sample t-test with a statistical significance level of 0.05: "the sensitivity of the suggested approach is not superior than the sensitivity of the surrogate predictor."

4.11 Experimental setup

This methodology was tuned on a machine equipped with an Intel Core i7-8700 3.20GHz processor, 30GB of RAM running on Windows 10 and using Python 3.7 on Spyder 4.0.1. The used libraries were: *numpy, pandas, pickle, datetime, time, matplotlib, seaborn, scipy, random, math, sklearn, xlwings,* and *xlsxwriter*.

4.12 Summary

The suggested methodology for seizure prediction is based on scalp EEG data from 37 patients undergoing pre-surgical monitoring and suffering from TLE, collected from the EPILEPSIAE database. In addition to signal data involving 19 electrodes (following the 10-20 system).

The time-series data was split using 5-second chunks with no overlap. A deep learning technique filters and rebuilds the signal free of noise and artefacts. Fifty-nine linear univariate features were extracted from each channel: four statistical moments, three Hjörth parameters, decorrelation time, relative spectral power (delta, theta, alpha, beta, gamma), spectral edge frequency and power, alpha peak frequency, total power, mean frequency and wavelet coefficient energy (from D1 to D5, considering the Daubechies 4 mother wavelet).

First, to evaluate the influence of retraining the models after each seizure, three methods are compared on the most common seizure prediction framework (Control). The developed algorithm is iteratively trained, accumulating data from previous seizures and testing on the following one, or learns from the first three chronological seizures, and it is tested in the remaining.

The classifier is the linear SVM, and the classification output is processed using the Firing Power method with a threshold of 0.5. A refractory period is also implemented to produce a more realistic alarm-triggering mechanism. The models are evaluated according to the sensitivity and FPR/h.

After evaluating the influence of retraining models over time and selecting the best retraining method, three seizure prediction pipelines to intrinsically adapt to

concept drifts were developed. The Backwards-Landmark Window aims to select the samples from and close to the current concept. The Seizure-batch Regression tries to choose the best combination of past seizures to train the classifier. The Dynamic Weighted Ensemble builds a dynamically integrated ensemble to recall older concepts.

Finally, the models are applied to the testing set to evaluate their performance sensitivity and FPR/h are used. Statistical validation is also performed by surrogate time series analysis.

Chapter 5

Results

This chapter presents the results obtained in the present work and their interpretative analysis. Section 5.1 is focused on evaluating the influence of retraining the model after each seizure (iterative retraining), while Section 5.2 concentrates on the proposed approaches to adapt to concept drifts during the learning process intrinsically.

5.1 Iterative retraining

This section presents the results of training and testing for the first analysis. This first analysis evaluated the influence of each data partitioning and iterative retraining method. Hence the Control approach was used as it does not adapt to concept drifts. This way, allowing for a better understanding of the influence of the sheer act of retraining the patient-specific model after each seizure.

Even though several Seizure Occurrence Period (SOP) values were considered and tested, the findings reported in this section only include one SOP duration for each patient. The value was chosen based on its highest metric (Equation 5.1).

$$\sqrt{SS_{sample} \times SP_{sample}} \tag{5.1}$$

5.1.1 Training phase

During the training phase, a grid search technique was employed to build a patient-tailored model for epilepsy seizure prediction using each patient's first three seizures.

Tables 5.1, 5.2, 5.3 list the validation results (sample sensitivity and sample specificity) for each data partitioning and iterative retraining method, together with the training parameters chosen for each patient during the grid-search (1^{st} variation) procedure (Support Vector Machine (SVM) cost and the number of features). Additionally, the average sample sensitivity and specificity values for all patients.

Furthermore, the SOP values used vary significantly amongst patients, from the shortest tested duration (10 minutes) to the highest value employed (50 minutes).

The best preictal duration may differ between seizures and patients since the preictal period includes the whole length of the SOP and Seizure Prediction Horizon (SPH).

Although the average sensitivity and specificity values are similar between the three methods, specificity was relatively higher than sensitivity, which means that there is a better classification of the inter-ictal samples. One would anticipate better results and a greater classifier's ability to discriminate between the two classes. However, it will only be in the testing phase, using unseen data, that it will be possible to verify their true predictive potential. The relatively poor training results highlight the complexity of the seizure prediction problem.

Table 5.1: Training parameters and performance obtained for each patient for the Control data partitioning.

Control partitioning					
Patient	**SOP**	**k**	**C**	**SS**$_{sample}$	**SP**$_{sample}$
402	10	20	2^{-10}	0.43	0.71
8902	20	40	2^{-10}	0.88	0.83
11002	15	10	2^{6}	0.45	0.71
16202	15	40	2^{-4}	0.64	0.83
21902	10	10	2^{-10}	0.67	0.61
23902	50	40	2^{-10}	0.68	0.53
26102	50	40	2^{8}	0.31	0.62
30802	50	30	2^{-10}	0.90	0.79
32702	15	10	2^{-10}	0.75	0.69
45402	15	40	2^{-10}	0.72	0.56
46702	40	40	2^{8}	0.22	0.67
50802	15	10	2^{-4}	0.83	0.84
53402	40	10	2^{-10}	0.48	0.66
55202	10	30	2^{-4}	0.53	0.72
56402	10	10	2^{-6}	0.74	0.68
58602	10	10	2^{0}	0.30	0.70
59102	15	10	2^{-10}	0.63	0.45
60002	15	10	2^{-4}	0.54	0.72
64702	35	20	2^{-10}	0.46	0.67
75202	30	30	2^{-8}	0.72	0.83
80702	45	30	2^{2}	0.41	0.73
85202	15	30	2^{2}	0.45	0.67
93402	50	10	2^{8}	0.52	0.54
93902	40	30	2^{-2}	0.62	0.56
94402	10	40	2^{2}	0.40	0.68
95202	10	10	2^{-10}	0.74	0.66
96002	40	10	2^{6}	0.84	0.64
98102	35	10	2^{8}	0.50	0.53
101702	10	10	2^{-8}	0.58	0.53
102202	50	10	2^{0}	0.35	0.66
104602	15	40	2^{8}	0.35	0.66
109502	10	10	2^{-10}	0.58	0.53
112802	10	30	2^{-10}	0.55	0.56
113902	45	30	2^{-8}	0.42	0.55
114702	35	20	2^{2}	0.28	0.70
114902	20	20	2^{0}	0.47	0.63
123902	10	40	2^{-6}	0.80	0.85
Overall	-	-	-	0.56 ± 0.18	0.66 ± 0.10

Table 5.2: Training parameters and performance obtained for each patient for the Add-One-Forget-One data partitioning and iterative retraining.

			Add-One-Forget-One		
Patient	**SOP**	**k**	**C**	**SS**$_{sample}$	**SP**$_{sample}$
402	40, 20	10, 10	$2^{-10}, 2^{-4}$	0.53, 0.42	0.62, 0.65
8902	20, 20	10, 10	$2^{-10}, 2^{-10}$	0.87, 0.91	0.84, 0.81
11002	15	20	2^{-10}	0.51	0.68
16202	15, 10, 50, 20	30, 30, 10, 40	$2^{-10}, 2^{-4}, 2^{-8}, 2^{-6}$	0.63, 0.42, 0.52, 0.24	0.82, 0.73, 0.79, 0.71
21902	10	20	2^{2}	0.66	0.61
23902	45, 10	40, 40	$2^{-10}, 2^{-10}$	0.69, 0.51	0.52, 0.66
26102	50	40	2^{6}	0.31	0.59
30802	50, 50, 35, 10, 50	40, 40, 30, 10, 10	$2^{-8}, 2^{-4}, 2^{-8}, 2^{-10}, 2^{-10}$	0.90, 0.58, 0.38, 0.56, 0.78	0.80, 0.79, 0.64, 0.71, 0.53
32702	15, 20	10, 20	$2^{-10}, 2^{-10}$	0.77, 0.75	0.67, 0.72
45402	15	10	2^{-2}	0.72	0.53
46702	45, 15	30, 10	$2^{-4}, 2^{-10}$	0.23, 0.65	0.67, 0.78
50802	15, 20	30, 20	$2^{-8}, 2^{-10}$	0.80, 0.58	0.83, 0.79
53402	40	30	2^{8}	0.47	0.65
55202	10, 15, 25, 10, 10	10, 10, 10, 10, 10	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-8}$	0.61, 0.65, 0.70, 0.89, 0.63	0.66, 0.70, 0.49, 0.73, 0.58
56402	10	10	2^{-8}	0.80	0.64
58602	10, 10, 10	10, 40, 40	$2^{4}, 2^{-4}, 2^{-10}$	0.24, 0.33, 0.53	0.71, 0.79, 0.75
59102	15, 50	40, 40	$2^{4}, 2^{-2}$	0.64, 0.36	0.45, 0.44
60002	15, 30, 40	10, 10, 40	$2^{-8}, 2^{-8}, 2^{-10}$	0.58, 0.20, 0.55	0.72, 0.67, 0.7
64702	30, 20	30, 40	$2^{-10}, 2^{-10}$	0.42, 0.80	0.68, 0.67
75202	30, 10, 10, 50	30, 20, 40, 40	$2^{-10}, 2^{-10}, 2^{-10}, 2^{2}$	0.71, 0.70, 0.68, 0.83	0.83, 0.80, 0.70, 0.56
80702	45, 45, 10	20, 20, 40	$2^{-10}, 2^{-10}, 2^{-8}$	0.38, 0.54, 0.73	0.73, 0.58, 0.36
85202	15, 40	30, 20	$2^{6}, 2^{8}$	0.45, 0.63	0.66, 0.67
93402	50, 20	10, 10	$2^{-10}, 2^{2}$	0.57, 0.45	0.52, 0.75
93902	40, 50, 15	30, 30, 10	$2^{-10}, 2^{-10}, 2^{-10}$	0.60, 0.73, 0.58	0.57, 0.59, 0.54
94402	10, 40, 15, 15	10, 10, 30, 40	$2^{4}, 2^{-10}, 2^{-6}, 2^{-10}$	0.41, 0.23, 0.13, 0.54	0.63, 0.64, 0.77, 0.84
95202	10, 45, 50, 25	10, 40, 10, 10	$2^{-10}, 2^{8}, 2^{-10}, 2^{-10}$	0.76, 0.33, 0.22, 0.55	0.65, 0.71, 0.64, 0.63
96002	35, 20, 15, 30	40, 30, 40, 20	$2^{-10}, 2^{-2}, 2^{0}, 2^{-2}$	0.81, 0.45, 0.52, 0.63	0.65, 0.66, 0.72, 0.79
98102	35, 45	40, 20	$2^{8}, 2^{-2}$	0.51, 0.6	0.52, 0.68
101702	10, 45	20, 30	$2^{-6}, 2^{8}$	0.62, 0.47	0.50, 0.60
102202	45, 50, 45, 10	30, 40, 30, 20	$2^{0}, 2^{2}, 2^{-2}, 2^{-6}$	0.33, 0.43, 0.31, 0.24	0.65, 0.51, 0.53, 0.57
104602	25, 50	10, 10	$2^{-10}, 2^{-10}$	0.43, 0.60	0.62, 0.67
109502	10	10	2^{-6}	0.57	0.53
112802	10, 10, 10	10, 30, 30	$2^{-10}, 2^{-4}, 2^{8}$	0.54, 0.26, 0.33	0.58, 0.63, 0.64
113902	45, 15, 45	20, 30, 40	$2^{-10}, 2^{-10}, 2^{-10}$	0.42, 0.64, 0.46	0.56, 0.73, 0.61
114702	35, 15, 10, 50, 50	30, 40, 20, 10, 40	$2^{-8}, 2^{6}, 2^{6}, 2^{0}, 2^{-6}$	0.26, 0.24, 0.18, 0.64, 0.62	0.73, 0.77, 0.85, 0.54, 0.63
114902	20, 35, 30, 35	10, 30, 10, 40	$2^{-10}, 2^{6}, 2^{-10}, 2^{-10}$	0.52, 0.15, 0.3, 0.43	0.62, 0.66, 0.76, 0.76
123902	10, 10	40, 10	$2^{-6}, 2^{-4}$	0.79, 0.65	0.84, 0.83
Overall	-	-	-	0.53 ± 0.19	0.66 ± 0.10

Table 5.3: Training parameters and performance obtained for each patient for the Chronological data partitioning and iterative retraining.

			Chronological		
Patient	**SOP**	**k**	**C**	**SS**$_{sample}$	**SP**$_{sample}$
402	40, 50	10, 40	$2^{-10}, 2^{6}$	0.54, 0.59	0.62, 0.59
8902	20, 20	10, 10	$2^{0}, 2^{4}$	0.87, 0.91	0.84, 0.85
11002	15	20	2^{-10}	0.50	0.69
16202	15, 10, 15, 10	10, 30, 30, 40	$2^{-10}, 2^{2}, 2^{2}, 2^{6}$	0.59, 0.49, 0.59, 0.47	0.83, 0.82, 0.74, 0.74
21902	10	40	2^{-10}	0.67	0.61
23902	45, 15	40, 10	$2^{-8}, 2^{-10}$	0.69, 0.48	0.52, 0.68
26102	50	40	2^{6}	0.31	0.58
30802	50, 50, 50, 35, 35	40, 20, 20, 40, 10	$2^{-8}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.90, 0.72, 0.76, 0.79, 0.79	0.79, 0.78, 0.65, 0.66, 0.59
32702	15, 15	10, 10	$2^{-2}, 2^{-10}$	0.74, 0.82	0.71, 0.71
45402	10	40	2^{-10}	0.66	0.60
46702	35, 25	40, 40	$2^{-2}, 2^{8}$	0.20, 0.49	0.68, 0.63
50802	15, 20	30, 30	$2^{-10}, 2^{-10}$	0.80, 0.73	0.84, 0.77
53402	45	20	2^{-4}	0.48	0.67
55202	10, 10, 10, 10, 10	10, 10, 10, 10, 10	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-8}$	0.50, 0.74, 0.69, 0.67, 0.71	0.73, 0.67, 0.55, 0.64, 0.61
56402	10	20	2^{-10}	0.79	0.64
58602	10, 10, 15	40, 40, 30	$2^{-10}, 2^{-10}, 2^{-10}$	0.26, 0.62, 0.56	0.72, 0.67, 0.62
59102	15, 50	10, 10	$2^{-10}, 2^{-10}$	0.66, 0.52	0.44, 0.43
60002	15, 20, 15	30, 20, 20	$2^{-10}, 2^{-10}, 2^{-8}$	0.55, 0.41, 0.36	0.71, 0.69, 0.63
64702	25, 15	30, 40	$2^{-6}, 2^{-4}$	0.41, 0.58	0.68, 0.68
75202	30, 10, 30, 40	10, 10, 10, 30	$2^{-4}, 2^{-10}, 2^{-10}, 2^{-4}$	0.71, 0.73, 0.69, 0.68	0.83, 0.82, 0.62, 0.61
80702	45, 45, 40	40, 40, 40	$2^{8}, 2^{4}, 2^{0}$	0.43, 0.54, 0.49	0.73, 0.69, 0.69
85202	15, 15	20, 30	$2^{0}, 2^{2}$	0.44, 0.57	0.66, 0.69
93402	50, 20	20, 40	$2^{4}, 2^{-6}$	0.57, 0.44	0.52, 0.75
93902	40, 45, 40	40, 10, 30	$2^{0}, 2^{-6}, 2^{-6}$	0.60, 0.78, 0.73	0.57, 0.53, 0.44
94402	20, 20, 50, 10	10, 10, 30, 10	$2^{0}, 2^{-4}, 2^{0}, 2^{-8}$	0.46, 0.23, 0.28, 0.27	0.57, 0.69, 0.58, 0.65
95202	10, 30, 10, 40	10, 30, 40, 10	$2^{-10}, 2^{4}, 2^{0}, 2^{-10}$	0.74, 0.28, 0.68, 0.55	0.66, 0.62, 0.66, 0.55
96002	30, 15, 25, 25	40, 40, 40, 30	$2^{-2}, 2^{4}, 2^{-10}, 2^{-10}$	0.82, 0.56, 0.62, 0.62	0.64, 0.66, 0.64, 0.65
98102	35, 50	40, 10	$2^{-8}, 2^{2}$	0.50, 0.69	0.52, 0.65
101702	10, 40	30, 40	$2^{-4}, 2^{-2}$	0.49, 0.60	0.60, 0.51
102202	45, 50, 20, 50	20, 10, 40, 40	$2^{-4}, 2^{-4}, 2^{-2}, 2^{0}$	0.35, 0.46, 0.52, 0.51	0.65, 0.75, 0.67, 0.55
104602	35, 35	20, 20	$2^{-10}, 2^{-4}$	0.42, 0.48	0.61, 0.66
109502	10	10	2^{-10}	0.56	0.54
112802	45, 50, 10	10, 30, 40	$2^{2}, 2^{-6}, 2^{-4}$	0.31, 0.30, 0.51	0.70, 0.62, 0.44
113902	45, 10, 50	40, 30, 10	$2^{6}, 2^{-4}, 2^{8}$	0.43, 0.50, 0.65	0.55, 0.69, 0.47
114702	35, 45, 10, 15, 15	30, 40, 40, 10, 30	$2^{-10}, 2^{-4}, 2^{-10}, 2^{-10}, 2^{-10}$	0.25, 0.36, 0.44, 0.4, 0.57	0.73, 0.55, 0.63, 0.63, 0.53
114902	20, 20, 35, 35	20, 30, 10, 10	$2^{-4}, 2^{2}, 2^{-10}, 2^{-10}$	0.46, 0.37, 0.54, 0.6	0.65, 0.59, 0.56, 0.61
123902	15, 10	40, 10	$2^{-10}, 2^{-10}$	0.80, 0.64	0.81, 0.85
Overall	-	-	-	0.56 ± 0.16	0.64 ± 0.10

5.1.2 Testing phase

The created patient-specific models were assessed during the testing phase by considering each patient's remaining seizures. As a result, Equations 2.3 and 2.4 were used to evaluate sensitivity and False Positive Rate per Hour (FPR/h), respectively. In addition to the performance evaluation, statistical validation was performed using the surrogate time series analysis.

Tables 5.4, 5.5, 5.6 present the seizure prediction results obtained for each patient.

These tables show that there were between 1 and 5 assessed seizures. As a result, comparing the sensitivity values between patients could be challenging. For instance, the sensitivity is only restricted to 0 (seizure not predicted) or 1 (seizure correctly predicted) when just one seizure is considered. Accordingly, if only three seizures are analysed, the sensitivity values are limited to 0, 0.33, 0.67, or 1. Therefore, 73% of the patients in this study have very limited sensitivity values. As a result, a sensitivity value of 1 in a patient with one testing seizure and a sensitivity value of 1 in a patient with five assessed seizures have different meanings. Consequently, it may be seen as a drawback of the current study that could only be overcome by the availability of more extensive data for each patient.

Considering the Control partitioning method (Table 5.4), the average sensitivity and FPR/h values obtained across the 37 patients were 0.35 ± 0.35 and 1.88 ± 2.05, respectively. Additionally, the sensitivity of the surrogate time series analysis method averaged at 0.25 ± 0.50, where only 17 patients (45.95%) achieved performance above the chance level.

The Add-One-Forget-One method (Table 5.5) obtained average values of 0.63 ± 0.34 for sensitivity and 1.72 ± 2.65 for FPR/h. The mean sensitivity of the surrogate time series analysis was 0.26 ± 0.51, and 31 patients (83.78%) were statistically validated.

Finally, in the Chronological method (Table 5.6), the average sensitivity and FPR/h values obtained were 0.68 ± 0.32 and 1.44 ± 2.06, respectively. The sensitivity of the surrogate time series analysis method averaged at 0.23 ± 0.48, where 34 patients (91.89%) achieved performance above the chance level.

Table 5.4: Testing performance obtained for each patient with the Control data partitioning method.

Control partitioning							
Patient	**Evaluated seizures**	**SOP**	**SS**	**FPR/h**	**SS Surrogate**	**p-value**	**Above chance**
402	2	10	0.50	5.34	0.17 ± 0.24	0.00	•
8902	2	20	1.00	0.16	0.10 ± 0.20	0.00	•
11002	1	15	1.00	0.57	0.07 ± 0.25	0.00	•
16202	4	15	0.00	0.07	0.03 ± 0.08	0.96	
21902	1	10	0.00	1.42	0.10 ± 0.30	0.96	
23902	2	50	0.00	2.93	0.68 ± 0.30	1.00	
26102	1	50	1.00	1.49	0.53 ± 0.50	0.00	•
30802	5	50	0.60	0.43	0.37 ± 0.16	0.00	•
32702	2	15	0.00	0.45	0.07 ± 0.17	0.98	
45402	1	15	0.00	3.40	0.33 ± 0.47	1.00	
46702	2	40	0.50	1.77	0.38 ± 0.36	0.04	•
50802	2	15	0.00	0.47	0.07 ± 0.17	0.98	
53402	1	40	0.00	0.99	0.30 ± 0.46	1.00	
55202	5	10	0.40	2.32	0.28 ± 0.20	0.00	•
56402	1	10	1.00	5.94	0.33 ± 0.47	0.00	•
58602	3	10	0.00	2.54	0.19 ± 0.28	1.00	
59102	2	15	0.50	10.04	0.48 ± 0.42	0.42	
60002	3	15	0.00	1.03	0.16 ± 0.22	1.00	
64702	2	35	0.50	0.89	0.15 ± 0.23	0.00	•
75202	4	30	0.00	0.11	0.03 ± 0.08	0.98	
80702	3	45	0.67	1.74	0.44 ± 0.22	0.00	•
85202	2	15	0.00	0.17	0.02 ± 0.09	0.84	
93402	2	50	0.50	3.89	0.85 ± 0.29	1.00	
93902	3	40	0.33	0.46	0.29 ± 0.29	0.21	
94402	4	10	0.00	3.14	0.23 ± 0.23	1.00	
95202	4	10	0.25	1.10	0.13 ± 0.14	0.00	•
96002	4	40	0.25	2.00	0.58 ± 0.28	1.00	
98102	2	35	0.50	0.21	0.10 ± 0.20	0.00	•
101702	2	10	0.50	2.39	0.25 ± 0.25	0.00	•
102202	4	50	0.50	0.32	0.13 ± 0.12	0.00	•
104602	2	15	0.50	0.88	0.23 ± 0.31	0.00	•
109502	1	10	1.00	3.23	0.33 ± 0.47	0.00	•
112802	3	10	0.33	4.49	0.38 ± 0.25	0.82	
113902	3	45	0.67	2.69	0.54 ± 0.24	0.00	•
114702	5	35	0.00	0.25	0.18 ± 0.17	1.00	
114902	4	20	0.00	0.13	0.01 ± 0.04	0.84	
123902	2	10	0.00	0.00	0.00 ± 0.00	-	
Overall	-	-	0.35 ± 0.35	1.88 ± 2.05	0.25 ± 0.50	-	17

Table 5.5: Testing parameters and performance obtained for each patient for with the Add-One-Forget-One data partitioning and iterative retraining method.

Add-One-Forget-One

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	40, 20	0.00	1.33	0.15 ± 0.23	1.00	
8902	2	20, 20	0.00	0.30	0.10 ± 0.20	0.99	
11002	1	15	0.00	4.53	0.37 ± 0.48	1.00	
16202	4	15, 10, 50, 20	0.75	0.34	0.11 ± 0.14	0.00	•
21902	1	10	1.00	16.19	0.40 ± 0.49	0.00	•
23902	2	45, 10	1.00	2.22	0.25 ± 0.25	0.00	•
26102	1	50	1.00	0.40	0.23 ± 0.42	0.00	•
30802	5	50, 50, 35, 10, 50	0.40	0.81	0.30 ± 0.17	0.00	•
32702	2	15, 20	1.00	1.69	0.20 ± 0.28	0.00	•
45402	1	15	1.00	3.63	0.33 ± 0.47	0.00	•
46702	2	45, 15	1.00	0.57	0.22 ± 0.31	0.00	•
50802	2	15, 20	1.00	0.59	0.12 ± 0.21	0.00	•
53402	1	40	1.00	0.84	0.20 ± 0.41	0.00	•
55202	5	10, 15, 25, 10, 10	0.60	1.38	0.21 ± 0.19	0.00	•
56402	1	10	1.00	4.57	0.27 ± 0.44	0.00	•
58602	3	10, 10, 10	0.67	1.68	0.14 ± 0.19	0.00	•
59102	2	15, 50	1.00	2.08	0.55 ± 0.30	0.00	•
60002	3	15, 30, 40	0.67	0.91	0.21 ± 0.20	0.00	•
64702	2	30, 20	0.50	1.89	0.32 ± 0.27	0.00	•
75202	4	30, 10, 10, 50	0.50	0.50	0.32 ± 0.18	0.00	•
80702	3	45, 45, 10	0.67	1.17	0.31 ± 0.27	0.00	•
85202	2	15, 40	0.00	1.04	0.37 ± 0.26	1.00	
93402	2	50, 20	0.50	0.45	0.22 ± 0.28	0.00	•
93902	3	40, 50, 15	0.67	2.21	0.32 ± 0.27	0.00	•
94402	4	10, 40, 15, 15	0.50	0.73	0.15 ± 0.14	0.00	•
95202	4	10, 45, 50, 25	0.75	0.53	0.16 ± 0.16	0.00	•
96002	4	35, 20, 15, 30	0.50	1.15	0.23 ± 0.22	0.00	•
98102	2	35, 45	0.50	0.54	0.30 ± 0.33	0.00	•
101702	2	10, 45	0.50	1.60	0.30 ± 0.28	0.00	•
102202	4	45, 50, 45, 10	0.50	0.42	0.15 ± 0.19	0.00	•
104602	2	25, 50	1.00	0.75	0.32 ± 0.33	0.00	•
109502	1	10	0.00	0.00	0.00 ± 0.00	-	
112802	3	10, 10, 10	0.67	3.06	0.22 ± 0.22	0.00	•
113902	3	45, 15, 45	0.67	1.20	0.32 ± 0.27	0.00	•
114702	5	35, 15, 10, 50, 50	1.00	0.99	0.31 ± 0.22	0.00	•
114902	4	20, 35, 30, 35	0.75	0.62	0.23 ± 0.23	0.00	•
123902	2	10, 10	0.00	0.64	0.05 ± 0.15	0.96	
Overall	-	-	0.63 ± 0.34	1.72 ± 2.65	0.26 ± 0.51	-	31

Table 5.6: Testing parameters and performance obtained for each patient for with the Chronological data partitioning and iterative retraining method.

Chronological							
Patient	**Evaluated seizures**	**SOP**	**SS**	**FPR/h**	**SS Surrogate**	**p-value**	**Above chance**
402	2	40, 50	0.50	0.00	0.00 ± 0.00	0.00	•
8902	2	20, 20	0.50	0.25	0.17 ± 0.24	0.00	•
11002	1	15	0.00	3.64	0.30 ± 0.46	1.00	
16202	4	15, 10, 15, 10	1.00	0.33	0.07 ± 0.13	0.00	•
21902	1	10	1.00	12.19	0.37 ± 0.48	0.00	•
23902	2	45, 15	0.50	1.95	0.38 ± 0.33	0.03	•
26102	1	50	1.00	0.31	0.13 ± 0.34	0.00	•
30802	5	50, 50, 50, 35, 35	0.60	0.70	0.35 ± 0.09	0.00	•
32702	2	15, 15	0.50	1.74	0.20 ± 0.28	0.00	•
45402	1	10	0.00	2.08	0.27 ± 0.44	1.00	
46702	2	35, 25	1.00	0.47	0.20 ± 0.24	0.00	•
50802	2	15, 20	1.00	0.61	0.02 ± 0.09	0.00	•
53402	1	45	1.00	0.89	0.23 ± 0.42	0.00	•
55202	5	10, 10, 10, 10, 10	0.40	1.99	0.25 ± 0.18	0.00	•
56402	1	10	1.00	4.53	0.30 ± 0.46	0.00	•
58602	3	10, 10, 15	0.33	1.87	0.20 ± 0.25	0.00	•
59102	2	15, 50	1.00	2.27	0.47 ± 0.22	0.00	•
60002	3	15, 20, 15	0.67	1.27	0.20 ± 0.2	0.00	•
64702	2	25, 15	1.00	2.26	0.35 ± 0.32	0.00	•
75202	4	30, 10, 30, 40	1.00	0.34	0.17 ± 0.19	0.00	•
80702	3	45, 45, 40	0.67	0.41	0.12 ± 0.16	0.00	•
85202	2	15, 15	0.50	0.38	0.03 ± 0.12	0.00	•
93402	2	50, 20	0.50	0.77	0.30 ± 0.28	0.00	•
93902	3	40, 45, 40	1.00	2.04	0.37 ± 0.29	0.00	•
94402	4	20, 20, 50, 10	0.50	0.50	0.10 ± 0.14	0.00	•
95202	4	10, 30, 10, 40	1.00	0.47	0.16 ± 0.15	0.00	•
96002	4	30, 15, 25, 25	0.50	2.49	0.36 ± 0.24	0.00	•
98102	2	35, 50	1.00	0.44	0.18 ± 0.27	0.00	•
101702	2	10, 40	0.50	1.24	0.25 ± 0.25	0.00	•
102202	4	45, 50, 20, 50	0.25	0.17	0.03 ± 0.08	0.00	•
104602	2	35, 35	1.00	0.57	0.25 ± 0.25	0.00	•
109502	1	10	0.00	0.00	0.00 ± 0.00	-	
112802	3	45, 50, 10	1.00	0.21	0.21 ± 0.20	0.00	•
113902	3	45, 10, 50	0.67	1.67	0.48 ± 0.29	0.00	•
114702	5	35, 45, 10, 15, 15	1.00	1.36	0.19 ± 0.13	0.00	•
114902	4	20, 20, 35, 35	0.75	0.34	0.11 ± 0.15	0.00	•
123902	2	15, 10	0.50	0.63	0.12 ± 0.21	0.00	•
Overall	-	-	0.68 ± 0.32	1.44 ± 2.06	0.23 ± 0.48	-	34

5.1.3 Comparative analysis between data partitioning and iterative retraining methods

Table 5.7 presents the overall performance of the implemented methodologies. Also, for a more intuitive overview, the obtained results for sensitivity and FPR/h are present in Figure 5.1

The average sensitivity values of the two proposed methods achieved in this study are almost identical and almost twice as higher as the Control partitioning. However, they have high standard deviations, suggesting that although some patients' models may have good predictive abilities, others may perform poorly.

Figure 5.1 shows a large dispersion of sensitivity values. Still, the Add-One-Forget-One and Chronological techniques are more restricted to higher sensitivity values, which may be a good sign of how well-predictive they are compared to the Control partitioning method. However, it is also apparent in Figure 5.1 that outliers had a distinct and considerable impact on the average values of FPR/h. It is simple to confirm that the values obtained for some patients are considerably large by carefully examining the Tables 5.4, 5.5, 5.6.

Table 5.7: Average seizure prediction performance across all patients for each data partitioning and iterative retraining method.

Retraining method	All patients			Validated patients		
	SS	FPR/h	SS Surrogate	%	SS	FPR/h
Control partitioning	0.35 ± 0.35	1.88 ± 2.05	0.25 ± 0.50	45.95	0.62 ± 0.28	1.75 ± 1.66
Add-One-Forget-One	0.63 ± 0.34	1.72 ± 2.65	0.26 ± 0.51	83.78	0.70 ± 0.27	1.69 ± 2.75
Chronological	0.68 ± 0.32	1.44 ± 2.06	0.23 ± 0.48	91.89	0.70 ± 0.30	1.32 ± 2.07

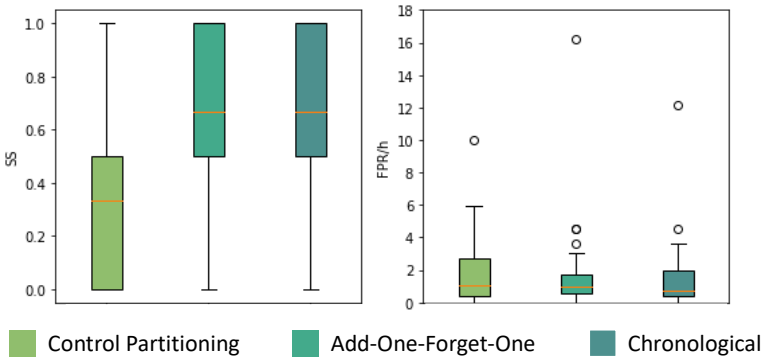

Figure 5.1: Seizure prediction performance across all patients for each data partitioning and iterative retraining method.

The average sensitivity for the three methods fared better than the surrogate time series analysis. All other methods—aside from the control partitioning method—were able to statistically validate more than 80% of the analysed patients.

Finally, as previously mentioned in Section 4.1, after accessing which is the best data partitioning and iterative retraining method, it would be used with the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted

Ensemble. As there was not a remarkable difference in the performance between the Add-One-Forget-One and the Chronological methods, they will be both used in the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble in the next section.

5.2 Concept drift adaptation

This section presents the training and testing results for the second analysis, where the Backwards-Landmark Window, Seizure-batch Regression and Dynamic Weighted Ensemble approaches were combined with the different partitioning methods and the respective Control. In this section, only the results for the Add-One-Forget-One procedure are presented as it fared better than the Chronological method for the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble (see Figure C.1). Refer to Appendix C for the results for the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble with the Chronological method.

Alike in Section 5.1, various SOP values were pondered and tested, and the results in this section are only present for the best SOP duration of each iteration for each patient.

5.2.1 Training phase

Once again, using each patient's first three seizures, a patient-tailored model was created using a grid search procedure.

The validation results for the Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble, are shown in Tables 5.8, 5.9, 5.10, as well as the training parameters chosen for each patient during the grid-search (2^{st} variation). The validation results for the Control are in Table 5.2.

It is again apparent that the SOP values range widely, from the smallest test length (10 minutes) to the greatest value used (50 minutes).

Compared to the other three, Seizure-batch Regression had a lower sensitivity and a higher specificity. Since the three proposed methodologies were designed to adapt to concept drifts during the grid search, better results could be expected. Although, conclusions regarding their predictive power can only be accessed during the testing phase.

Table 5.8: Training parameters and performance obtained for each patient with the Backwards-Landmark Window and the Add-One-Forget-One data partitioning and iterative retraining method.

Backwards-Landmark Window					
Patient	**SOP**	**k**	**C**	**SS**$_{sample}$	**SP**$_{sample}$
402	40, 40	40, 40	$2^{-2}, 2^{-4}$	0.50, 0.07	0.84, 0.53
8902	20, 30	10, 10	$2^{-10}, 2^{0}$	0.82, 0.97	0.78, 0.78
11002	20	10	2^{-4}	0.53	0.72
16202	15, 25, 50, 10	40, 30, 40, 20	$2^{-10}, 2^{-6}, 2^{-2}, 2^{-10}$	0.47, 0.70, 0.25, 0.36	0.82, 0.59, 0.75, 0.74
21902	40	10	2^{6}	0.71	0.73
23902	50, 10	20, 40	$2^{-2}, 2^{-6}$	0.81, 0.44	0.56, 0.75
26102	50	20	2^{-10}	0.47	0.72
30802	10, 10, 50, 25, 50	40, 10, 40, 20, 40	$2^{-10}, 2^{-10}, 2^{2}, 2^{-10}, 2^{-10}$	0.59, 0.26, 0.45, 0.59, 0.65	0.84, 0.86, 0.53, 0.70, 0.63
32702	15, 15	20, 10	$2^{-10}, 2^{-2}$	0.82, 0.83	0.76, 0.75
45402	50	20	2^{2}	0.67	0.57
46702	30, 10	20, 10	$2^{-8}, 2^{-10}$	0.37, 0.51	0.76, 0.85
50802	15, 25	10, 20	$2^{-4}, 2^{-10}$	0.73, 0.62	0.84, 0.87
53402	45	40	2^{0}	0.52	0.66
55202	35, 10, 25, 10, 10	10, 20, 30, 10, 30	$2^{-6}, 2^{-10}, 2^{2}, 2^{-10}, 2^{-6}$	0.62, 0.51, 0.60, 0.94, 0.53	0.53, 0.72, 0.56, 0.67, 0.69
56402	15	10	2^{-4}	0.80	0.67
58602	10, 10, 20	10, 20, 30	$2^{-10}, 2^{-10}, 2^{-10}$	0.45, 0.57, 0.50	0.79, 0.85, 0.73
59102	20, 15	20, 40	$2^{-10}, 2^{-4}$	0.48, 0.48	0.59, 0.37
60002	20, 15, 35	20, 20, 40	$2^{-6}, 2^{-8}, 2^{2}$	0.46, 0.51, 0.58	0.80, 0.66, 0.54
64702	50, 15	20, 10	$2^{-10}, 2^{-4}$	0.37, 0.54	0.56, 0.62
75202	30, 10, 45, 35	10, 10, 10, 30	$2^{-4}, 2^{-10}, 2^{-10}, 2^{-4}$	0.68, 0.53, 0.31, 0.56	0.85, 0.83, 0.85, 0.49
80702	10, 50, 50	40, 40, 40	$2^{8}, 2^{4}, 2^{0}$	0.66, 0.54, 0.54	0.34, 0.47, 0.54
85202	10, 30	20, 30	$2^{0}, 2^{2}$	0.21, 0.59	0.85, 0.55
93402	15, 20	20, 40	$2^{4}, 2^{-6}$	0.45, 0.67	0.71, 0.71
93902	30, 45, 20	40, 10, 30	$2^{0}, 2^{-6}, 2^{-6}$	0.42, 0.41, 0.75	0.34, 0.24, 0.76
94402	10, 20, 35, 15	10, 10, 30, 10	$2^{0}, 2^{-4}, 2^{0}, 2^{-8}$	0.51, 0.17, 0.39, 0.46	0.52, 0.79, 0.82, 0.99
95202	10, 35, 40, 10	10, 30, 40, 10	$2^{-10}, 2^{4}, 2^{0}, 2^{-10}$	0.82, 0.35, 0.39, 0.76	0.57, 0.66, 0.68, 0.90
96002	40, 20, 15, 50	40, 40, 40, 30	$2^{-2}, 2^{4}, 2^{-10}, 2^{-10}$	0.43, 0.23, 0.85, 0.33	0.63, 0.83, 0.66, 0.87
98102	25, 10	40, 10	$2^{-8}, 2^{2}$	0.22, 0.71	0.57, 0.68
101702	30, 50	30, 40	$2^{-4}, 2^{-2}$	0.63, 0.70	0.51, 0.58
102202	10, 20, 10, 15	20, 10, 40, 40	$2^{-4}, 2^{-4}, 2^{-2}, 2^{0}$	0.54, 0.49, 0.22, 0.56	0.63, 0.7, 0.9, 0.54
104602	30, 40	20, 20	$2^{-10}, 2^{-4}$	0.76, 0.79	0.55, 0.66
109502	10	10	2^{-10}	0.37	0.79
112802	10, 40, 25	10, 30, 40	$2^{2}, 2^{-6}, 2^{-4}$	0.43, 0.35, 0.51	0.56, 0.73, 0.41
113902	50, 10, 50	40, 30, 10	$2^{6}, 2^{-4}, 2^{8}$	0.44, 0.72, 0.73	0.52, 0.71, 0.56
114702	35, 15, 15, 50, 15	30, 40, 40, 10, 30	$2^{-10}, 2^{-4}, 2^{-10}, 2^{-10}, 2^{-10}$	0.22, 0.17, 0.67, 0.41, 0.56	0.76, 0.93, 0.7, 0.53, 0.74
114902	25, 50, 15, 20	20, 30, 10, 10	$2^{-4}, 2^{2}, 2^{-10}, 2^{-10}$	0.67, 0.05, 0.25, 0.75	0.44, 0.81, 0.93, 0.73
123902	40, 10	40, 10	$2^{-10}, 2^{-10}$	0.80, 0.33	0.77, 0.9
Overall	-	-	-	0.53 ± 0.19	0.70 ± 0.15

Table 5.9: Training parameters and performance obtained for each patient with the Seizure-batch Regression and the Add-One-Forget-One data partitioning and iterative retraining method.

			Seizure-batch Regression		
Patient	**SOP**	**k**	**C**	**SS**$_{sample}$	**SP**$_{sample}$
402	10, 10	10, 10	$2^{-10}, 2^{-4}$	0.52, 0.25	0.78, 0.53
8902	25, 30	10, 10	$2^{-10}, 2^{-10}$	0.77, 0.91	0.76, 0.85
11002	10	20	2^{-10}	0.42	0.71
16202	40, 10, 15, 30	30, 30, 10, 40	$2^{-10}, 2^{-4}, 2^{-8}, 2^{-6}$	0.65, 0.22, 0.15, 0.14	0.76, 0.85, 0.81, 0.81
21902	50	20	2^{2}	0.68	0.7
23902	50, 10	40, 40	$2^{-10}, 2^{-10}$	0.37, 0.55	0.62, 0.67
26102	50	40	2^{6}	0.26	0.74
30802	35, 15, 20, 50, 50	40, 40, 30, 10, 10	$2^{-8}, 2^{-4}, 2^{-8}, 2^{-10}, 2^{-10}$	0.66, 0.31, 0.56, 0.54, 0.65	0.85, 0.9, 0.56, 0.77, 0.64
32702	15, 25	10, 20	$2^{-10}, 2^{-10}$	0.57, 0.61	0.74, 0.71
45402	15	10	2^{-2}	0.36	0.76
46702	40, 10	30, 10	$2^{-4}, 2^{-10}$	05, 0.44	0.87, 0.82
50802	15, 25	30, 20	$2^{-8}, 2^{-10}$	0.7, 0.39	0.86, 0.9
53402	30	30	2^{8}	0.26	0.76
55202	10, 10, 10, 10, 50	10, 10, 10, 10, 10	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-8}$	0.73, 0.49, 0.52, 0.62, 0.55	0.55, 0.9, 0.6, 0.6, 0.51
56402	25	10	2^{-8}	0.41	0.74
58602	10, 10, 20	10, 40, 40	$2^{4}, 2^{-4}, 2^{-10}$	0.16, 0.2, 0.36	0.92, 0.95, 0.8
59102	50, 15	40, 40	$2^{4}, 2^{-2}$	0.38, 0.54	0.42, 0.32
60002	20, 20, 40	10, 10, 40	$2^{-8}, 2^{-8}, 2^{-10}$	0.44, 0.26, 0.44	0.77, 0.77, 0.83
64702	40, 25	30, 40	$2^{-10}, 2^{-10}$	0.44, 0.43	0.67, 0.79
75202	30, 10, 50, 50	30, 20, 30, 10	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-8}$	0.74, 0.6, 0.51, 0.84	0.83, 0.87, 0.83, 0.43
80702	35, 50, 15	40, 40, 30	$2^{-10}, 2^{-8}, 2^{-10}$	0.3, 0.63, 0.53	0.79, 0.55, 0.51
85202	10, 15	20, 10	$2^{-10}, 2^{-10}$	0.26, 0.5	0.85, 0.63
93402	20, 15	10, 10	$2^{0}, 2^{8}$	0.4, 0.46	0.66, 0.53
93902	50, 50, 15	10, 20, 20	$2^{-10}, 2^{-10}, 2^{-8}$	0.56, 0.45, 0.43	0.62, 0.84, 0.75
94402	10, 35, 30, 10	10, 20, 20, 30	$2^{8}, 2^{-8}, 2^{8}, 2^{-2}$	0.42, 0.5, 0.36, 0.27	0.56, 0.55, 0.69, 0.95
95202	10, 40, 45, 25	10, 40, 30, 30	$2^{-10}, 2^{8}, 2^{6}, 2^{-10}$	0.51, 0.36, 0.33, 0.66	0.65, 0.67, 0.56, 0.67
96002	25, 40, 35, 45	20, 30, 40, 20	$2^{-2}, 2^{-6}, 2^{-6}, 2^{8}$	0.52, 0.29, 0.61, 0.39	0.7, 0.63, 0.73, 0.86
98102	25, 35	20, 30	$2^{-4}, 2^{-2}$	0.31, 0.53	0.57, 0.7
101702	35, 50	20, 40	$2^{-2}, 2^{-8}$	0.4, 0.29	0.62, 0.77
102202	50, 45, 50, 50	30, 40, 30, 10	$2^{-6}, 2^{-4}, 2^{-2}, 2^{-6}$	06, 0.32, 0.27, 09	0.97, 0.48, 0.48, 0.75
104602	25, 45	10, 30	$2^{-8}, 2^{-10}$	0.38, 0.75	0.68, 0.67
109502	20	10	2^{-10}	0.35	0.57
112802	10, 15, 15	10, 10, 40	$2^{-10}, 2^{-6}, 2^{-10}$	0.65, 0.38, 0.19	0.49, 0.58, 0.79
113902	50, 10, 40	20, 40, 20	$2^{-10}, 2^{-6}, 2^{-2}$	0.13, 0.29, 0.42	0.77, 0.89, 0.66
114702	40, 30, 15, 45, 50	20, 20, 40, 10, 40	$2^{8}, 2^{-8}, 2^{-8}, 2^{-10}, 2^{-10}$	0.27, 04, 0.28, 0.52, 0.75	0.62, 0.84, 0.86, 0.55, 0.63
114902	25, 50, 50, 20	20, 30, 10, 10	$2^{0}, 2^{8}, 2^{-10}, 2^{-10}$	0.53, 04, 0.15, 0.67	0.44, 0.77, 0.91, 0.84
123902	50, 35	10, 20	$2^{-8}, 2^{-10}$	0.57, 0.2	0.83, 0.9
Overall	-	-	-	0.43 ± 0.20	0.71 ± 0.14

Table 5.10: Training parameters and performance obtained for each patient with the Dynamic Weighted Ensemble and the Add-One-Forget-One data partitioning and iterative retraining method.

Dynamic Weighted Ensemble

Patient	SOP	k	C	SS_{sample}	SP_{sample}
402	10, 15	10, 10	$2^{-10}, 2^{8}$	0.69, 0.19	0.65, 0.59
8902	15, 20	10, 10	$2^{0}, 2^{6}$	0.90, 0.97	0.68, 0.78
11002	10	20	2^{-10}	0.53	0.61
16202	45, 15, 10, 30	10, 40, 20, 40	$2^{-10}, 2^{8}, 2^{8}, 2^{6}$	0.83, 0.55, 0.46, 0.21	0.61, 0.59, 0.69, 0.71
21902	40	40	2^{-10}	0.84	0.58
23902	50, 10	40, 30	$2^{-8}, 2^{0}$	0.50, 0.63	0.53, 0.53
26102	50	40	2^{6}	0.66	0.35
30802	35, 15, 20, 25, 50	40, 10, 20, 30, 10	$2^{-8}, 2^{-10}, 2^{0}, 2^{-10}, 2^{-10}$	0.92, 0.40, 0.53, 0.59, 0.74	0.78, 0.85, 0.56, 0.73, 0.46
32702	15, 20	10, 20	$2^{-2}, 2^{-10}$	0.67, 0.86	0.70, 0.57
45402	50	40	2^{-10}	0.67	0.55
46702	30, 10	40, 10	$2^{-2}, 2^{4}$	0.37, 0.66	0.69, 0.75
50802	15, 20	30, 40	$2^{-10}, 2^{-10}$	0.82, 0.75	0.70, 0.8
53402	15	20	2^{-4}	0.22	0.72
55202	10, 10, 10, 10, 50	10, 10, 10, 20, 10	$2^{-10}, 2^{-10}, 2^{-6}, 2^{-2}, 2^{-10}$	0.84, 0.63, 0.67, 0.71, 0.62	0.52, 0.83, 0.56, 0.56, 0.45
56402	25	20	2^{-10}	0.46	0.72
58602	10, 10, 20	40, 40, 40	$2^{-10}, 2^{-10}, 2^{-8}$	0.15, 0.17, 0.65	0.78, 0.87, 0.73
59102	50, 50	10, 10	$2^{-10}, 2^{-10}$	0.41, 0.73	0.42, 0.14
60002	25, 25, 50	30, 10, 40	$2^{-10}, 2^{-10}, 2^{-10}$	0.56, 0.40, 0.68	0.65, 0.67, 0.69
64702	50, 25	30, 30	$2^{-6}, 2^{-6}$	0.64, 0.71	0.46, 0.55
75202	30, 10, 10, 10	30, 20, 40, 40	$2^{-10}, 2^{-10}, 2^{-10}, 2^{2}$	0.82, 0.70, 0.84, 0.57	0.75, 0.81, 0.75, 0.57
80702	30, 30, 45	20, 20, 40	$2^{-10}, 2^{-10}, 2^{-8}$	0.51, 0.67, 0.61	0.63, 0.40, 0.51
85202	10, 15	30, 20	$2^{6}, 2^{8}$	0.25, 0.65	0.79, 0.53
93402	20, 15	10, 10	$2^{-10}, 2^{2}$	0.64, 0.56	0.43, 0.37
93902	50, 50, 40	30, 30, 10	$2^{-10}, 2^{-10}, 2^{-10}$	0.71, 0.69, 0.43	0.51, 0.72, 0.63
94402	10, 50, 25, 10	10, 10, 30, 40	$2^{4}, 2^{-10}, 2^{-6}, 2^{-10}$	0.48, 0.82, 0.29, 0.49	0.62, 0.36, 0.71, 0.86
95202	10, 40, 35, 25	10, 40, 10, 10	$2^{-10}, 2^{8}, 2^{-10}, 2^{-10}$	0.77, 0.48, 0.42, 0.74	0.43, 0.57, 0.50, 0.59
96002	10, 40, 50, 45	40, 30, 40, 20	$2^{-10}, 2^{-2}, 2^{0}, 2^{-2}$	0.84, 0.55, 0.92, 0.55	0.56, 0.45, 0.64, 0.77
98102	25, 50	40, 20	$2^{8}, 2^{-2}$	0.58, 0.63	0.32, 0.58
101702	10, 50	20, 30	$2^{-6}, 2^{8}$	0.50, 0.40	0.57, 0.67
102202	45, 25, 50, 50	30, 40, 30, 20	$2^{0}, 2^{2}, 2^{-2}, 2^{-6}$	0.12, 0.17, 0.4, 0.11	0.88, 0.66, 0.28, 0.73
104602	20, 30	10, 10	$2^{-10}, 2^{-10}$	0.51, 0.80	0.58, 0.60
109502	30	10	2^{-6}	0.34	0.71
112802	10, 10, 30	10, 30, 30	$2^{-10}, 2^{-4}, 2^{8}$	0.77, 0.36, 0.26	0.37, 0.56, 0.71
113902	15, 15, 10	20, 30, 40	$2^{-10}, 2^{-10}, 2^{-10}$	0.35, 0.60, 0.59	0.65, 0.76, 0.64
114702	50, 30, 10, 45, 50	30, 40, 20, 10, 40	$2^{-8}, 2^{6}, 2^{6}, 2^{0}, 2^{-6}$	0.46, 0.17, 0.24, 0.6, 0.84	0.41, 0.66, 0.87, 0.51, 0.51
114902	25, 45, 35, 35	10, 30, 10, 40	$2^{-10}, 2^{6}, 2^{-10}, 2^{-10}$	0.70, 03, 0.25, 0.85	0.23, 0.76, 0.89, 0.58
123902	10, 20	40, 10	$2^{-6}, 2^{-4}$	0.92, 0.31	0.80, 0.88
Overall	-	-	-	0.56 ± 0.22	0.62 ± 0.15

5.2.2 Testing phase

The seizure prediction results obtained for each patient are presented in tables 5.5, 5.11, 5.12, 5.13.

Regarding the Control (Table 5.5), the average sensitivity and FPR/h values obtained across the 37 patients were 0.63 ± 0.34 for sensitivity and 1.72 ± 2.65, respectively. Additionally, the mean sensitivity of the surrogate time series analysis was 0.26 ± 0.51 and 31 patients (83.78%) were statistically validated.

Considering Backwards-Landmark Window (Table 5.11), the average sensitivity and FPR/h values obtained were 0.75 ± 0.33 and 1.03 ± 1.00, respectively. The sensitivity of the surrogate time series analysis method averaged at 0.22 ± 0.47, where 33 patients (89.20%) achieved performance above the chance level.

The Seizure-batch Regression (Table 5.12), obtained average values of 0.64 ± 0.31 for sensitivity and 3.73 ± 15.82 for FPR/h. The mean sensitivity of the surrogate time series analysis was 0.23 ± 0.48, and 32 patients (86.49%) were statistically validated.

Finally, the Dynamic Weighted Ensemble (Table 5.13) presents results with average values of 0.69 ± 0.36 for sensitivity and 1.60 ± 2.26 for FPR/h. The sensitivity of the surrogate time series analysis averaged at 0.25 ± 0.50, and 31 patients (83.78%) were statistically validated.

Table 5.11: Testing parameters and performance obtained for each patient with the Backwards-Landmark Window and the Add-One-Forget-One data partitioning and iterative retraining method.

Backwards-Landmark Window

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	40, 40	1.00	0.00	0.00 ± 0.00	0.00	•
8902	2	20, 30	0.00	0.19	0.10 ± 0.20	0.99	
11002	1	20	1.00	0.00	0.00 ± 0.00	0.00	•
16202	4	15, 25, 50, 10	1.00	1.82	0.28 ± 0.15	0.00	•
21902	1	40	1.00	1.23	0.33 ± 0.47	0.00	•
23902	2	50, 10	1.00	2.99	0.32 ± 0.35	0.00	•
26102	1	50	1.00	2.88	0.53 ± 0.50	0.00	•
30802	5	10, 10, 50, 25, 50	0.60	0.79	0.10 ± 0.14	0.00	•
32702	2	15, 15	1.00	1.90	0.22 ± 0.31	0.00	•
45402	1	50	1.00	0.23	0.17 ± 0.37	0.00	•
46702	2	30, 10	0.50	1.46	0.23 ± 0.25	0.00	•
50802	2	15, 25	1.00	0.96	0.15 ± 0.26	0.00	•
53402	1	45	1.00	0.03	0.03 ± 0.18	0.00	•
55202	5	35, 10, 25, 10, 10	0.80	0.34	0.11 ± 0.13	0.00	•
56402	1	15	0.00	4.77	0.23 ± 0.42	1.00	
58602	3	10, 10, 20	0.67	0.73	0.16 ± 0.21	0.00	•
59102	2	20, 15	1.00	0.17	0.05 ± 0.15	0.00	•
60002	3	20, 15, 35	1.00	1.30	0.27 ± 0.22	0.00	•
64702	2	50, 15	0.50	1.49	0.37 ± 0.34	0.02	•
75202	4	30, 10, 45, 35	0.75	0.23	0.27 ± 0.14	0.00	•
80702	3	10, 50, 50	1.00	0.91	0.36 ± 0.23	0.00	•
85202	2	10, 30	0.00	1.61	0.22 ± 0.28	1.00	
93402	2	15, 20	0.50	0.11	0.03 ± 0.12	0.00	•
93902	3	30, 45, 20	1.00	1.42	0.38 ± 0.24	0.00	•
94402	4	10, 20, 35, 15	0.75	0.12	0.03 ± 0.08	0.00	•
95202	4	10, 35, 40, 10	0.75	1.83	0.35 ± 0.22	0.00	•
96002	4	40, 20, 15, 50	0.75	0.73	0.16 ± 0.16	0.00	•
98102	2	25, 10	1.00	2.37	0.40 ± 0.24	0.00	•
101702	2	30, 50	1.00	0.64	0.32 ± 0.27	0.00	•
102202	4	10, 20, 10, 15	0.25	0.36	0.06 ± 0.11	0.00	•
104602	2	30, 40	1.00	0.81	0.33 ± 0.27	0.00	•
109502	1	10	0.00	0.47	0.07 ± 0.25	0.92	
112802	3	10, 40, 25	1.00	0.83	0.18 ± 0.22	0.00	•
113902	3	50, 10, 50	0.67	1.13	0.37 ± 0.20	0.00	•
114702	5	35, 15, 15, 50, 15	0.60	0.74	0.19 ± 0.17	0.00	•
114902	4	25, 50, 15, 20	1.00	0.22	0.17 ± 0.21	0.00	•
123902	2	40, 10	0.50	0.17	0.03 ± 0.12	0.00	•
Overall	-	-	0.75 ± 0.33	1.03 ± 1.00	0.22 ± 0.47	-	33

Table 5.12: Testing parameters and performance obtained for each patient with the Seizure-batch Regression and the Add-One-Forget-One data partitioning and iterative retraining method.

Seizure-batch Regression

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	10, 10	0.00	5.07	0.17 ± 0.24	1.00	
8902	2	25, 30	0.50	0.31	0.13 ± 0.22	0.00	•
11002	1	10	0.00	11.96	0.47 ± 0.50	1.00	
16202	4	40, 10, 15, 30	0.25	0.61	0.13 ± 0.15	0.00	•
21902	1	50	1.00	0.19	0.13 ± 0.34	0.00	•
23902	2	50, 10	1.00	0.67	0.10 ± 0.24	0.00	•
26102	1	50	1.00	0.05	0.03 ± 0.18	0.00	•
30802	5	35, 15, 20, 50, 50	0.60	0.49	0.23 ± 0.17	0.00	•
32702	2	15, 25	1.00	1.81	0.30 ± 0.33	0.00	•
45402	1	15	1.00	97.86	0.60 ± 0.49	0.00	•
46702	2	40, 10	0.50	0.61	0.12 ± 0.25	0.00	•
50802	2	15, 25	1.00	0.41	0.03 ± 0.12	0.00	•
53402	1	30	1.00	0.03	0.00 ± 0.00	0.00	•
55202	5	10, 10, 10, 10, 50	0.60	1.20	0.23 ± 0.19	0.00	•
56402	1	25	1.00	1.55	0.30 ± 0.46	0.00	•
58602	3	10, 10, 20	0.33	0.49	0.13 ± 0.16	0.00	•
59102	2	50, 15	1.00	1.47	0.20 ± 0.28	0.00	•
60002	3	20, 20, 40	0.67	0.85	0.21 ± 0.20	0.00	•
64702	2	40, 25	0.50	0.66	0.27 ± 0.28	0.00	•
75202	4	30, 10, 50, 50	0.50	0.40	0.34 ± 0.14	0.00	•
80702	3	35, 50, 15	0.67	1.08	0.27 ± 0.22	0.00	•
85202	2	10, 15	1.00	0.88	0.20 ± 0.24	0.00	•
93402	2	20, 15	0.50	1.55	0.37 ± 0.31	0.01	•
93902	3	50, 50, 15	0.67	1.85	0.26 ± 0.27	0.00	•
94402	4	10, 35, 30, 10	0.25	0.83	0.21 ± 0.17	0.10	
95202	4	10, 40, 45, 25	0.75	0.42	0.12 ± 0.12	0.00	•
96002	4	25, 40, 35, 45	0.50	0.26	0.13 ± 0.14	0.00	•
98102	2	25, 35	0.50	0.50	0.12 ± 0.25	0.00	•
101702	2	35, 50	1.00	0.34	0.32 ± 0.27	0.00	•
102202	4	50, 45, 50, 50	0.75	0.22	0.07 ± 0.14	0.00	•
104602	2	25, 45	1.00	0.38	0.23 ± 0.31	0.00	•
109502	1	20	0.00	0.00	0.00 ± 0.00	-	
112802	3	10, 15, 15	0.67	1.15	0.18 ± 0.22	0.00	•
113902	3	50, 10, 40	0.33	0.55	0.21 ± 0.22	0.00	•
114702	5	40, 30, 15, 45, 50	0.60	0.62	0.33 ± 0.17	0.00	•
114902	4	25, 50, 50, 20	0.50	0.37	0.16 ± 0.18	0.00	•
123902	2	50, 35	0.50	0.18	0.08 ± 0.19	0.00	•
Overall	-	-	0.64 ± 0.31	3.73 ± 15.82	0.23 ± 0.48	-	32

Even though the sensitivity was above the one from the Surrogate, it was not considered as statistically validated as the FPR/h is extremely large (•).

Table 5.13: Testing parameters and performance obtained for each patient with the Dynamic Weighted Ensemble and the Add-One-Forget-One data partitioning and iterative retraining method.

Dynamic Weighted Ensemble

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	10, 15	1.00	2.77	0.25 ± 0.28	0.00	•
8902	2	15, 20	1.00	0.55	0.17 ± 0.24	0.00	•
11002	1	10	0.00	14.37	0.43 ± 0.50	1.00	
16202	4	45, 15, 10, 30	0.75	0.52	0.15 ± 0.15	0.00	•
21902	1	40	0.00	1.50	0.37 ± 0.48	1.00	
23902	2	50, 10	1.00	1.39	0.28 ± 0.28	0.00	•
26102	1	50	1.00	0.88	0.33 ± 0.47	0.00	•
30802	5	35, 15, 20, 25, 50	0.80	0.52	0.19 ± 0.21	0.00	•
32702	2	15, 20	1.00	1.80	0.20 ± 0.28	0.00	•
45402	1	50	1.00	0.00	0.00 ± 0.00	0.00	•
46702	2	30, 10	0.00	0.88	0.13 ± 0.26	1.00	
50802	2	15, 20	1.00	1.27	0.12 ± 0.21	0.00	•
53402	1	15	1.00	2.03	0.20 ± 0.40	0.00	•
55202	5	10, 10, 10, 10, 50	1.00	1.67	0.32 ± 0.15	0.00	•
56402	1	25	1.00	3.30	0.50 ± 0.50	0.00	•
58602	3	10, 10, 20	0.67	0.63	0.16 ± 0.19	0.00	•
59102	2	50, 50	1.00	1.21	0.43 ± 0.33	0.00	•
60002	3	25, 25, 50	1.00	1.06	0.27 ± 0.18	0.00	•
64702	2	50, 25	1.00	2.00	0.53 ± 0.36	0.00	•
75202	4	30, 10, 10, 10	0.50	0.98	0.23 ± 0.21	0.00	•
80702	3	30, 30, 45	1.00	1.19	0.52 ± 0.24	0.00	•
85202	2	10, 15	0.50	1.16	0.17 ± 0.27	0.00	•
93402	2	20, 15	0.50	1.16	0.35 ± 0.29	0.00	•
93902	3	50, 50, 40	0.67	2.54	0.53 ± 0.22	0.00	•
94402	4	10, 50, 25, 10	0.75	0.86	0.23 ± 0.13	0.00	•
95202	4	10, 40, 35, 25	1.00	1.21	0.36 ± 0.26	0.00	•
96002	4	10, 40, 50, 45	0.75	0.58	0.30 ± 0.18	0.00	•
98102	2	25, 50	1.00	0.96	0.25 ± 0.31	0.00	•
101702	2	10, 50	0.50	1.84	0.33 ± 0.3	0.00	•
102202	4	45, 25, 50, 50	0.50	0.18	0.08 ± 0.11	0.00	•
104602	2	20, 30	0.00	1.74	0.37 ± 0.41	1.00	
109502	1	30	0.00	0.00	0.00 ± 0.00	-	
112802	3	10, 10, 30	0.33	2.89	0.28 ± 0.23	0.10	
113902	3	15, 15, 10	0.67	1.31	0.22 ± 0.22	0.00	•
114702	5	50, 30, 10, 45, 50	0.60	0.52	0.26 ± 0.13	0.00	•
114902	4	25, 45, 35, 35	1.00	0.59	0.31 ± 0.19	0.00	•
123902	2	10, 20	0.00	0.98	0.05 ± 0.15	0.96	•
Overall	-	-	0.69 ± 0.36	1.60 ± 2.26	0.25 ± 0.50	-	31

5.2.3 Comparative analysis between approaches

The performance of the adopted methodologies is shown in Table 5.14. The results obtained for sensitivity and FPR/h are also shown in Figure 5.2 for a more visual overview.

In general, the average sensitivity values of the implemented approaches in this study, are very similar. They do, however, exhibit substantial standard deviations, indicating that although some patients' models could be strong predictors, others may perform poorly.

Figure 5.2 shows a large dispersion of sensitivity values but restricted for 75% of the patients and all approaches to values above 0.4.

As stated in 2.3.2.2, in the scope of presurgical monitoring, as in the case of the present study, a maximum FPR/h value of 0.15 can be seen as reasonable in terms of a real-life warning system [7]. This criterion was not overall satisfied after looking at Table 5.14.

Table 5.14: Average seizure prediction performance across all patients for each approach.

Approach	All patients			Validated patients		
	SS	**FPR/h**	**SS Surrogate**	**%**	**SS**	**FPR/h**
Control	0.63 ± 0.34	1.72 ± 2.65	0.26 ± 0.51	83.78	0.70 ± 0.27	1.69 ± 2.75
Backwards-Landmark Window	0.75 ± 0.33	1.03 ± 1.00	0.22 ± 0.47	89.19	0.81 ± 0.25	0.91 ± 0.81
Seizure-batch Regression	0.64 ± 0.31	3.73 ± 15.82	0.23 ± 0.48	86.49	0.68 ± 0.27	0.67 ± 0.51
Dynamic Weighted Ensemble	0.69 ± 0.36	1.60 ± 2.26	0.25 ± 0.50	83.78	0.79 ± 0.28	1.18 ± 0.76

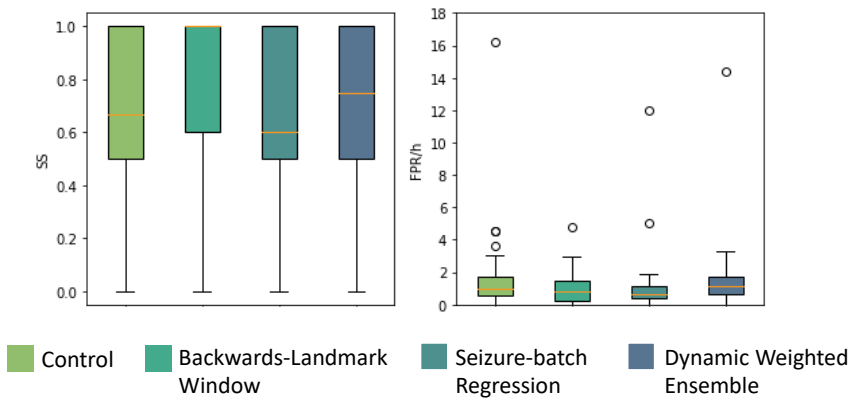

Figure 5.2: Seizure prediction performance across all patients for each approach. Seizure-batch Regression has an outlier of 97.86 FPR/h.

However, it is also apparent in Figure 5.2 that outliers had a distinct and significant impact on the average values of FPR/h. It is easy to confirm that the values obtained for some patients are considerably large by carefully examining the Tables 5.5, 5.11, 5.12, 5.13, while others are close to zero.

As shown in Figure 5.2, the Backwards-Landmark Window and the Dynamic Weighted Ensemble comprise lower FPR/h values in comparison with the Control. The Seizure-batch Regression has an outlier with an FPR/h of 97.86, which, if removed, gives the Seizure-batch Regression a smaller average FPR/h than the control.

Nevertheless, the seizure prediction models from this study raised a high number of false alarms, making them unsuitable for a warning system.

The average sensitivity for the four approaches transcended the one from surrogate time series analysis. All approaches were able to statistically validate more than 80% of the analysed patients. With the Backwards-Landmark Window, 33 out of 37 patients (89.19%) performed above the chance level.

Moreover, as expected the metrics derived for the group of validated patients were significantly better than those for all patients. But even for these patients, the FPR/h values are generally higher than the desirable.

In general, the Backwards-Landmark Window may be the best approach for the 37 studied patients, as it had the highest sensitivity, lowest FPR/h, and the most significant percentage of statistically validated patients.

Figure 5.3 illustrates a comparison of achieved performances for each patient between the proposed methodologies.

Twenty three (62%) patients (16202, 23902, 26102, 30802, 32702, 50802, 53402, 55202, 58602, 59102, 60002, 64702, 75202, 80702, 93902, 95202, 96002, 98102, 101702, 102202, 113902, 114702, 114902), performed above chance level for all approaches.

In contrast, a single patient (109502) did not perform above the chance level for any of the proposed methods. This may result from more complex brain dynamics and, consequently, may be more challenging to predict seizures. Additionally, 16% of the patients (402, 8902, 11002, 85202, 98102, 123902) were able to achieve performance above the chance level with at least one of the approaches that adapted concept drifts but not with the control approach. Particularly, the Backwards-Landmark Window was the only one that statistically validated patient 11002. The best approach, the Backwards-Landmark Window, only failed to validate patient 56402 compared to the control.

5.2.4 Patient stratification

The EPILEPSIAE database includes clinical annotations in addition to the long-term records of the Electroencephalogram (EEG), as was indicated in Chapter 4. This made it possible to group individuals with similar traits to identify trends or differences in the overall results.

Seizure classification, seizure activity pattern, vigilance level, and circadian cycle were used for stratification (included in Table 4.1 and used the following rules):

- patients suffering only from Focal Onset Aware (FOA) and Focal Onset Impaired Awareness (FOIA) seizures;
- patients with pre-seizure activity annotated as rhythmic by clinicians;
- patients that only experienced seizures while awake.

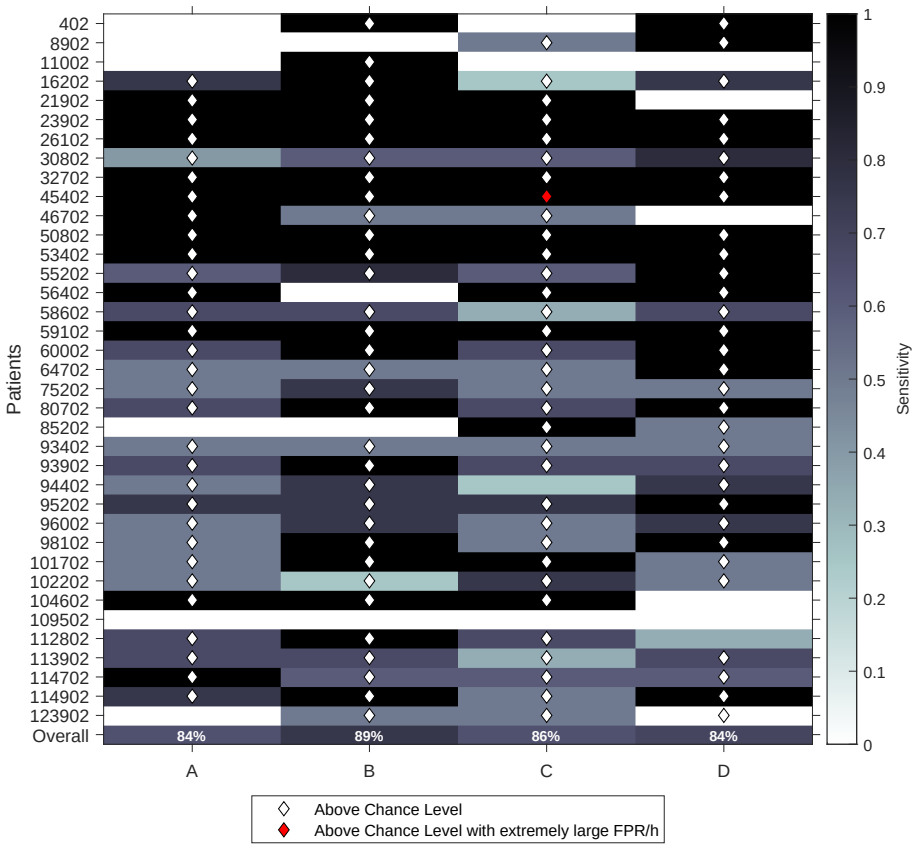

Figure 5.3: Sensitivity of seizure prediction and statistical validation results for the proposed approaches for each patient. The blue colour scale refers to the sensitivity achieved, the white diamond shape is present when statistically validated and the red diamond shape when the performance is above chance level, but not considered statistically validated because of the large FPR/h. On the overall column, for each approach, one can see the percentage of patients whose models performed above the chance level along with the average sensitivity given by the colour of the cell. A stands for Control, B for Backwards-Landmark Window, C for Seizure-batch Regression, and D for Dynamic Weighted Ensemble.

The overall values were determined for each methodology taking into account each patient group, as shown in Table 5.15, assuming that seizures with unknown classification ("UC") and unclear patterns ("?") from Table 4.1 comply with the first two criteria, respectively.

By analysing the table, it appears that the percentage of validated patients

Table 5.15: Test results for the overall set of patients, and for stratified sets of patients.

Approach	Stratification	Number of patients	SS	FPR/h	Validated patients
Control	Only FOA or FOIA seizures	25	0.63 ± 0.35	1.94 ± 3.10	84.00%
	Rhythmic activity only	36	0.64 ± 0.27	1.22 ± 0.96	86.11%
	Awake-only onset seizures	16	0.67 ± 0.37	1.55 ± 1.25	81.25%
	Overall	**37**	**0.63 ± 0.34**	**1.72 ± 2.65**	**83.78%**
Backwards-Landmark Window	Only FOA or FOIA seizures	25	0.73± 0.33	0.94 ± 0.84	88.00%
	Rhythmic activity only	36	0.74 ± 0.30	0.92 ± 0.98	91.67%
	Awake-only onset seizures	16	0.76 ± 0.38	1.33 ± 1.30	81.25%
	Overall	**37**	**0.75 ± 0.33**	**1.03 ± 1.00**	**89.19%**
Seizure-batch Regression	Only FOA or FOIA seizures	25	0.65 ± 0.31	4.95 ± 19.10	84.00%
	Rhythmic activity only	36	0.63± 0.26	0.69 ± 0.50	86.11%
	Awake-only onset seizures	16	0.64 ± 0.34	7.14± 23.45	81.25%
	Overall	**37**	**0.64 ± 0.31**	**3.73 ± 15.82**	**86.49%**
Dynamic Weighted Ensemble	Only FOA or FOIA seizures	25	0.66 ± 0.35	1.59 ± 2.69	80.00%
	Rhythmic activity only	36	0.70 ± 0.33	1.31 ± 0.78	83.33%
	Awake-only onset seizures	16	0.84 ± 0.29	1.26 ± 0.97	87.50%
	Overall	**37**	**0.69 ± 0.36**	**1.6 ± 2.26**	**83.78%**

only improved for the Control in the FOA or FOIA seizure stratified groups. For the second group, composed of patients with rhythmic activity patterns in their seizure, the ratio of validated patients was higher for the Control and the Backwards-Landmark Window. For the group of patients that only had seizures while awake, the percentage of validated patients was higher only for the Dynamic Weighted Ensemble.

5.2.5 Concept drift adaptation analysis

To understand how the models and the preictal period change after each seizure, the selected SVM cost, features, scalp EEG channels and SOP duration were evaluated for the best approach (Backwards-Landmark Window). For this reason, the relative frequencies of the selected parameters overall patients were calculated and are presented in figure 5.4. However, in general, no conclusion was obtained. Therefore, the three patients (30802, 55202, 114702) with eight seizures were selected to evaluate if conclusions could be drawn individually. Their relative frequencies of the set parameters are shown in figures 5.5, 5.6 and 5.7.

Starting with patient 30802, a clear preference for the SVM cost of 2^{-10} is noticeable, but the preictal duration after the 5^{th} seizure shifts from 10 to 50 minutes. Nevertheless, the number of selected features varied, but there was a preference for 40. It also was verified that for this patient, the algorithm selected mostly ratio-related features (see Figure 5.5). Initially, the alpha, beta, and gamma bands showed more importance, but after the 6^{th} seizure, the more important were the delta, theta and gamma. Also, some Hjörth parameters were selected from the 6^{th} seizure onward. Regarding the electrode selection, all of them were chosen, ones more frequently than others.

Continuing to patient 55202, as seizures occur, the preictal period tends to a duration of 10 minutes, and the SVM costs are in the majority to the power of -6 and -10. Nonetheless, the algorithm, in addition to never selecting the channels F4 and P8 (see Figure 5.6), the selected channels overall vary significantly between

seizures, this may highlight different seizure generating processes. For all, except the 6^{th}, evaluated seizures, a predominance of the relative power features of the beta, gamma1, gamma2, and gamma3 bands is noticeable. Also, some Hjörth parameters were selected. For the 6^{th} seizure in particular, no relative power features were selected, but the ratios between the beta, gamma1, gamma2, and gamma3 bands, were chosen instead. Also, the number of selected features varied between 10 and 30, but there was a preference for 10 and 30.

Finally, for patient 114702, there is also a clear preference for the SVM cost of 2^{-10}; the preictal duration is usually 15 minutes but is once 35 and other time 50 minutes. Regarding the electrode selection, all of them were chosen, ones more frequently than others. In particular, for the 7^{th} seizure, only two electrodes from the temporal lobes were selected (see Figure 5.7). However, for features other than a predominance for measures related to the theta, beta, gamma1, gamma2, and gamma3 bands, other features were selected, such as Hjörth parameters and wavelets, where five energy levels were kept. Also, the number of selected features varied between 10 and 40, but there was a preference for 30 and 40.

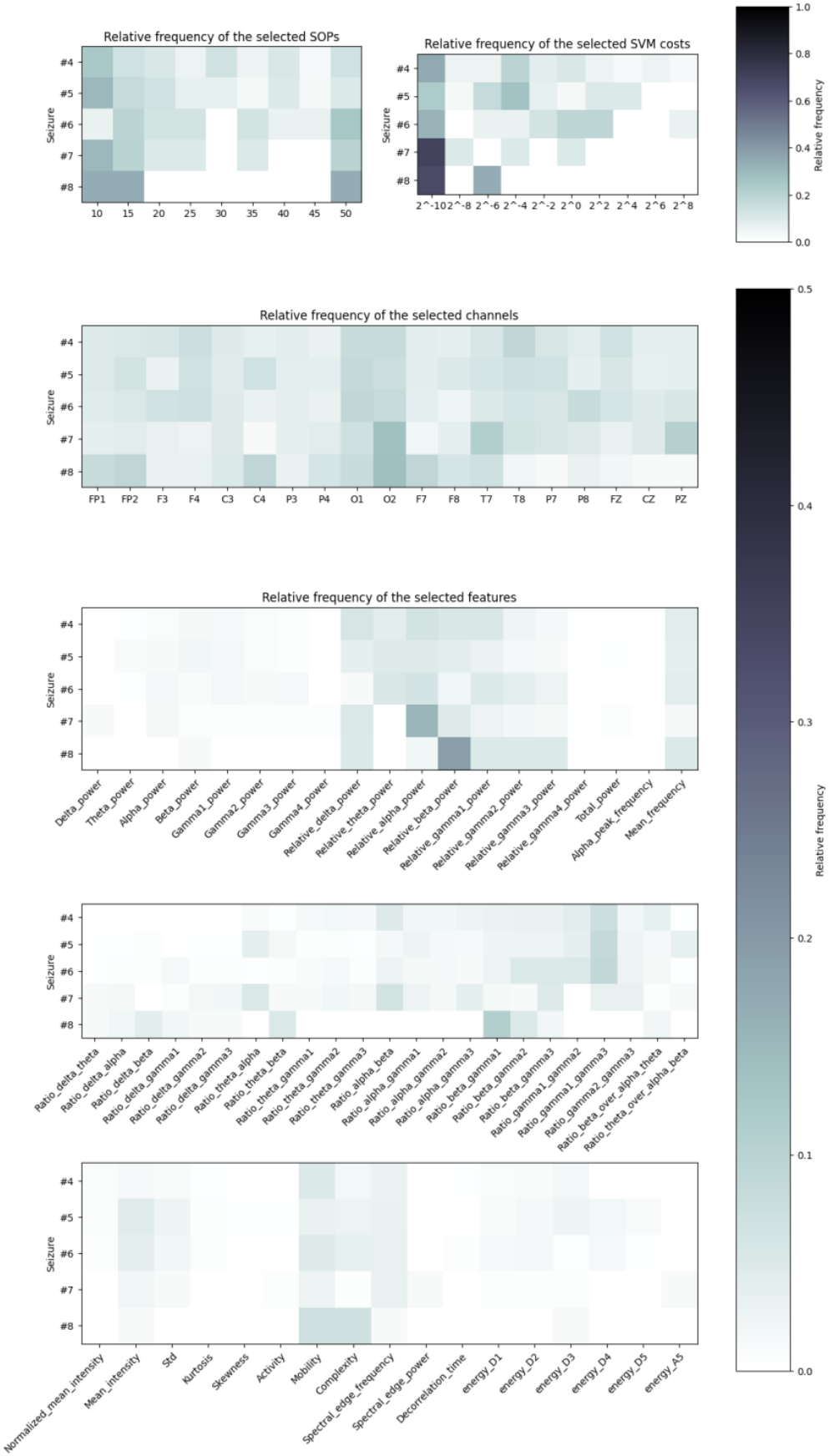

Figure 5.4: Relative frequency of the selected SOP duration, SVM costs, channels, and features for all 37 patients.

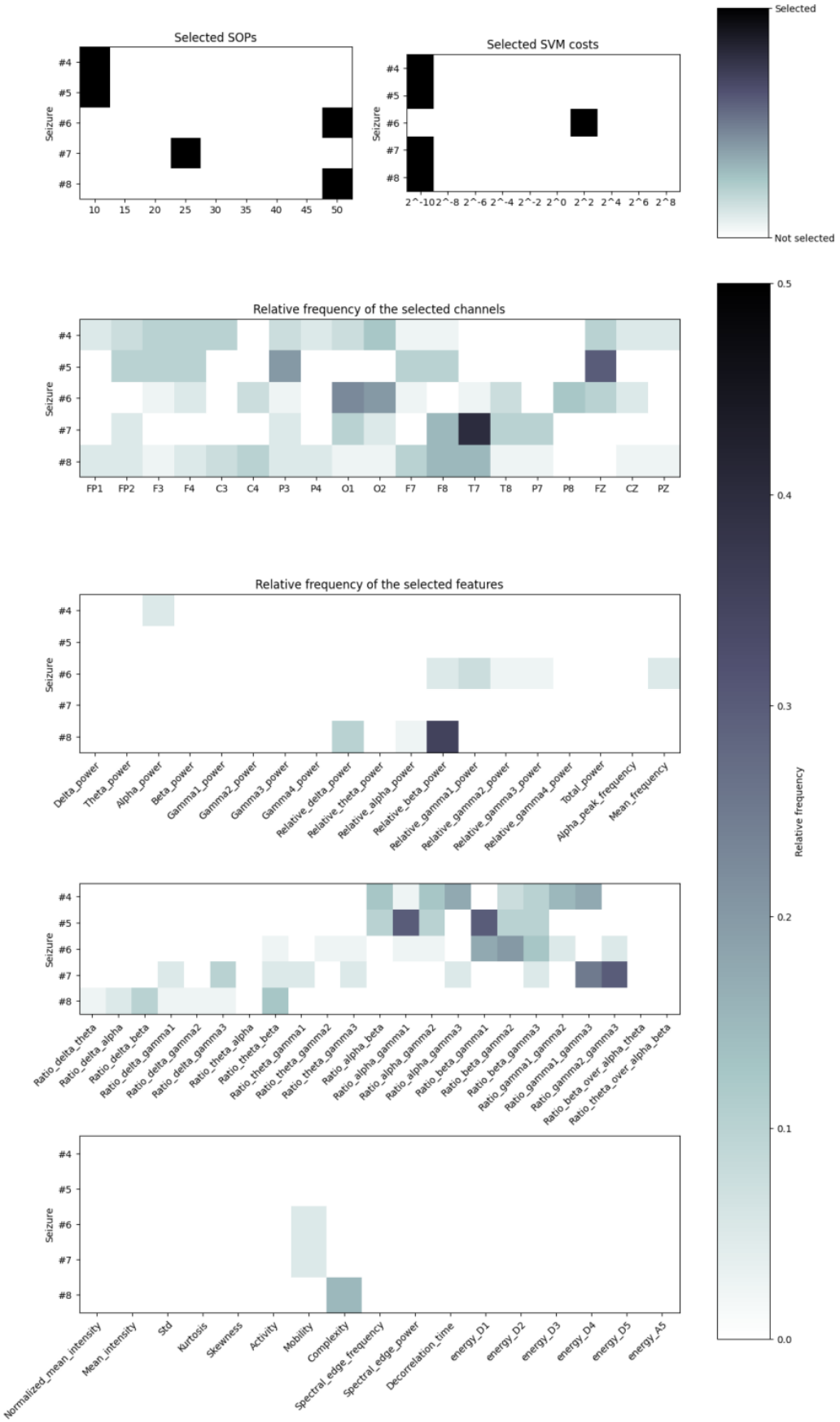

Figure 5.5: Relative frequency of the selected SOP duration, SVM costs, channels, and features for patient 30802.

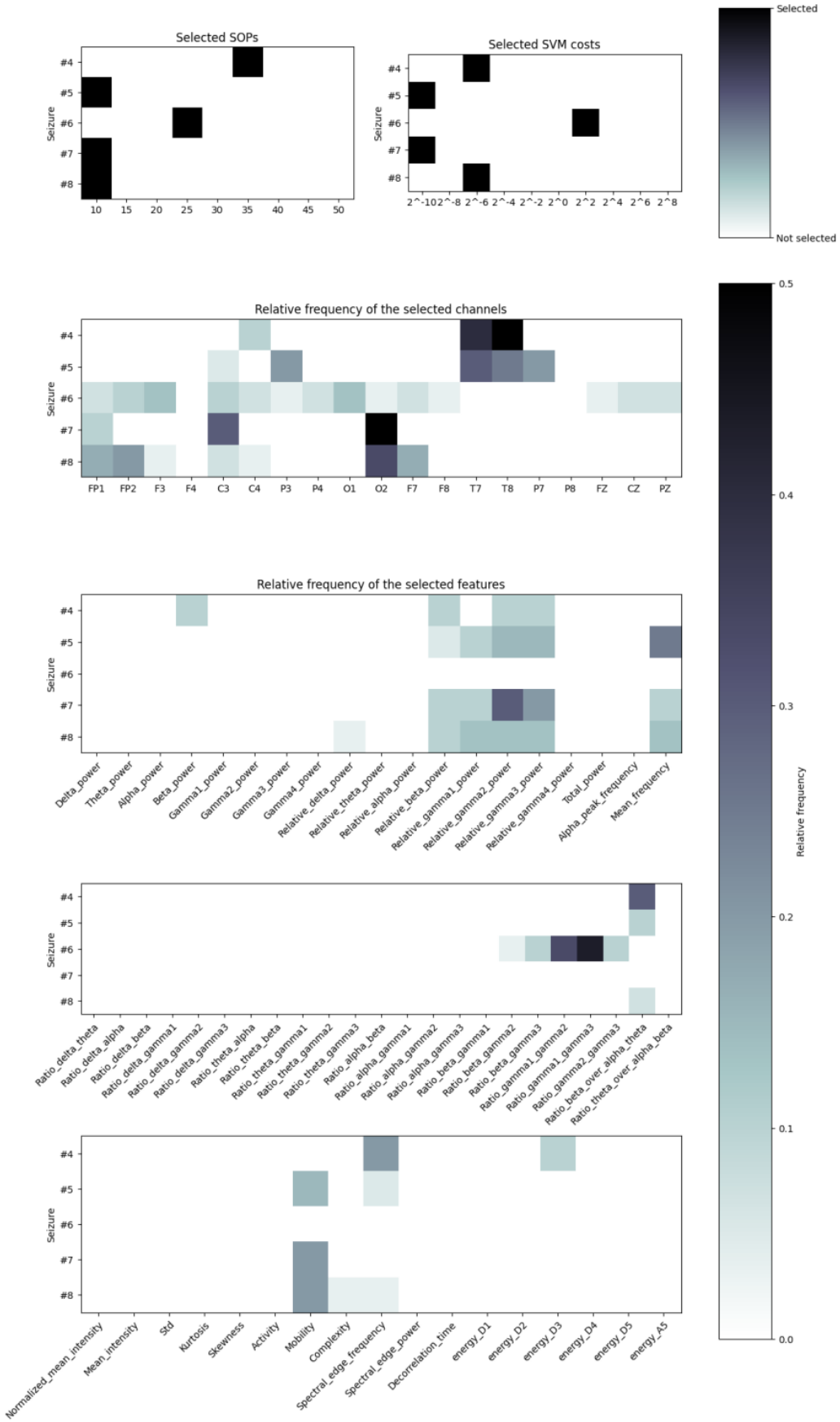

Figure 5.6: Relative frequency of the selected SOP duration, SVM costs, channels, and features for patient 55202.

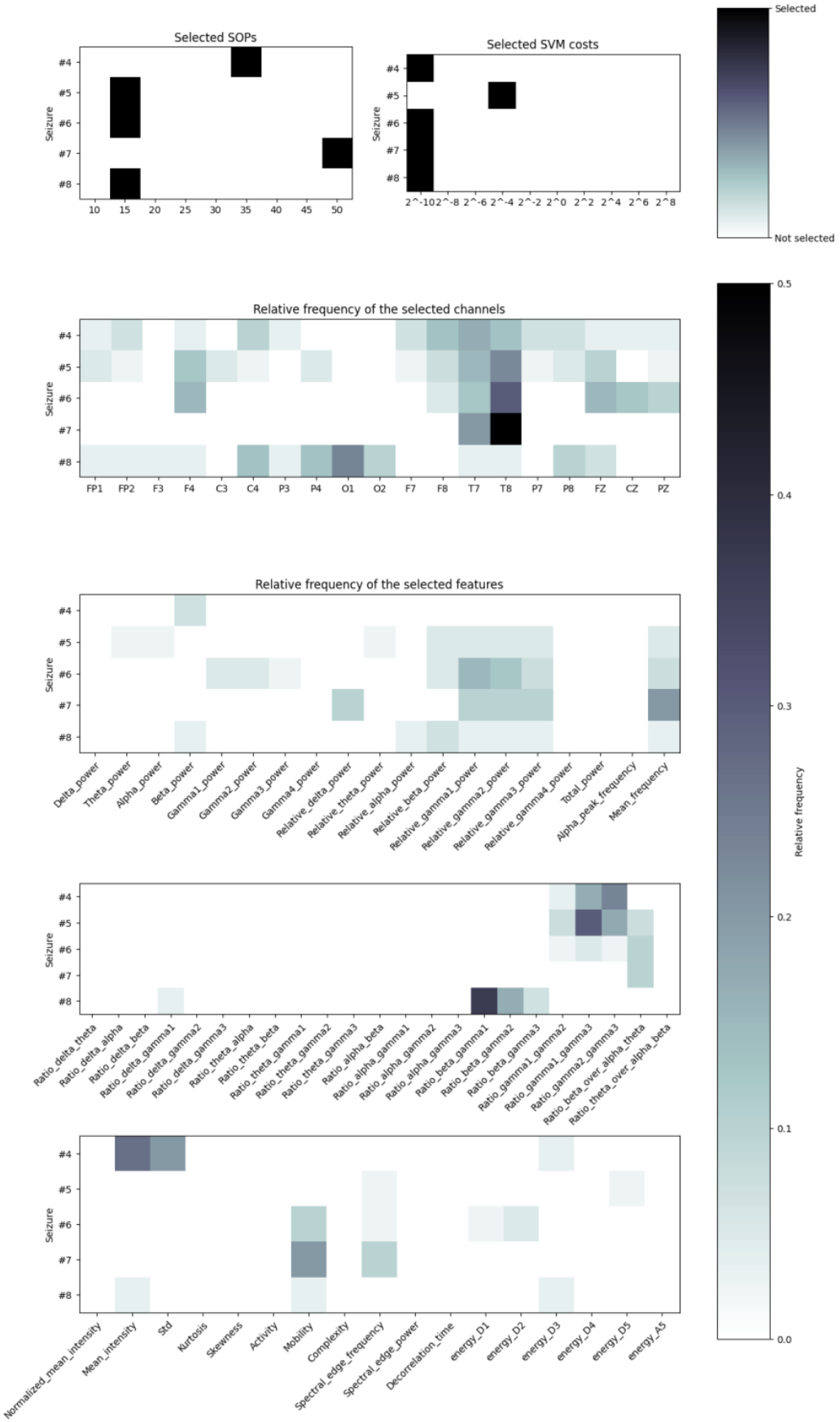

Figure 5.7: Relative frequency of the selected SOP duration, SVM costs, channels, and features for patient 114702.

Chapter 6

Discussion

In this chapter, the most pertinent topics touched in this thesis are here discussed. Section 6.1 provides a detailed discussion about the first analysis on the influence of retraining models, Section 6.2 discusses the obtained findings for the second analysis. In Section 6.3 a comparison with other studies is made, Section 6.4 discusses this thesis limitations, and Section 6.5 provides some final reflections.

6.1 Iterative retraining

The results for the first analysis suggest that retraining the Support Vector Machine (SVM) after a new seizure occurs improves seizure prediction performance. It is believed that the set of features is re-selected after each new seizure. Therefore, handling concept drifts by accounting for the evolving brain dynamics instead of assuming that the first three seizures belong to a static dataset. Moreover, running one execution of the Control algorithm takes about 6h26min or 12h32min for the Add-One-Forget-One and Chronological methods, respectively.

6.2 Concept drift adaptation

The results of the second analysis suggest that taking into account changes in context within the interictal period preceding seizures while retraining the model after each seizure increases the classifier's performance. This is indicated because the two best-performing approaches (Backwards-Landmark Window and Dynamic Weighted Ensemble), out of the four, are the only ones to adapt to changes of context within each seizure. In contrast, the Control and Seizure-batch Regression do not account for context changes within seizures, only between them. Furthermore, running one execution of the proposed approaches for the Add-One-Forget-One method takes approximately 1h6min, 1h16min, and 2h20min for the Backwards-Landmark Window, Seizure-batch Regression and Dynamic Weighted Ensemble, respectively. Since the

running time of each of the three proposed methods is shorter than the examined inter-seizure independence period (4h30min), all can be considered for real-life applications. These three methods are computationally faster than the Control, as the Control builds 31 models to combat class imbalance while the others only train one.

Additionally, a 10-minute intervention time was chosen to grant the patient enough time to take necessary measures for a seizure [7]. This time frame is also suitable for administering diazepam rectal gel, which takes 5–10 minutes to start working.

It is also essential to discuss the obtained preictal periods and how the classifiers evolve. The results presented in figure 5.4 suggest that shorter Seizure Occurrence Periods (SOPs) may be more frequent (10 and 15 minutes), but its also noticeable a relatively high frequency for SOPs of 50 minutes. There could be a propensity for the algorithms to choose long SOPs in an attempt to combat class imbalance, as the preictal period has significantly less samples than the interictal. Long SOPs are consistent with what has been observed in the literature; hence, there are increasingly more studies with longer SOPs. But does it make sense, given an application or the patients' will? Patients prefer shorter SOPs since they have less impact on their day-to-day activities when a seizure is imminent [164].

Across all patients, a particular subset of more discriminatory features does not exist, but the delta, alpha, and gamma frequency bands appear to be more important. On the one hand, measures extracted from the alpha waves have a higher relative frequency probably because they are the most prominent brain rhythm. On the other hand, because the scalp Electroencephalogram (EEG) may not completely capture gamma rhythms, the relatively high frequency of gamma band features may cause some scepticism since muscle artefacts could cause them. This findings lead us to believe that the patient may have muscular jerks due to pre-seizure dynamics recorded by gamma-related features.

6.3 Comparative analysis with other studies

The achieved findings for the seizure prediction pipeline may be compared with the results of earlier studies shown in Chapter 3. Five studies using the EPILEPSIAE database and applied statistical validation were chosen. The performance of the selected studies and the Control and the best approach implemented in the present are shown in Table 6.1.

By observing the performances of the selected studies, it is notable that the best-proposed methodology (Backwards-Landmark Window) achieved the highest sensitivity and number of statistically validated patients. Although the False Positive Rate per Hour (FPR/h) obtained were better for all of the other studies.

The random predictor was used in all of the selected studies to perform statistical validation except Pinto et al. [82,89], who performed a surrogate time series analysis,

Table 6.1: Seizure prediction performance for studies under comparison.

Study	Number of patients	SS	FPR/h	Validated patients
Pinto et al. [89] (2022)	93	0.16	0.21	32.00%
Pinto et al. [82] (2021)	19	0.37	0.79	32.00%
Direito et al. [83] (2017)	216	0.38	0.20	11.11%
Rasekhi et al. [85] (2015)	10	0.61	0.11	80.00%
Alvarado-Rojas et al. [69] (2014)	53	0.47	0.94	13.21%
Control + Control partitioning	37	0.35	1.88	45.95%
Control + Add-One-Forget-One	37	0.63	1.72	83.78%
Best proposed methodology (Backwards-Landmark Window)	37	0.75	1.03	89.19%

presenting a statistical validation of 32%. Direito et al. [83], and Alvarado-Rojas et al. [69] only reached performances above the chance level for 11.11% and 13.21%, respectively. Significantly lower than best achieved in this work (89.19%).

Alike the proposed methodology, the studies under comparison tested a range of SOP values and selected for each patient the best duration. While Rasekhi et al. [85] chose the SOP based on the best testing performance, the selection in the current study and Pinto et al. [82] were based on a specified training metric. However, in Rasekhi et al. [85], since the model parameters are selected based on the test findings, which are a priori unknown, it may result in a bias of the reported results and interference of real-life applications. Additionally, Direito et al. [83] took into account an Seizure Prediction Horizon (SPH) length of 10 seconds, which is inappropriate for a warning system as it does not give the patient enough time to take preventative measures.

It is important to note that comparing studies becomes a difficult task when there is high variability in the chosen patient population and the large variety of available parameters and alternatives incorporated throughout the process.

6.4 Limitations

To reduce computation time, an early-stopping mechanism was implemented in the Backwards-Landmark Window; this way, during the search for the more adequate window size, if the window estimate did not improve during the last 12 hours of recordings, the algorithm would stop. Stopping the search before going through the entire dataset is a limitation because there is no way of knowing that a window with a better estimate would not appear. Also, other parameters, such as the 1-hour window step for the Backwards-Landmark Window and the 1-hour window size for Dynamic Weighted Ensemble, were reasonably decided based on the computation time without being adjusted based on the test outcome. Additionally, on average, 2.59 seizures were tested per patient, and 7 of the 37 patients only had one seizure for testing.

Moreover, the data used is from presurgical monitoring circumstances. Patients in clinics experience medication withdrawal and sleep deprivation, which might lead

to more seizures that may not be characteristic of everyday life. Therefore, this study has to be reproduced using ultra long-term recordings made throughout the patients' daily lives, such as the data from the Neurovista, to access seizure prediction performance fully.

6.5 Final reflections

This work contributed to epilepsy seizure prediction by providing a complete pipeline for patient-specific prediction models while addressing changes in context. As expected, the proposed pipeline showed a significant performance improvement, but it was in presurgical monitoring, meaning there was a shorter inter-seizure time. As mentioned in Section 2.3.2.2, under normal conditions, patients with pharmacorefractory focal epilepsy have a mean seizure frequency of about three seizures per month. Thus, in those situations adopting the Add-One-Forget-One method to retrain the model would mean that approximately one month of recordings would be used. That would significantly increase the inter-seizure time and the model training time. But only after reproducing this study using ultra long-term recordings can it be inferred whether the training time is still inferior than the inter-seizure time, making it applicable in a medical device. Also, only then can it be evaluated whether parameters such as the window step or size should remain 1 hour or be increased to hours or even days.

Chapter 7

Conclusion

In this thesis, efforts were made to assess the impact of concept drifts in seizure prediction. To this end, two independent analysis were done. One to access the impact of retraining the algorithm through hard-rules, and another to understand the impact of concept drift adaptation techniques. For the first analysis, two methods to retrain the patient-specific models were used, the Add-One-Forget-One and Chronological, these were compared with the Control partitioning method, and the best one was used on the second analysis. On the second analysis, three seizure prediction approaches (Backwards-Landmark Window, Seizure-batch Regression, and Dynamic Weighted Ensemble) able to adapt to concept drifts were proposed and compared to the Control.

Between the two retraining methods, Add-One-Forget-One achieved better performance in terms of sensitivity, False Positive Rate per Hour (FPR/h) and percentage of validated patients for the four approaches.

Overall, the proposed approaches that integrated concept drift adaptation techniques outperformed the Control, evidencing the relevance of adapting to concept drifts to predict seizures. In particular, the best-proposed methodology (Backwards-Landmark Window) achieved results of 0.75 ± 0.33 for sensitivity, 1.03 ± 1.00 for FPR/h and 89.19% of the patients were statistically validated. This approach aimed to detect the data associated with last concept before the last labelled seizure. Therefore, taking into account changes in the hidden context over time may be the path to follow in the seizure prediction.

The FPR/h values found in this study are not appropriate for real-world warning systems. Still, depending on the effects on the patient's health, their application in intervention systems (such as electrical stimulation) might be. However, further research needs to be done to determine the highest FPR/h deemed appropriate for each intervention technique.

It is important to stress that because the current study was based on data from presurgical monitoring, it can only be considered a proof of concept. Therefore, to determine its relevance in a real-world scenario, ultra long-term data gathered from

conditions encountered in daily life must be used.

Concerning future work, the use of the same methodology in various datasets would enable comparisons to earlier seizure prediction studies.Ultra long-term data with naturally occurring seizures, like the one published by Cook et al. [4], would be extremely important for developing and validating novel prediction algorithms, advancing wider clinical acceptability, and evaluating current methodologies. In the words of Mormann et al. [20], the long and winding road continues.

References

[1] 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. xv, 6, 7, 8

[2] 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. xv, 8

[3] P. Ryvlin, S. Rheims, L. J. Hirsch, A. Sokolov, and L. Jehi, "Neuromodulation in epilepsy: state-of-the-art approved therapies," *The Lancet Neurology*, vol. 20, no. 12, pp. 1038–1047, 2021. xv, 10, 11

[4] 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. xv, 2, 12, 17, 32, 100

[5] A. Varsavsky, I. Mareels, and M. Cook, *Epileptic seizures and the EEG: measurement, models, detection and prediction.* Taylor & Francis, 2011. xv, 12, 14, 15, 16, 35, 41

[6] N. Moghim and D. W. Corne, "Predicting epileptic seizures in advance," *PLoS One*, vol. 9, no. 6, 2014. xv, 17, 18, 32, 33, 35, 36, 40, 41, 42, 44, 120, 121, 123

[7] 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. xv, 2, 18, 19, 20, 21, 22, 23, 25, 44, 86, 96

[8] 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*, pp. 237–248, 2008. xvi, 22, 23, 24

[9] 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. xvi, 1, 25, 26, 47, 49, 50, 129

[10] I. Žliobaitė, "Learning under concept drift: an overview," *arXiv preprint arXiv:1010.4784*, 2010. xvi, 26

[11] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM computing surveys (CSUR)*, vol. 46, no. 4, pp. 1–37, 2014. xvi, 26

[12] E. Bou 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. xvi, 2, 3, 16, 17, 20, 21, 22, 24, 29, 30, 32, 34, 35, 37, 40, 41, 43, 44, 119, 122, 123, 124

[13] I. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira, "Discussion and review on evolving data streams and concept drift adapting," *Evolving systems*, vol. 9, no. 1, pp. 1–23, 2018. xvi, xvii, 46, 47, 48, 49, 50, 51, 125, 126, 127, 128, 129, 130, 131, 132

[14] World Health Organization, "Epilepsy - Detailed Fact Sheets," 2018. 1, 5

[15] E. J. Jr, "What can we do for people with drug-resistant epilepsy? the 2016 wartenberg lecture," *American Academy of Neurology*, vol. 87, no. 23, pp. 2483–2489, 2016. 1, 9, 10

[16] C. Rathore and K. Radhakrishnan, "Concept of epilepsy surgery and presurgical evaluation," *Epileptic disorders*, vol. 17, no. 1, pp. 19–31, 2015. 1, 10

[17] 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. 2, 3, 16, 24, 25, 29, 119, 120, 121, 122

[18] P. J. Karoly, H. Ung, D. B. Grayden, L. Kuhlmann, K. Leyde, M. J. Cook, and D. R. Freestone, "The circadian profile of epilepsy improves seizure forecasting," *Brain*, vol. 140, no. 8, pp. 2169–2182, 2017. 2, 3, 17, 25, 33, 35, 39, 40, 41, 42, 44, 46

[19] 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. 2, 3, 16, 25

[20] 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. 2, 3, 16, 17, 18, 22, 25, 26, 29, 37, 100, 119, 122, 123, 124

[21] 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. 3, 25, 48, 49, 50, 54, 64, 130, 131

[22] S. Khan, L. Nobili, R. Khatami, T. Loddenkemper, C. Cajochen, D.-J. Dijk, and S. H. Eriksson, "Circadian rhythm and epilepsy," *The Lancet Neurology*, vol. 17, no. 12, pp. 1098–1108, 2018. 3, 25, 46

[23] R. Klinkenberg and T. Joachims, "Detecting concept drift with support vector machines.," in *ICML*, pp. 487–494, 2000. 3, 48, 54, 63, 127

[24] K. Yeon, M. S. Song, Y. Kim, H. Choi, and C. Park, "Model averaging via penalized regression for tracking concept drift," *Journal of Computational and Graphical Statistics*, vol. 19, no. 2, pp. 457–473, 2010. 3, 54, 64

[25] J. Klatt, H. Feldwisch-Drentrup, M. Ihle, V. Navarro, M. Neufang, C. Teixeira, C. Adam, M. Valderrama, C. Alvarado-Rojas, A. Witon, M. Le Van Quyen, F. Sales, A. Dourado, J. Timmer, A. Schulze-Bonhage, and B. Schelter, "The EPILEPSIAE database: An extensive electroencephalography database of epilepsy patients," *Epilepsia*, vol. 53, no. 9, pp. 1669–1676, 2012. 5, 32, 55

[26] R. S. Fisher, W. V. E. Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. Engel, "Epileptic Seizures and Epilepsy: Definitions Proposed by the International League Against Epilepsy (ILAE) and the International Bureau for Epilepsy (IBE)," vol. 46, no. 4, pp. 1–3, 2005. 5

[27] P. Jiruska, M. De Curtis, J. G. Jefferys, C. A. Schevon, S. J. Schiff, and K. Schindler, "Synchronization and desynchronization in epilepsy: controversies and hypotheses," *The Journal of physiology*, vol. 591, no. 4, pp. 787–797, 2013. 5

[28] S. Z. Wu Y, Liu D, "Neuronal networks and energy bursts in epilepsy," *Journal of Physiology*, vol. 287, pp. 175–186, 2015. 5

[29] 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. 6

[30] I. L. A. E. (ILAE), "Epilepsy syndromes," 2019. Last accessed 02 August 2019. 8

[31] N. I. of Neurological Disorders and S. (NIHDS), "The epilepsies and seizures: Hope through research," 2019. Last accessed 02 August 2019. 9

[32] C. J. H. Klein, "Temporal lobe epilepsy." Last accessed: 2021-07-12. 9

[33] W. H. Organization *et al.*, *Epilepsy: a public health imperative*. World Health Organization, 2019. 9

[34] Y. Wang and Z. Chen, "An update for epilepsy research and antiepileptic drug development: Toward precise circuit therapy," *Pharmacology & therapeutics*, vol. 201, pp. 77–93, 2019. 9

[35] P. Kwan, A. Arzimanoglou, A. T. Berg, M. J. Brodie, W. Allen Hauser, G. Mathern, S. L. Moshé, E. Perucca, S. Wiebe, and J. French, "Definition of drug resistant epilepsy: consensus proposal by the ad hoc task force of the ilae commission on therapeutic strategies," *Epilepsia*, vol. 51, no. 6, pp. 1069–1077, 2010. 9

[36] P. Ryvlin, J. H. Cross, and S. Rheims, "Epilepsy surgery in children and adults," *The Lancet Neurology*, vol. 13, no. 11, pp. 1114–1126, 2014. 10

[37] J. Engel Jr, "The current place of epilepsy surgery," *Current opinion in neurology*, vol. 31, no. 2, p. 192, 2018. 10

[38] O. Devinsky, A. Vezzani, T. J. O'Brien, N. Jette, I. E. Scheffer, M. de Curtis, and P. Perucca, "Epilepsy," *Nature Reviews Disease Primers*, vol. 4, p. 18024, 6 2018. 10

[39] N. Rincon, D. Barr, and N. Velez-Ruiz, "Neuromodulation in drug resistant epilepsy," *Aging and disease*, vol. 12, no. 4, p. 1070, 2021. 10

[40] M. D. Bigelow and A. Z. Kouzani, "Neural stimulation systems for the control of refractory epilepsy: a review," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, pp. 1–17, 2019. 10

[41] V. Krishna, F. Sammartino, N. K. K. King, R. Q. Y. So, and R. Wennberg, "Neuromodulation for epilepsy," *Neurosurgery Clinics*, vol. 27, no. 1, pp. 123–131, 2016. 10

[42] P. Boon, E. De Cock, A. Mertens, and E. Trinka, "Neurostimulation for drug-resistant epilepsy: a systematic review of clinical evidence for efficacy, safety, contraindications and predictors for response," *Current opinion in neurology*, vol. 31, no. 2, pp. 198–210, 2018. 10

[43] S. Beniczky, P. Karoly, E. Nurse, P. Ryvlin, and M. Cook, "Machine learning and wearable devices of the future," *Epilepsia*, vol. 62, pp. S116–S124, 2021. 10, 11

[44] M. O. Baud, T. Proix, V. R. Rao, and K. Schindler, "Chance and risk in epilepsy," *Current opinion in neurology*, vol. 33, no. 2, pp. 163–172, 2020. 10

[45] M. Gaínza-Lein, R. Benjamin, C. Stredny, M. McGurl, K. Kapur, and T. Loddenkemper, "Rescue medications in epilepsy patients: a family perspective," *Seizure*, vol. 52, pp. 188–194, 2017. 10

[46] J. Riss, J. Cloyd, J. Gates, and S. Collins, "Benzodiazepines in epilepsy: pharmacology and pharmacokinetics," *Acta neurologica scandinavica*, vol. 118, no. 2, pp. 69–86, 2008. 11

[47] M. Nasseri, E. Nurse, M. Glasstetter, S. Böttcher, N. M. Gregg, A. Laks Nandakumar, B. Joseph, T. Pal Attia, P. F. Viana, E. Bruno, *et al.*, "Signal quality and patient experience with wearable devices for epilepsy management," *Epilepsia*, vol. 61, pp. S25–S35, 2020. 11, 12

[48] M. O. Baud, T. Proix, N. M. Gregg, B. H. Brinkmann, E. S. Nurse, M. J. Cook, and P. J. Karoly, "Seizure forecasting: bifurcations in the long and winding road," *Epilepsia*, 2022. 12

[49] K. Gadhoumi, J. M. Lina, F. Mormann, and J. Gotman, "Seizure Prediction for Therapeutic Devices: A Review," *Journal of Neuroscience Methods*, vol. 260, no. 029, pp. 270–282, 2016. 12

[50] S. Ramgopal, S. Thome-Souza, M. Jackson, N. E. Kadish, I. S. Fernández, J. Klehm, W. Bosl, C. Reinsberger, S. Schachter, and T. Loddenkemper, "Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy," *Epilepsy & behavior*, vol. 37, pp. 291–307, 2014. 12

[51] B. Brinkmann, E. Nurse, M. Nasseri, P. F. Viana, P. Karoly, T. P. Attia, N. Gregg, B. Joseph, C. Grzeskowiak, M. Dümpelmann, *et al.*, "Seizure forecasting and detection with wearable devices and subcutaneous eeg-a practical seizure gauge (n2. 004)," 2022. 12

[52] J. W. C. Medithe and U. R. Nelakuditi, "Study of normal and abnormal eeg," *2016 3rd International Conference on Advanced Computing and Communication Systems (ICACCS)*, vol. 01, pp. 1–4, 2016. 12, 13, 14

[53] I. Osorio, H. P. Zaveri, M. G. Frei, and S. Arthurs, *Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics*. CRC press, 2019. 12, 13, 14, 15, 16, 17, 20, 21

[54] T. Kirschstein and R. Köhling, "What is the source of the eeg?," *Clinical EEG and neuroscience*, vol. 40, no. 3, pp. 146–149, 2009. 12

[55] 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. 13

[56] A. S. Oliveira, B. R. Schlink, W. D. Hairston, P. König, and D. P. Ferris, "Induction and separation of motion artifacts in eeg data using a mobile phantom head device," *Journal of neural engineering*, vol. 13, no. 3, p. 036014, 2016. 14

[57] T. N. Alotaiby, S. A. Alshebeili, T. Alshawi, I. Ahmad, and F. E. Abd El-Samie, "Eeg seizure detection and prediction algorithms: a survey," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–21, 2014. 14, 15, 16

[58] 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. 14, 16

[59] M. Sazgar and M. G. Young, "Overview of eeg, electrode placement, and montages," in *Absolute epilepsy and EEG rotation review*, pp. 117–125, Springer, 2019. 15, 16

[60] J. N. Acharya and V. J. Acharya, "Overview of eeg montages and principles of localization," *Journal of Clinical Neurophysiology*, vol. 36, no. 5, pp. 325–329, 2019. 15, 16

[61] J. Parvizi and S. Kastner, "Promises and limitations of human intracranial electroencephalography," *Nature neuroscience*, vol. 21, no. 4, pp. 474–483, 2018. 15, 16

[62] E. Foundation, "Tests Before Surgery." https://www.epilepsy.com/learn/treating-seizures-and-epilepsy/surgery/tests-surgery/video-eeg-monitoring-invasive-electrodes., last accessed: 18 January 2022. 15, 16

[63] J. Duun-Henriksen, M. Baud, M. P. Richardson, M. Cook, G. Kouvas, J. M. Heasman, D. Friedman, J. Peltola, I. C. Zibrandtsen, and T. W. Kjaer, "A new era in electroencephalographic monitoring? subscalp devices for ultra–long-term recordings," *Epilepsia*, vol. 61, no. 9, pp. 1805–1817, 2020. 15

[64] K. M. Grande, S. K. Ihnen, and R. Arya, "Electrical stimulation mapping of brain function: a comparison of subdural electrodes and stereo-eeg," *Frontiers in Human Neuroscience*, p. 538, 2020. 15

[65] A. K. Health, "Invasive electroencephalography (EEG) monitoring before epilepsy surgery," 2017. 15

[66] 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. 16

[67] H.-H. Chen and V. Cherkassky, "Performance metrics for online seizure prediction," *Neural Networks*, vol. 128, pp. 22–32, 2020. 17

[68] R. E. Stirling, D. B. Grayden, W. D'Souza, M. J. Cook, E. Nurse, D. R. Freestone, D. E. Payne, B. H. Brinkmann, T. Pal Attia, P. F. Viana, *et al.*, "Forecasting seizure likelihood with wearable technology," *Frontiers in neurology*, p. 1170, 2021. 17, 32, 33, 35, 36, 38, 39, 40, 41, 42, 43, 44

[69] 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, no. 1, pp. 1–9, 2014. 17, 32, 33, 35, 36, 39, 41, 42, 43, 97

[70] M. Nasseri, T. Pal Attia, B. Joseph, N. M. Gregg, E. S. Nurse, P. F. Viana, G. Worrell, M. Dümpelmann, M. P. Richardson, D. R. Freestone, *et al.*, "Ambulatory seizure forecasting with a wrist-worn device using long-short term memory deep learning," *Scientific reports*, vol. 11, no. 1, pp. 1–9, 2021. 17, 32, 33, 35, 36, 38, 39, 40, 41, 42, 43, 44

[71] F. Lopes, A. Leal, J. Medeiros, M. F. Pinto, A. Dourado, M. Dümpelmann, and C. Teixeira, "Automatic electroencephalogram artifact removal using deep convolutional neural networks," *IEEE Access*, vol. 9, pp. 149955–149970, 2021. 17, 36, 57

[72] M. O. Baud, J. K. Kleen, E. A. Mirro, J. C. Andrechak, D. King-Stephens, E. F. Chang, and V. R. Rao, "Multi-day rhythms modulate seizure risk in epilepsy," *Nature communications*, vol. 9, no. 1, pp. 1–10, 2018. 18, 45, 46

[73] B. Schelter, M. Winterhalder, H. F. genannt Drentrup, J. Wohlmuth, J. Nawrath, A. Brandt, A. Schulze-Bonhage, and J. Timmer, "Seizure prediction: the impact of long prediction horizons," *Epilepsy research*, vol. 73, no. 2, pp. 213–217, 2007. 19, 21

[74] 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. 22, 23

[75] 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. 22, 23, 24

[76] 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. 22, 24

[77] Y. LeCun, Y. Bengio, G. Hinton, *et al.*, "Deep learning. nature, 521 (7553), 436-444," *Google Scholar Google Scholar Cross Ref Cross Ref*, 2015. 31, 42, 43

[78] A. Aarabi and B. He, "Seizure prediction in patients with focal hippocampal epilepsy," *Clinical Neurophysiology*, vol. 128, no. 7, pp. 1299–1307, 2017. 32, 33, 35, 36, 39, 42, 123, 124

[79] T. Tamanna, M. A. Rahman, S. Sultana, M. H. Haque, and M. Z. Parvez, "Predicting seizure onset based on time-frequency analysis of eeg signals," *Chaos, Solitons & Fractals*, vol. 145, p. 110796, 2021. 32, 33, 35, 36, 39, 40, 42, 123

[80] S. M. Usman, S. Khalid, and Z. Bashir, "Epileptic seizure prediction using scalp electroencephalogram signals," *Biocybernetics and Biomedical Engineering*, vol. 41, no. 1, pp. 211–220, 2021. 32, 33, 35, 36, 38, 39, 40, 42, 44

[81] K. M. Tsiouris, V. C. Pezoulas, M. Zervakis, S. Konitsiotis, D. D. Koutsouris, and D. I. Fotiadis, "A long short-term memory deep learning network for the prediction of epileptic seizures using eeg signals," *Computers in biology and medicine*, vol. 99, pp. 24–37, 2018. 32, 33, 35, 36, 38, 39, 40, 42

[82] M. Pinto, A. Leal, F. Lopes, A. Dourado, P. Martins, C. A. Teixeira, *et al.*, "A personalized and evolutionary algorithm for interpretable eeg epilepsy seizure prediction," *Scientific reports*, vol. 11, no. 1, pp. 1–12, 2021. 32, 33, 35, 36, 39, 40, 41, 42, 43, 96, 97

[83] 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. 32, 33, 35, 36, 39, 40, 41, 42, 43, 61, 97, 119, 120, 121, 122

[84] 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. 32, 33, 35, 36, 39, 40, 41, 42, 43, 66, 121

[85] 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. 32, 33, 35, 36, 37, 39, 40, 41, 42, 43, 97, 119, 120, 121, 124

[86] 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. 32, 33, 35, 36, 38, 39, 40, 41, 42, 43, 119, 120, 121, 122

[87] 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. 32, 33, 35, 36, 38, 39, 40, 41, 42, 43, 119, 120, 121, 122

[88] A. F. Rabbi, L. Azinfar, and R. Fazel-Rezai, "Seizure prediction using adaptive neuro-fuzzy inference system," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 2100–2103, 2013. 32, 33, 35, 36, 39, 40, 41, 42, 123, 124

[89] M. Pinto, T. Coelho, A. Leal, F. Lopes, A. Dourado, P. Martins, and C. Teixeira, "Interpretable eeg seizure prediction using a multiobjective evolutionary algorithm," *Scientific reports*, vol. 12, no. 1, pp. 1–15, 2022. 32, 33, 35, 36, 39, 40, 41, 42, 43, 96, 97

[90] 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, 2018. 32, 33, 35, 36, 38, 39, 40, 42, 43, 44

[91] E. B. Assi, M. Sawan, D. K. 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. 32, 33, 35, 36, 39, 40, 41, 42, 44, 121

[92] P. Viana, M. Nasseri, J. Duun-Henriksen, A. Biondi, J. Winston, P. Martins, E. Nurse, M. Dümpelmann, G. Worrell, A. Schulze-Bonhage, *et al.*, "Seizure forecasting using minimally invasive, ultra-long-term subcutaneous eeg: Generalizable cross-patient models.," *Epilepsia*, 2022. 32, 33, 35, 36, 38, 39, 40, 41, 42, 44

[93] T. Pal Attia, P. F. Viana, M. Nasseri, J. Duun-Henriksen, A. Biondi, J. S. Winston, I. P. Martins, E. S. Nurse, M. Dümpelmann, G. A. Worrell, *et al.*, "Seizure forecasting using minimally invasive, ultra-long-term subcutaneous

eeg: Generalizable cross-patient models," *Epilepsia*, 2022. 32, 33, 35, 36, 38, 39, 40, 41, 42, 44

[94] 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. 33, 35, 36, 38, 39, 40, 42, 44

[95] 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. 33, 35, 36, 38, 39, 40, 41, 42

[96] 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. 33, 35, 36, 38, 39, 40, 42

[97] H. Adeli, S. Ghosh-Dastidar, and N. Dadmehr, "A wavelet-chaos methodology for analysis of eegs and eeg subbands to detect seizure and epilepsy," *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 2, pp. 205–211, 2007. 35

[98] A. Leal, J. Curty, F. Lopes, M. F. Pinto, A. Oliveira, F. Sales, A. M. Bianchi, M. G. Ruano, A. Dourado, J. Henriques, *et al.*, "Unsupervised eeg preictal interval identification in patients with drug-resistant epilepsy," 2022. 36

[99] U. R. Acharya, S. V. Sree, A. P. C. Alvin, and J. S. Suri, "Use of principal component analysis for automatic classification of epileptic eeg activities in wavelet framework," *Expert Systems with Applications*, vol. 39, no. 10, pp. 9072–9078, 2012. 40

[100] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE transactions on biomedical circuits and systems*, vol. 13, no. 5, pp. 804–813, 2019. 40, 43

[101] 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. 41

[102] N. D. Truong, L. Kuhlmann, M. R. Bonyadi, D. Querlioz, L. Zhou, and O. Kavehei, "Epileptic seizure forecasting with generative adversarial networks," *IEEE Access*, vol. 7, pp. 143999–144009, 2019. 41, 42

[103] 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. 41, 119, 123, 124

[104] 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. 41

[105] 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. 42, 123, 124

[106] C. Teixeira, B. Direito, M. Bandarabadi, and A. Dourado, "Output regularization of SVM seizure predictors: Kalman Filter versus the Firing Power method," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 6530–6533, 2012. 43, 44

[107] 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. 43

[108] M. Dümpelmann, "Early seizure detection for closed loop direct neurostimulation devices in epilepsy," *Journal of neural engineering*, vol. 16, no. 4, p. 041001, 2019. 45

[109] V. R. Rao, M. G Leguia, T. K. Tcheng, and M. O. Baud, "Cues for seizure timing," *Epilepsia*, vol. 62, pp. S15–S31, 2021. 45

[110] P. J. Karoly, V. R. Rao, N. M. Gregg, G. A. Worrell, C. Bernard, M. J. Cook, and M. O. Baud, "Cycles in epilepsy," *Nature Reviews Neurology*, vol. 17, no. 5, pp. 267–284, 2021. 45, 46

[111] M. G. Leguia, R. G. Andrzejak, C. Rummel, J. M. Fan, E. A. Mirro, T. K. Tcheng, V. R. Rao, and M. O. Baud, "Seizure cycles in focal epilepsy," *JAMA neurology*, vol. 78, no. 4, pp. 454–463, 2021. 45

[112] P. J. Karoly, D. M. Goldenholz, D. R. Freestone, R. E. Moss, D. B. Grayden, W. H. Theodore, and M. J. Cook, "Circadian and circaseptan rhythms in human epilepsy: a retrospective cohort study," *The Lancet Neurology*, vol. 17, no. 11, pp. 977–985, 2018. 45

[113] V. Ferastraoaru, D. M. Goldenholz, S. Chiang, R. Moss, W. H. Theodore, and S. R. Haut, "Characteristics of large patient-reported outcomes: where can one million seizures get us?," *Epilepsia open*, vol. 3, no. 3, pp. 364–373, 2018. 45

[114] D.-H. Tran, "Automated change detection and reactive clustering in multivariate streaming data," in *2019 IEEE-RIVF International Conference on Computing and Communication Technologies (RIVF)*, pp. 1–6, IEEE, 2019. 48, 49, 50, 128, 129, 131

[115] D. Martínez-Rego, D. Fernández-Francos, O. Fontenla-Romero, and A. Alonso-Betanzos, "Stream change detection via passive-aggressive classification and bernoulli cusum," *Information Sciences*, vol. 305, pp. 130–145, 2015. 48

[116] H. Toubakh and M. Sayed-Mouchaweh, "Hybrid dynamic data-driven approach for drift-like fault detection in wind turbines," *Evolving Systems*, vol. 6, no. 2, pp. 115–129, 2015. 48, 50, 131

[117] P. M. Gonçalves Jr, S. G. de Carvalho Santos, R. S. Barros, and D. C. Vieira, "A comparative study on concept drift detectors," *Expert Systems with Applications*, vol. 41, no. 18, pp. 8144–8156, 2014. 48, 50, 51, 131, 132

[118] D. Mejri, M. Limam, and C. Weihs, "Adaptive control chart with time varying control limits based on online classification methods for data streams," in *12th workshop on quality improvement methods in Dortmund, Germany*, 2013. 48, 127

[119] G. J. Ross and N. M. Adams, "Two nonparametric control charts for detecting arbitrary distribution changes," *Journal of Quality Technology*, vol. 44, no. 2, pp. 102–116, 2012. 48

[120] G. Ditzler and R. Polikar, "Hellinger distance based drift detection for nonstationary environments," in *2011 IEEE symposium on computational intelligence in dynamic and uncertain environments (CIDUE)*, pp. 41–48, IEEE, 2011. 48, 50, 51, 127, 131, 132

[121] P. Sobhani and H. Beigy, "New drift detection method for data streams," in *International conference on adaptive and intelligent systems*, pp. 88–97, Springer, 2011. 48

[122] R. N. Lichtenwalter and N. V. Chawla, "Adaptive methods for classification in arbitrarily imbalanced and drifting data streams," in *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp. 53–75, Springer, 2009. 48, 50, 51, 127, 131, 132

[123] Y. Luo, Z. Li, and Z. Wang, "Adaptive cusum control chart with variable sampling intervals," *Computational Statistics & Data Analysis*, vol. 53, no. 7, pp. 2693–2701, 2009. 48

[124] A. Dries and U. Rückert, "Adaptive concept drift detection," *Statistical Analysis and Data Mining: The ASA Data Science Journal*, vol. 2, no. 5-6, pp. 311–327, 2009. 48

[125] D. A. Cieslak and N. V. Chawla, "A framework for monitoring classifiers' performance: when and why failure occurs?," *Knowledge and Information Systems*, vol. 18, no. 1, pp. 83–108, 2009. 48, 50, 51, 127, 131, 132

[126] S. Muthukrishnan, E. Van Den Berg, and Y. Wu, "Sequential change detection on data streams," in *Seventh IEEE International Conference on Data Mining Workshops (ICDMW 2007)*, pp. 551–550, IEEE, 2007. 48, 51, 132

[127] K. Nishida and K. Yamauchi, "Detecting concept drift using statistical testing," in *International conference on discovery science*, pp. 264–269, Springer, 2007. 48

[128] I. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira, "Self-adaptive windowing approach for handling complex concept drift," *Cognitive Computation*, vol. 7, no. 6, pp. 772–790, 2015. 48, 50, 127, 129, 131

[129] I. Khamassi and M. Sayed-Mouchaweh, "Drift detection and monitoring in non-stationary environments," in *2014 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, pp. 1–6, IEEE, 2014. 48, 50, 51, 127, 128, 131, 132

[130] A. Bifet and R. Gavalda, "Adaptive learning from evolving data streams," in *International Symposium on Intelligent Data Analysis*, pp. 249–260, Springer, 2009. 48, 49, 127

[131] S. H. Bach and M. A. Maloof, "Paired learners for concept drift," in *2008 Eighth IEEE International Conference on Data Mining*, pp. 23–32, IEEE, 2008. 48, 128

[132] B. Pfahringer, G. Holmes, and R. Kirkby, "New options for hoeffding trees," in *Australasian Joint Conference on Artificial Intelligence*, pp. 90–99, Springer, 2007. 48, 49, 127

[133] A. Bifet and R. Gavalda, "Learning from time-changing data with adaptive windowing," in *Proceedings of the 2007 SIAM international conference on data mining*, pp. 443–448, SIAM, 2007. 48, 50, 127, 128, 131

[134] S. Hoeglinger and R. Pears, "Use of hoeffding trees in concept based data stream mining," in *2007 Third International Conference on Information and Automation for Sustainability*, pp. 57–62, IEEE, 2007. 48, 49, 127

[135] J. Gama and G. Castillo, "Learning with local drift detection," in *International conference on advanced data mining and applications*, pp. 42–55, Springer, 2006. 48, 50, 51, 127, 131, 132

[136] M. Baena-Garcıa, J. del Campo-Ávila, R. Fidalgo, A. Bifet, R. Gavalda, and R. Morales-Bueno, "Early drift detection method," in *Fourth international workshop on knowledge discovery from data streams*, vol. 6, pp. 77–86, 2006. 48, 50, 51, 127, 131, 132

[137] M. M. Lazarescu, S. Venkatesh, and H. H. Bui, "Using multiple windows to track concept drift," *Intelligent data analysis*, vol. 8, no. 1, pp. 29–59, 2004. 48, 129

[138] D. Kifer, S. Ben-David, and J. Gehrke, "Detecting change in data streams," in *VLDB*, vol. 4, pp. 180–191, Toronto, Canada, 2004. 48, 127, 128

[139] M. Black and R. Hickey, "Learning classification rules for telecom customer call data under concept drift," *Soft Computing*, vol. 8, no. 2, pp. 102–108, 2003. 48, 49, 127

[140] B. Babcock, S. Babu, M. Datar, R. Motwani, and J. Widom, "Models and issues in data stream systems," in *Proceedings of the twenty-first ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems*, pp. 1–16, 2002. 48, 127

[141] G. Hulten, L. Spencer, and P. Domingos, "Mining time-changing data streams," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 97–106, 2001. 48, 49, 127

[142] P. Domingos and G. Hulten, "Mining high-speed data streams," in *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 71–80, 2000. 48, 49, 127

[143] C. Alippi and M. Roveri, "Just-in-time adaptive classifiers in non-stationary conditions," in *2007 International Joint Conference on Neural Networks*, pp. 1014–1019, IEEE, 2007. 49, 131

[144] C. Alippi and M. Roveri, "Just-in-time adaptive classifiers—part ii: Designing the classifier," *IEEE Transactions on Neural Networks*, vol. 19, no. 12, pp. 2053–2064, 2008. 49, 50, 131

[145] A. Bifet, G. Holmes, B. Pfahringer, R. Kirkby, and R. Gavalda, "New ensemble methods for evolving data streams," in *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 139–148, 2009. 49, 128, 129

[146] G. Song, Y. Ye, H. Zhang, X. Xu, R. Y. Lau, and F. Liu, "Dynamic clustering forest: an ensemble framework to efficiently classify textual data stream with concept drift," *Information Sciences*, vol. 357, pp. 125–143, 2016. 49, 130

[147] D. Brzezinski and J. Stefanowski, "Combining block-based and online methods in learning ensembles from concept drifting data streams," *Information Sciences*, vol. 265, pp. 50–67, 2014. 49, 130

[148] K. Jackowski, "Fixed-size ensemble classifier system evolutionarily adapted to a recurring context with an unlimited pool of classifiers," *Pattern Analysis and Applications*, vol. 17, no. 4, pp. 709–724, 2014. 49, 130

[149] P. Sobolewski and M. Wozniak, "Concept drift detection and model selection with simulated recurrence and ensembles of statistical detectors.," *J. Univers. Comput. Sci.*, vol. 19, no. 4, pp. 462–483, 2013. 49, 130

[150] I. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira, "Ensemble classifiers for drift detection and monitoring in dynamical environments," in *Annual Conference of the PHM Society*, vol. 5, 2013. 49, 129, 130

[151] D. Brzezinski and J. Stefanowski, "Reacting to different types of concept drift: The accuracy updated ensemble algorithm," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 25, no. 1, pp. 81–94, 2013. 49, 129, 130

[152] D. Mejri, R. Khanchel, and M. Limam, "An ensemble method for concept drift in nonstationary environment," *Journal of Statistical computation and Simulation*, vol. 83, no. 6, pp. 1115–1128, 2013. 49, 130

[153] M. Woźniak and B. Krawczyk, "Combined classifier based on feature space partitioning," *International Journal of Applied Mathematics and Computer Science*, vol. 22, no. 4, pp. 855–866, 2012. 49, 129

[154] R. Elwell and R. Polikar, "Incremental learning of concept drift in nonstationary environments," *IEEE Transactions on Neural Networks*, vol. 22, no. 10, pp. 1517–1531, 2011. 49, 129

[155] J. Z. Kolter and M. A. Maloof, "Dynamic weighted majority: An ensemble method for drifting concepts," *The Journal of Machine Learning Research*, vol. 8, pp. 2755–2790, 2007. 49, 129, 130

[156] F. Chu and C. Zaniolo, "Fast and light boosting for adaptive mining of data streams," in *Pacific-Asia conference on knowledge discovery and data mining*, pp. 282–292, Springer, 2004. 49, 129

[157] R. Polikar, L. Upda, S. S. Upda, and V. Honavar, "Learn++: An incremental learning algorithm for supervised neural networks," *IEEE transactions on systems, man, and cybernetics, part C (applications and reviews)*, vol. 31, no. 4, pp. 497–508, 2001. 49, 129

[158] A. Shaker and E. Lughofer, "Self-adaptive and local strategies for a smooth treatment of drifts in data streams," *Evolving Systems*, vol. 5, no. 4, pp. 239–257, 2014. 50, 131

[159] L. I. Kuncheva and I. Žliobaitė, "On the window size for classification in changing environments," *Intelligent Data Analysis*, vol. 13, no. 6, pp. 861–872, 2009. 50, 127, 131

[160] G. Cauwenberghs and T. Poggio, "Incremental and decremental support vector machine learning," *Advances in neural information processing systems*, vol. 13, 2000. 50, 51, 131, 132

[161] B. Krawczyk and M. Woźniak, "One-class classifiers with incremental learning and forgetting for data streams with concept drift," *Soft Computing*, vol. 19, no. 12, pp. 3387–3400, 2015. 51, 132

[162] C. Pinto and J. Gama, "Incremental discretization, application to data with concept drift," in *Proceedings of the 2007 ACM symposium on Applied computing*, pp. 467–468, 2007. 51, 132

[163] L. I. Kuncheva, "Classifier ensembles for changing environments," in *International Workshop on Multiple Classifier Systems*, pp. 1–15, Springer, 2004. 51, 130, 132

[164] 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. 59, 96

[165] 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. 120

[166] 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. 121

[167] 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. 121, 122, 123

[168] P. Grassberger and I. Procaccia, "Characterization of strange attractors," *Physical review letters*, vol. 50, no. 5, p. 346, 1983. 122

[169] L. D. Iasemidis, J. Chris 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. 123

[170] 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. 123

[171] M. Bandarabadi, C. A. Teixeira, B. Direito, and A. Dourado, "Epileptic seizure prediction based on a bivariate spectral power methodology," *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 5943–5946, 2012. 124

[172] L. D. Iasemidis, D.-S. Shiau, J. C. Sackellares, P. M. Pardalos, and A. Prasad, "Dynamical resetting of the human brain at epileptic seizures: application of nonlinear dynamics and global optimization techniques," *IEEE transactions on biomedical engineering*, vol. 51, no. 3, pp. 493–506, 2004. 124

[173] G. J. Ross, N. M. Adams, D. K. Tasoulis, and D. J. Hand, "Exponentially weighted moving average charts for detecting concept drift," *Pattern recognition letters*, vol. 33, no. 2, pp. 191–198, 2012. 126

[174] J. Gama, P. Medas, G. Castillo, and P. Rodrigues, "Learning with drift detection," in *Brazilian symposium on artificial intelligence*, pp. 286–295, Springer, 2004. 127

[175] G. Widmer and M. Kubat, "Learning flexible concepts from streams of examples: Flora2," 1992. 127

[176] G. Widmer and M. Kubat, "Learning in the presence of concept drift and hidden contexts," *Machine learning*, vol. 23, no. 1, pp. 69–101, 1996. 127

[177] H. Wang, J. Yin, J. Pei, P. S. Yu, and J. X. Yu, "Suppressing model overfitting in mining concept-drifting data streams," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 736–741, 2006. 129

[178] W. Fan, "Systematic data selection to mine concept-drifting data streams," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 128–137, 2004. 129

[179] W. N. Street and Y. Kim, "A streaming ensemble algorithm (sea) for large-scale classification," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 377–382, 2001. 130

[180] A. Bifet, G. Holmes, and B. Pfahringer, "Leveraging bagging for evolving data streams," in *Joint European conference on machine learning and knowledge discovery in databases*, pp. 135–150, Springer, 2010. 130

[181] K. O. Stanley, "Learning concept drift with a committee of decision trees," *Informe técnico: UT-AI-TR-03-302, Department of Computer Sciences, University of Texas at Austin, USA*, 2003. 130

[182] H. Wang, W. Fan, P. S. Yu, and J. Han, "Mining concept-drifting data streams using ensemble classifiers," in *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 226–235, 2003. 130

Appendix A

Feature description

This chapter provides a detailed description of the most extracted features from the state-of-the-art and a more detailed description of the used features in this thesis. It is worth noting that, in this thesis, only linear univariate features were used.

Univariate Linear features

Linear features are mathematical measures that extract linear dynamics from signals by using phase/frequency and amplitude information. When these features are extracted, the Electroencephalogram (EEG) signal is considered quasi-stationary within each time window.

Statistical Moments

The distribution of the EEG time series amplitude can be determined using statistical moments. The first four statistical moments (mean, variance, skewness, kurtosis) are summarised in Table A.1, where N is the number of samples in the sliding window and x is a vector of the input values [17, 83, 85, 86, 103]. As previously stated in Section 3.1.4 univariate linear measures are light to compute. They can also be used on characteristics other than the electrical amplitude, such as spectral information [12, 83]. For symmetric amplitude distributions, the skewness is zero; for asymmetric distributions, it is non-zero. The kurtosis assesses how peaked or flat an amplitude distribution is [20]. According to Rasekhi et al. [87], these statistical measures have demonstrated considerable changes between the interictal and preictal states. Compared to interictal data, the preictal period showed a decrease in variance, and an increase in kurtosis [86].

Accumulated energy

The main idea behind adopting accumulated energy as a feature is that seizure-generation mechanisms intensify brain activity, resulting in an accumulation of en-

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)	$\kappa = \frac{\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu)^4}{(N-1)\sigma^4} - 3$ (A.4)	Measures the relative flatness or peakedness of the amplitude distribution

ergy [85, 87, 165].

To compute accumulated energy, instantaneous energy is first obtained by squaring the EEG values. The resulting instantaneous energy sequence is then subjected to an average window of length N (A.5).

$$E_a = \frac{1}{N}\sum_{i=1}^{N} x_i^2 \tag{A.5}$$

Finally, the accumulated energy (A.6) is given by the average of the E_a sequence using other window with length W:

$$E_{ac_k} = \frac{1}{W}\sum_{i=1}^{W} E_a + E_{ac_{k-1}} \tag{A.6}$$

where E_{ac_k} is k^{th} value of the accumulated energy and $E_{ac_{k-1}}$ is the previous value [85].

Hjörth parameters

Hjörth parameters concern activity, mobility, and complexity, which are measures of mean power, root-mean-squared frequency, and root-mean-square frequency spread, respectively. These monitor, an increase in brain activity that increases energy [6,17,83,85,86]. The mobility and complexity of the EEG, were shown to significantly increase during the preictal stage [87]. Hjörth parameters can be determined by taking into account the variance $\sigma^2(x)$ of the input signal x, and the first and second derivatives, x' and x'', respectively, by the Equations A.7, A.8 and A.9:

$$H_a = \sigma^2(x) \tag{A.7}$$

$$H_m = \sqrt{\frac{\sigma^2(x')}{\sigma^2(x)}} \tag{A.8}$$

$$H_c = \sqrt{\frac{\sigma^2(x'')\sigma^2(x)}{(\sigma^2(x'))^2}} \tag{A.9}$$

Auto-regressive models

Auto-regressive models assume the model's output in instant t is a weighted sum of previous p values (p determines the order of the model) plus a constant term c and white noise ε. It is necessary to consider that the signal is stationary. An auto-regressive model can be characterised by considering Y_t, the predicted value for instant t, and ar_i, the model's parameters, as describe by Equation A.10:

$$Y_t = const + \sum_{i=1}^{P} ar_i Y_{t-1} + \varepsilon_t \tag{A.10}$$

By modeling the EEG, these models have been used to evaluate neuronal synchronisation [17, 83, 85–87]. The modelling error as a result of a seizure-generation process [87] or the modelling coefficient values as features have both been employed by authors [166].

Decorrelation time

The decorrelation time is defined as the first zero crossing of the autocorrelation function and gives information about the data variability's usual time scale. It estimates the periodicity of data and the strength of linear relationships; the lower its values, the less correlated the signal. This function may also be used to assess signal stochasticity since a temporal value of zero decorrelation indicates that a particular signal is entirely stochastic (white noise). A reduction in decorrelation time has been recorded before seizures [83, 86, 87, 167].

Relative spectral power

The signal power associated with particular frequency ranges (delta, theta, alpha, beta, and gamma) is quantified by spectral power. Authors then utilise these ranges to extract features that capture low to high-frequency transitions. These are the most commonly used characteristics [6, 17, 83, 85–87, 91] and may be determined using the Power Spectral Density (PSD). The PSD can be calculated by performing the Fast Fourier Transform (FFT) on the EEG time series and then averaging the squared coefficients of the frequency range of interest. It is critical to note that the PSD calculation presupposes that the signal in each window is both short enough to be deemed quasi-stationarity and long enough to capture the brain's low-frequency activity. The power of a certain frequency band divided by the overall power of the EEG signal characterises the relative spectral power. Because there is more power in low frequencies than in high frequencies, normalised spectral power gives a more robust measurement. Some authors documented a power shift from lower to higher frequencies before seizure onset [83, 84, 86, 87]. For example, Mormann et al. [167] demonstrated a drop in Delta band power, which was accompanied by a fall in relative power in the other sub-bands. Bandarabadi et al. [84] showed that

relative combinations of sub-band spectral energies across channel pairs might be used to follow progressive changes before seizures.

Spectral edge frequency

The minimal frequency below which a specified proportion of the total power of the signal is contained is usually referred to as the Spectral edge frequency (SEF), and it is frequently used in seizure prediction [86,87,167]. Most of the spectral strength in the EEG signal is contained in the 0-40Hz band [20]. SEF 50 is the frequency at which 50% of the signal's overall power is situated. As a result, SEF may be able to detect a power shift from low to high frequencies during the preictal phase [83,86].

Wavelet transform

The wavelet transform is a time-frequency domain transform that can be used in place of the FFT. It can show the spectral and temporal features of a signal. The wavelet transform divides the signal into several resolution levels based on the frequency ranges [17,83,86,87]. The first decomposition layers correspond to higher frequencies, whereas the latter levels correspond to lower frequencies. After the signal decomposition, it is possible to compute discriminant features from distinct frequency bands by applying the wavelet coefficients. A characteristic that may be acquired with the wavelet transform is the measurement of energy in different frequency ranges [83,86,87].

Univariate Non-linear features

Because the EEG is a noisy and nonstationarity time series, chaotic measures can assist in describing brain dynamics. In theory, a decrease in chaos may suggest upcoming seizures because the predictability of brain dynamics tends to rise before a seizure. The measures described below are designed to capture the increased brain synchronisation that occurs before seizures. Furthermore, nonlinear features may be too computationally costly to be used in an online system [12].

Correlation sum and dimension

The correlation sum, originally proposed by Grassberger and Procaccia [168], expresses the likelihood that two vectors in the state space trajectory are close to one another. One of the ways for fractal dimension measurement is the correlation dimension of a signal, which evaluates the space dimension filled by signal samples. It offers a mathematical estimate of the complexity of attractors.

Lyapunov exponent

Lyapunov exponents quantify the exponential divergence of neighbouring state space trajectories, representing all potential system states. One basic sign of deterministic chaos is the exponential divergence, which is affected by the initial state conditions [6, 12]. Fitting an exponential regression can be used to compute it. Some studies [6,78] have used Lyapunov exponents. The early findings showed a drop in the largest Lyapunov exponent several minutes before seizure start [169]; however, a subsequent analysis found the contrary, suggesting an increase in the largest Lyapunov exponent 30 minutes before seizure onset [167].

Entropy

Entropy [79] quantifies the regularity and predictability of a signal's variations. There are various entropy metrics, including approximation entropy, sample entropy, permutation entropy, and spectral entropy. Because a synchronised brain state characterises a seizure, entropy has been found to detect shifts from the interictal to the preictal state.

Dynamic Similarity Index

Le Van Quyen et al. [170] introduced the Dynamic Similarity Index to quantify the dynamical similarity between two segments of an EEG signal. One of the segments is a reference segment, which is generally a few minutes long and chosen at a time apart from the seizure (an interictal segment). The other is a segment provided by a moving test window [12, 20, 170]. When the difference in the dynamics of these windows exceeds a certain threshold, it is presumed to identify the preictal period.

Multivariate Linear features

As the preictal stage is a spatio-temporal complicated condition and seizures are known to be electrical discharges caused by brain synchronisation. Multivariate features are of particular interest, as these describe interactions between various brain areas and, hence, different electrodes [12, 88].

Maximum Linear cross-correlation

Maximum linear cross-correlation measures the degree of lag synchronisation of two electrode channels, which is the situation in which two identical signals are shifted by a temporal lag τ. As a result, this characteristic measures the similarity of two time series [12, 20, 103, 105]. The maximum normalised cross-correlation maintains the measure's independence from signal variance. Values close to 1 suggest a high

degree of similarity between the two signals under consideration, but with a potential temporal lag τ, whereas asynchronous signals will produce values near 0 [12,20,103].

Ratio and Differences

To measure the interrelationship between distinct channels, relative and differential techniques were used [85, 171]. Relative and differential features are calculated by dividing and subtracting features from one channel to/from features from another. Given N features in the feature vector F for each of the M recording channels, relative and differential features may be represented as A.11 and A.12, respectively, where $f_k(i)$ represents the k-th feature from the i-th channel.

$$f_{rel_k}(i,j) = \frac{f_k(i)}{f_k(j)}, k = 1,2,...,N; i,j = 1,2,...,M; i \neq j \tag{A.11}$$

$$f_{dif_k}(i,j) = f_k(i) - f_k(j), k = 1,2,...,N; i,j = 1,2,...,M; i \neq j \tag{A.12}$$

Multivariate Non-linear features

These aim to capture synchrony changes using similarity and mutual information measures by inspecting information in several electrodes simultaneously.

Non-linear interdependence

The non-linear interdependence between two EEG signals from distinct channels is regarded as a measure of generalised synchronisation. Generalised synchronisation occurs when the dynamical state of one of the coupled systems is determined by the other [78, 88].

Mean phase coherence

Mean Phase Coherence (MPC) estimates the degree of phase synchronisation between two time series. The MPC is restricted between 0 and 1, with values close to 1 indicating a high degree of synchronisation [20, 88, 103]. Some studies observed a drop in mean phase coherence values before seizure onset [88, 103].

Dynamical entrainment

Dynamical entrainment used in [105], was initially proposed by Iasemidis et al. [172]. It is based on chaos theory as it is a multivariate version of the Lyapunov exponent. This measure aims to quantify the non-linear behaviour of two series.

Appendix B

Concept Drifts Adaptation detailed description

B.1 Data process

Sequential methods - Change detection

The following section explains how these techniques were modified for the learning task and how their statistical processing works [13].

Given a series of independent random observations $X_1, ..., X_n$, where each observation X_i is produced from a particular distribution D_i, the online statistical analysis can be applied to one of two scenarios [13]:

- The distribution is said to be stationary if all of the observations $X_1, ..., X_n$ are produced from the same distribution D_0.
- If, for a sub-sequence $X_1, ..., X_k$ where $1 < k < n$, it exists a change point $k < \lambda < n$ such that X_λ is generated according to another distribution D_1, where $D_0 \neq D_1$; Consequently, D_0 and D_1 are referred to as the pre-and post-change distributions, respectively. At this point, we declare the distribution to be non-stationary.

The basic goal of the sequential techniques is to determine if the distribution at point λ differs from distributions D_0 and D_1. Two hypotheses are put forth for this reason [13]:

- The null hypothesis $H_0 : D_0 = D_1$ there is no change in the original distribution D_0.
- The alternative hypothesis $H_1 : D_0 \neq D_1$ there is a change in D_0.

As a result, two fundamental factors are established to specify the degree of reliability of the hypothesis testing:

- $\alpha = P(H_1|H_0)$ false alarm rate: the probability of accepting H_1 when H_0 is true, i.e., the probability of detecting a change when the distribution is stationary.
- $\beta = P(H_0|H_1)$ missed detection rate: the probability of accepting H_0 when H_1 is true, i.e., the probability of wrongly considering that the distribution is stationary.

According to conventional wisdom, a change happens when D_0 and D_1 diverge significantly. Using some dissimilarity measure, $Diss_\lambda(D_0, D_1)$, this change is measured. The hypothesis testing evaluates the greatest $Diss_\lambda(D_0, D_1)$ against a τ change threshold to determine whether a difference is significant. Thus, the two primary elements of hypothesis testing, the change threshold and the dissimilarity measure, used by the former [13]. Both defined in the following sections.

Dissimilarity Measure

The difference between the two data distributions D_0 and D_1 at a certain point λ is measured using the formula $Diss_\lambda(D_0, D_1)$. The dissimilarity measure for change detection is often applied in one of two ways [13]:

- It could operate directly on data, for instance, by calculating how different an incoming instance is from a set of data.
- It might be applicable to work on statistics like mean, variance, and covariance summarised from the two distributions. For instance, the Exponentially Weighted Moving Average chart was used in Ross et al. [173] (2012) to identify a significant rise in the original distribution's mean.

Change threshold

If the dissimilarity measure $Diss_\lambda(D_0, D_1)$ at a certain point λ is greater than the change threshold τ, a drift is identified in sequential techniques. This threshold can be either fixed or variable [13].

The first is typically predefined by the user and is connected to the drift's specificity. Larger thresholds are better suited for slow drifts, whereas lower thresholds are better suited for sudden changes. The variable threshold is more suited to handle various kinds of drift. Sequential methods that employ varying thresholds tend to be more autonomous and capable of accurately identifying the change point. They must, however, keep an eye on the false alarm and missed detection rates [13].

Windowing techniques

According to Khamassi et al. [13], window characteristics can be defined according to four dimensions:

Specificity

This dimension entails the position of the window regarding the learning system, which is inside [130,132,134,139,141,142], or independent [23,120,122,125,128,133, 135,136,174–176] of the base learner. For the former, the essential statistics of the learner are estimated using this window, and they are kept up to date with current concepts. Whilst for the latter, the window is not learner specific, therefore, can be applied to any learner. These windows monitor indicators external and independent of the learner (e.g. prediction performance or characteristics of incoming data distribution) [13].

Nature

Windows, where the number of instances characterises their size, are known as Data-based or sequence-based windows. In contrast, when the length of time determines the window size, they are known as time-based, or timestamp-based windows [13, 140].

Size

Regarding window size, windows can be either *fixed* [138, 175] or *variable* [23, 118, 128, 129, 133, 159, 176]. The former have their size fixed and previously defined. A small size is good for detecting sudden drifts, whilst larger sizes are suited for detecting gradual drifts. While the latter are windows of dynamic size to handle different types of drift. Several methods exist to determine the window size [13].

Positioning strategy

Positioning strategy refers to how the window or windows evolve throughout the concept drift tracking (see Figure B.1) [13]. To properly characterise the algorithms, they are first divided according to the number of windows used:

- *Single windows* can be further separated into *Sliding* [141, 175] or *Landmark* [23, 135, 136, 176], in a *sliding window* a fixed number of instances is recorded in a *First-In-First-Out (FIFO)* data structure. The learner is periodically updated because the oldest instance is deleted whenever a new one appears and is saved in memory. In contrast, beginning at a specific moment, the *landmark window* begins to store instances until a specific condition is met. When there is no change, they can continue to store instances and increase

Figure B.1: Different positioning strategies illustrated. Adapted from Khamassi et al. 2018 [13].

their size; nevertheless, once a drift is detected, the window size is reduced to contain the most representative data [13].

- With *Two windows*, keeping one window as a reference and the other as the current data batch is the principle behind this strategy. These windows can be: *separated*, *adjacent*, *overlapping*, or *discontinued*.

  - *Separated* where as data is processed the *reference window* diverges from the *current window* [131]. The offline learner is initially trained on the *reference window* before being continuously tested on the current batch of data. A new learner is constructed from the current data batch if the performance has fallen below a certain threshold.
  - *Adjacent* where as data is processed the *reference* and *current windows* remain joined. This strategy needs an online learner because the comparison between both windows is done sequentially, and the resizing process can be done in three different manners: both windows remain with a fixed size [138]; the *current* window has a fixed size and the *reference* a variable size, storing instances while no change is detected [129]; and both the *reference* and *current windows* have a variable size [133, 145].
  - *Overlapping* where as data is processed, the *reference* and *current windows* have data samples in common [114, 138].
  - *Discontinued* where as data is processed, only subsets of the reference window are used for comparison with the current window. This strategy selects subsets of data from the reference window and compares them

with the current window. Subsets can be chosen according to their spatial similarity, temporal similarity, or representativeness [128].

- *Multiple windows* were also found in the literature. Three windows of varying sizes (small, medium, and large) were used in Lazarescu et al. [137] (2004) to handle various types of drift. Each window has a designated drift speed: the small window is designed to handle concepts that change very quickly, the intermediate window is for concepts that change slower, and the large window is used to handle concepts that move very slowly.

B.2 Learning process

Ensemble learners

Training set management

Effectively managing the training set is one approach to guarantee that the ensemble will react well to changes [13]. To train base learners differently and assure diversity in the ensemble, the training set partition can be satisfied by using a variety of strategies:

- *Block-based techniques* present the training set as blocks or pieces of data at a time. Very frequently, after each data block, ensemble learners evaluate their components and replace the one who performs the weakest with a new (candidate) learner [151,155,157]. This method maintains the ensemble's adaptability, so learners who have received training in recent blocks are best suited to represent the current concept.

- *Weighting data techniques* train the base learners according to weighted instances from the training set [9, 13]. Two popular instance weighted ensemble algorithms used are Bagging and Boosting with several extensions [145,150,154,156,157,177,178].

- *Filtering data techniques* choose data from the training set based on a predetermined criterion, such as feature space similarity [114,153]. This method allows base learners to be trained under several sub-spaces to take advantage of various features of the overall feature space.

Structure management

Effectively managing the ensemble's structure is another approach to guarantee strong adaptability to changes [13]. As a result, the adaptability can be fulfilled by:

- *Fixed size* are ensembles with *a priori* fixed number of base learners. The weakest learner is replaced with a new one trained on recent data. It is one of the well-known methods for managing fixed-size ensembles [151, 155, 163, 179]. Similar to this, some systems employ an ensemble drift detection mechanism to only replace the weakest learner when a drift is detected [147, 150, 180].

- *Variable size* are ensemble solutions, where classifiers are left in the pool, and historical context information is retained for use in the future. For recurrent drifts, this is advantageous. For instance, an evolutionary-based optimisation method that seeks to reduce the system misclassification rate can be used to implement the process of choosing classifiers and updating their weights [148, 181].

Final decision management

An appealing technique in ensemble methods is ensuring flexibility when aggregating the decisions of the base learners. In the current state of the art in non-stationary learning, creating a final decision is also subject to significant efforts. This can be done either by *dynamic weighting*, *dynamic selection*, or by *combining weighting and selecting strategies*:

- *Dynamic weighting technique* is a weighted majority vote that combines the decisions of all the base learners. This method's main goal is to ensure that the learners' decisions are aggregated in non-stationary contexts [146, 147, 150–152, 155, 180, 182]. The accuracy performance on the most recent data blocks is used to continuously update each learner's weight, giving the learner who obtains the highest accuracy the best weight. This technique is suitable for dealing with gradual, continuous drifts where changes are minor as they are only apparent over a long period [13].

- *Dynamic selection technique* is based on choosing one learner from the ensemble pool who is most suited to make predictions [13]. This can be the learner trained on the most recent data block, the learner with the highest weight based on accuracy performance, or the learner that is the most similar (similarity in feature space) to the instance to be classified [149].

- *Combining dynamic weighting and selection techniques* is founded on picking a subgroup of learners and then adding up their predictions. Learners can typically be chosen based on their accuracy performance, age, or proximity to the incoming instances, and the final selection is then compiled using a weighted majority vote [21, 146, 148].

B.3 Monitoring process

Methods based on supervised indicators

These methods often focus on preserving the learner's performance by handling *Real Concept Drifts*. The primary key for handling this type of drift relies on monitoring the learner feedback indicators, such as *accuracy* [128, 129, 133, 135, 136], *recall*, *precision*, *sensitivity* and *specificity*.

These indicators have the benefit of being trustworthy and method-independent. However, because they function in supervised mode, they require a genuine class label. If the correct label is not immediately available, as in most real applications, a delay in the detection of changes can occur [13].

Methods based on unsupervised indicators

In many real-world situations where the data is unlabeled, methods based on unsupervised indicators help detect changes when the prediction feedback is delayed. They can also help handle *Virtual Concept Drifts* because they have no impacts on the decision boundaries [13] as seen in Section 2.4. These methods can be based on the following:

- *Similarity in Time*: Evaluates how a data distribution evolves from one time stamp to another. The similarity in time can be accessed using hypothesis tests such as Page-Hinkley test [158], Sequential probability ratio test [159], or CUSUM test [143, 144].

- *Similarity in Space*: Evaluates how a data distribution evolves regarding the feature space. The similarity in space can be examined using distance functions such as Euclidean distance [114], Heterogeneous Euclidean-overlap distance [21], Mahalanobis distance [116, 117], Hellinger distance [120, 122, 125].

- *Model Complexity Measure*: is built on observing the model's structure and/or parameters. An odd model behaviour may be revealed by the increase of the number of rules for rule-based classifiers or the number of support vectors for the Support Vector Machine (SVM) algorithm [160].

B.4 Adapting process

Informed methods

In general, the strategy chosen depends on the goal of dealing with concept drift. For instance, it is crucial to identify unusual activity and out-of-control behaviour in monitoring and control applications. An indication of the drift is required in such

a situation, which is frequently stated as a detection problem. The informed methods—also known as drift detection mechanisms—are best suited for this purpose since they expressly detect drifts through triggering mechanisms. The latter is helpful when we hope to describe the incidence and the drifts' time of discovery. From the standpoint of machine learning, these methods may keep an eye on a learner's performance [117, 129, 135, 136], data distributions [120, 122, 125], or the structure and parameters of the learner [160] to spot a drift. They are reactive because they may either start over and learn the model from scratch or update it with a current set of data (data window) when a drift is identified. In conclusion, informed methods process as follows: Detecting the drift; Deciding which data should be kept and which one should be forgotten; Retraining the current learner when a significant change has been detected [13].

Blind methods

Without any drift detection, the blind methods automatically adapt the learner to the current concept at regular intervals. Regardless of whether changes have occurred, they reject outdated concepts at a steady rate. These methods can help manage gradual, continuous drifts when the differences between consecutive data sources do not sufficiently warrant a change [13]. The following are some techniques employed by blind approaches to managing drift:

- *Fixed size sliding window* the learner is periodically updated according to a fixed number of instances stored in the FIFO data structure [126].
- *Instance weighting* the learner is periodically updated according to weighted instances from the training set. For example, the more recent data instances should have the highest weights [161, 162].
- *Ensemble learners* individual learners are re-evaluated, and the worst one is replaced by a new one trained on recent data [163].

Appendix C

Supplementary Results

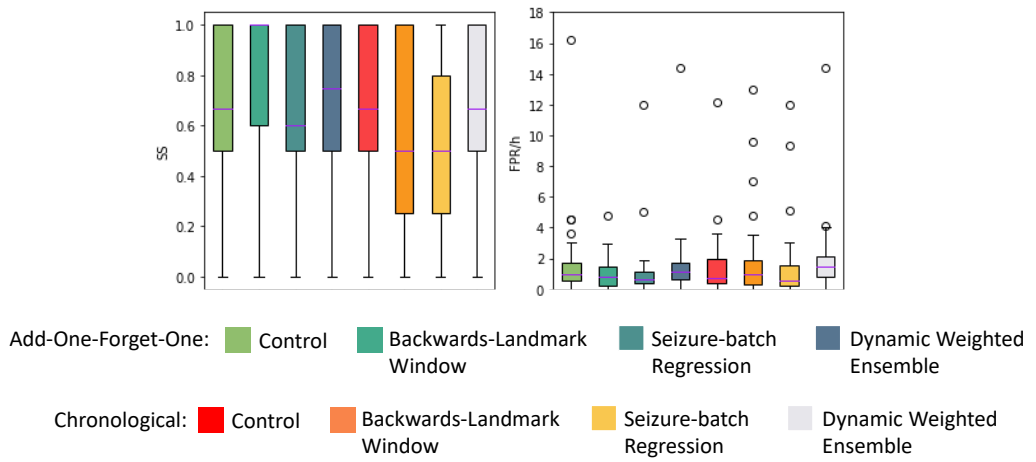

Figure C.1: Seizure prediction performance across all patients for each approach, both Add-One-Forget-One and Chronological methods. Seizure-batch Regression has an outlier of 97.86 False Positive Rate per Hour (FPR/h) for both iterative retraining methods.

Table C.1: Average seizure prediction performance across all patients for each approach, both Add-One-Forget-One and Chronological methods.

Approach	Retraining method	All patients			Validated patients		
		SS	FPR/h	SS Surrogate	%	SS	FPR/h
Control	Add-One-Forget-One	0.63 ± 0.34	1.72 ± 2.65	0.26 ± 0.51	83.78	0.70 ± 0.27	1.69 ± 2.75
	Chronological	0.68 ± 0.31	1.44 ± 2.06	0.23 ± 0.48	91.89	0.70 ± 0.30	1.32 ± 2.07
Backwards-Landmark Window	Add-One-Forget-One	0.75 ± 0.33	1.03 ± 1.00	0.22 ± 0.47	89.19	0.81 ± 0.25	0.91 ± 0.81
	Chronological	0.53 ± 0.33	1.83 ± 2.71	0.23 ± 0.48	75.68	0.66 ± 0.29	1.92 ± 3.05
Seizure-batch Regression	Add-One-Forget-One	0.64 ± 0.31	3.73 ± 15.82	0.23 ± 0.48	86.49	0.68 ± 0.27	0.67 ± 0.51
	Chronological	0.50 ± 0.36	4.07 ± 15.82	0.23 ± 0.49	64.86	0.67 ± 0.26	0.72 ± 0.59
Dynamic Weighted Ensemble	Add-One-Forget-One	0.69 ± 0.36	1.6 ± 2.26	0.25 ± 0.50	83.78	0.79 ± 0.28	1.18 ± 0.76
	Chronological	0.61 ± 0.36	1.84 ± 2.31	0.25 ± 0.51	75.68	0.64 ± 0.34	1.49 ± 1.05

Table C.2: Training parameters and performance obtained for each patient with the Backwards-Landmark Window and the Chronological data partitioning and iterative re-training method.

Chronological - Backwards-Landmark Window

Patient	SOP	k	C	SS_{sample}	SP_{sample}
402	40, 20	10, 30	$2^{-10}, 2^{0}$	0.53, 0.42	0.62, 0.65
8902	20, 30	10, 10	$2^{-10}, 2^{0}$	0.82, 0.97	0.78, 0.78
11002	10	20	2^{-10}	0.42	0.71
16202	15, 10, 50, 20	30, 30, 10, 30	$2^{0}, 2^{0}, 2^{-10}, 2^{0}$	0.63, 0.42, 0.52, 0.24	0.82, 0.73, 0.79, 0.71
21902	40	10	2^{6}	0.71	0.73
23902	50, 10	40, 40	$2^{-10}, 2^{-10}$	0.37, 0.55	0.62, 0.67
26102	50	30	2^{0}	0.31	0.59
30802	10, 10, 50, 25, 50	40, 10, 40, 20, 40	$2^{-10}, 2^{-10}, 2^{2}, 2^{-10}, 2^{-10}$	0.59, 0.26, 0.45, 0.59, 0.65	0.84, 0.86, 0.53, 0.70, 0.63
32702	15, 25	10, 20	$2^{-10}, 2^{-10}$	0.57, 0.61	0.74, 0.71
45402	15	30	2^{-10}	0.72	0.53
46702	30, 10	20, 10	$2^{-8}, 2^{-10}$	0.37, 0.51	0.76, 0.85
50802	15, 25	30, 20	$2^{-8}, 2^{-10}$	0.70, 0.39	0.86, 0.90
53402	40	10	2^{-10}	0.47	0.65
55202	35, 10, 25, 10, 10	10, 20, 30, 10, 30	$2^{-6}, 2^{-10}, 2^{2}, 2^{-10}, 2^{-6}$	0.62, 0.51, 0.60, 0.94, 0.53	0.53, 0.72, 0.56, 0.67, 0.69
56402	25	10	2^{-8}	0.41	0.74
58602	10, 10, 10	10, 30, 30	$2^{0}, 2^{0}, 2^{0}$	0.24, 0.33, 0.53	0.71, 0.79, 0.75
59102	20, 15	20, 40	$2^{-10}, 2^{-4}$	0.48, 0.48	0.59, 0.37
60002	20, 20, 40	10, 10, 40	$2^{-8}, 2^{-8}, 2^{-10}$	0.44, 0.26, 0.44	0.77, 0.77, 0.83
64702	30, 20	30, 10	$2^{-10}, 2^{-10}$	0.42, 0.80	0.68, 0.67
75202	30, 10, 45, 35	10, 10, 10, 30	$2^{-4}, 2^{-10}, 2^{-10}, 2^{-4}$	0.68, 0.53, 0.31, 0.56	0.85, 0.83, 0.85, 0.49
80702	35, 50, 15	40, 40, 30	$2^{-10}, 2^{-8}, 2^{-10}$	0.3, 0.63, 0.53	0.79, 0.55, 0.51
85202	15, 40	30, 10	$2^{0}, 2^{0}$	0.45, 0.63	0.66, 0.67
93402	15, 20	20, 40	$2^{4}, 2^{-6}$	0.45, 0.67	0.71, 0.71
93902	50, 50, 15	10, 20, 20	$2^{-10}, 2^{-10}, 2^{-8}$	0.56, 0.45, 0.43	0.62, 0.84, 0.75
94402	10, 40, 15, 15	30, 30, 30, 10	$2^{0}, 2^{0}, 2^{0}, 2^{-10}$	0.41, 0.23, 0.13, 0.54	0.63, 0.64, 0.77, 0.84
95202	10, 35, 40, 10	10, 30, 40, 10	$2^{-10}, 2^{4}, 2^{0}, 2^{-10}$	0.82, 0.35, 0.39, 0.76	0.57, 0.66, 0.68, 0.9
96002	25, 40, 35, 45	20, 30, 40, 20	$2^{-2}, 2^{-6}, 2^{-6}, 2^{8}$	0.52, 0.29, 0.61, 0.39	0.70, 0.63, 0.73, 0.86
98102	35, 45	10, 10	$2^{0}, 2^{-10}$	0.51, 0.60	0.52, 0.68
101702	30, 50	30, 40	$2^{-4}, 2^{-2}$	0.63, 0.70	0.51, 0.58
102202	50, 45, 50, 50	30, 40, 30, 10	$2^{-6}, 2^{-4}, 2^{-2}, 2^{-6}$	0.06, 0.32, 0.27, 0.09	0.97, 0.48, 0.48, 0.75
104602	25, 50	10, 30	$2^{-10}, 2^{0}$	0.43, 0.60	0.62, 0.67
109502	10	10	2^{-10}	0.37	0.79
112802	10, 15, 15	10, 10, 40	$2^{-10}, 2^{-6}, 2^{-10}$	0.65, 0.38, 0.19	0.49, 0.58, 0.79
113902	45, 15, 45	30, 30, 10	$2^{-10}, 2^{-10}, 2^{-10}$	0.42, 0.64, 0.46	0.56, 0.73, 0.61
114702	35, 15, 15, 50, 15	30, 40, 40, 10, 30	$2^{-10}, 2^{-4}, 2^{-10}, 2^{-10}, 2^{-10}$	0.22, 0.17, 0.67, 0.41, 0.56	0.76, 0.93, 0.70, 0.53, 0.74
114902	25, 50, 50, 20	20, 30, 10, 10	$2^{0}, 2^{8}, 2^{-10}, 2^{-10}$	0.53, 0.04, 0.15, 0.67	0.44, 0.77, 0.91, 0.84
123902	10, 10	30, 30	$2^{-10}, 2^{-10}$	0.79, 0.65	0.84, 0.83
Overall	-	-	-	0.49 ± 0.19	0.70 ± 0.12

Table C.3: Training parameters and performance obtained for each patient with the Seizure-batch Regression and the Chronological data partitioning and iterative retraining method.

Chronological - Seizure-batch Regression

Patient	SOP	k	C	SS_{sample}	SP_{sample}
402	10, 15	10, 10	$2^{-10}, 2^{8}$	0.69, 0.19	0.65, 0.59
8902	20, 20	30, 30	$2^{-10}, 2^{-10}$	0.87, 0.91	0.84, 0.85
11002	10	20	2^{-10}	0.53	0.61
16202	45, 15, 10, 30	10, 40, 20, 40	$2^{-10}, 2^{8}, 2^{8}, 2^{6}$	0.83, 0.55, 0.46, 0.21	0.61, 0.59, 0.69, 0.71
21902	10	10	2^{-10}	0.67	0.61
23902	50, 45	40, 10	$2^{-8}, 2^{-10}$	0.50, 0.66	0.53, 0.44
26102	50	40	2^{6}	0.66	0.35
30802	50, 50, 50, 35, 35	30, 10, 10, 10, 10	$2^{-10}, 2^{0}, 2^{-10}, 2^{0}, 2^{-10}$	0.90, 0.72, 0.76, 0.79, 0.79	0.79, 0.78, 0.65, 0.66, 0.59
32702	15, 15	10, 10	$2^{-2}, 2^{-10}$	0.67, 0.81	0.70, 0.63
45402	50	40	2^{-10}	0.67	0.55
46702	35, 25	30, 30	$2^{0}, 2^{0}$	0.20, 0.49	0.68, 0.63
50802	15, 20	30, 30	$2^{-10}, 2^{-10}$	0.82, 0.73	0.70, 0.8
53402	15	20	2^{-4}	0.22	0.72
55202	10, 10, 10, 10, 10	30, 10, 30, 30, 30	$2^{0}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.50, 0.74, 0.69, 0.67, 0.71	0.73, 0.67, 0.55, 0.64, 0.61
56402	25	20	2^{-10}	0.46	0.72
58602	10, 10, 20	40, 40, 40	$2^{-10}, 2^{-10}, 2^{-8}$	0.15, 0.17, 0.65	0.78, 0.87, 0.73
59102	15, 50	10, 30	$2^{-10}, 2^{0}$	0.66, 0.52	0.44, 0.43
60002	25, 15, 15	30, 20, 20	$2^{-10}, 2^{-10}, 2^{-8}$	0.56, 0.44, 0.44	0.65, 0.66, 0.7
64702	50, 25	30, 30	$2^{-6}, 2^{-6}$	0.64, 0.71	0.46, 0.55
75202	30, 10, 30, 40	30, 10, 30, 10	$2^{-10}, 2^{-10}, 2^{0}, 2^{-10}$	0.71, 0.73, 0.69, 0.68	0.83, 0.82, 0.62, 0.61
80702	30, 30, 50	20, 10, 30	$2^{-10}, 2^{-10}, 2^{-6}$	0.51, 0.63, 0.61	0.63, 0.45, 0.43
85202	10, 15	30, 20	$2^{6}, 2^{8}$	0.25, 0.65	0.79, 0.53
93402	50, 20	10, 30	$2^{-10}, 2^{0}$	0.57, 0.44	0.52, 0.75
93902	50, 50, 40	30, 30, 10	$2^{-10}, 2^{-10}, 2^{-10}$	0.71, 0.75, 0.46	0.51, 0.61, 0.68
94402	10, 50, 25, 10	10, 10, 30, 40	$2^{4}, 2^{-10}, 2^{-6}, 2^{-10}$	0.48, 0.82, 0.29, 0.49	0.62, 0.36, 0.71, 0.86
95202	10, 30, 10, 40	10, 30, 10, 30	$2^{-10}, 2^{0}, 2^{-10}, 2^{0}$	0.74, 0.28, 0.68, 0.55	0.66, 0.62, 0.66, 0.55
96002	10, 20, 30, 30	40, 40, 30, 30	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.84, 0.46, 0.58, 0.56	0.56, 0.67, 0.57, 0.60
98102	25, 50	40, 20	$2^{8}, 2^{-2}$	0.58, 0.63	0.32, 0.58
101702	10, 40	10, 10	$2^{0}, 2^{0}$	0.49, 0.60	0.60, 0.51
102202	45, 25, 50, 50	30, 40, 30, 10	$2^{0}, 2^{6}, 2^{4}, 2^{4}$	0.12, 0.25, 0.35, 0.24	0.88, 0.78, 0.59, 0.72
104602	20, 30	10, 10	$2^{-10}, 2^{-10}$	0.51, 0.8	0.58, 0.60
109502	10	10	2^{-10}	0.56	0.54
112802	10, 15, 15	10, 10, 10	$2^{-10}, 2^{8}, 2^{8}$	0.77, 0.49, 0.42	0.37, 0.42, 0.5
113902	15, 15, 10	20, 30, 40	$2^{-10}, 2^{-10}, 2^{-10}$	0.35, 0.60, 0.59	0.65, 0.76, 0.64
114702	35, 45, 10, 15, 15	10, 10, 30, 30, 10	$2^{0}, 2^{0}, 2^{0}, 2^{0}, 2^{0}$	0.25, 0.36, 0.44, 0.4, 0.57	0.73, 0.55, 0.63, 0.63, 0.53
114902	25, 35, 35, 35	10, 10, 10, 40	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.70, 0.31, 0.31, 0.57	0.23, 0.60, 0.71, 0.66
123902	10, 20	40, 10	$2^{-6}, 2^{-4}$	0.92, 0.31	0.80, 0.88
Overall	-	-	-	0.56 ± 0.19	0.63 ± 0.13

Table C.4: Training parameters and performance obtained for each patient with the Dynamic Weighted Ensemble and the Chronological data partitioning and iterative retraining method.

Chronological - Dynamic Weighted Ensemble

Patient	SOP	k	C	SS_{sample}	SP_{sample}
402	10, 50	10, 40	$2^{-10}, 2^{6}$	0.69, 0.32	0.65, 0.68
8902	15, 20	10, 10	$2^{0}, 2^{4}$	0.90, 0.95	0.68, 0.80
11002	10	20	2^{-10}	0.53	0.61
16202	45, 15, 15, 50	10, 30, 30, 40	$2^{-10}, 2^{2}, 2^{2}, 2^{6}$	0.83, 0.47, 0.43, 0.45	0.61, 0.70, 0.70, 0.69
21902	40	40	2^{-10}	0.84	0.58
23902	50, 45	40, 10	$2^{-8}, 2^{-10}$	0.50, 0.66	0.53, 0.44
26102	50	40	2^{6}	0.66	0.35
30802	35, 10, 10, 25, 50	40, 20, 20, 40, 10	$2^{-8}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.92, 0.58, 0.75, 0.65, 0.76	0.78, 0.83, 0.53, 0.68, 0.52
32702	15, 15	10, 10	$2^{-2}, 2^{-10}$	0.67, 0.81	0.70, 0.63
45402	50	40	2^{-10}	0.67	0.55
46702	30, 10	40, 40	$2^{-2}, 2^{8}$	0.37, 0.59	0.69, 0.69
50802	15, 20	30, 30	$2^{-10}, 2^{-10}$	0.82, 0.73	0.70, 0.8
53402	15	20	2^{-4}	0.22	0.72
55202	10, 10, 10, 10, 10	10, 10, 10, 10, 10	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}, 2^{-8}$	0.84, 0.73, 0.70, 0.76, 0.75	0.52, 0.74, 0.56, 0.62, 0.57
56402	25	20	2^{-10}	0.46	0.72
58602	10, 10, 20	40, 40, 30	$2^{-10}, 2^{-10}, 2^{-10}$	0.15, 0.26, 0.33	0.78, 0.83, 0.76
59102	50, 45	10, 10	$2^{-10}, 2^{-10}$	0.41, 0.70	0.42, 0.26
60002	25, 15, 15	30, 20, 20	$2^{-10}, 2^{-10}, 2^{-8}$	0.56, 0.44, 0.44	0.65, 0.66, 0.70
64702	50, 25	30, 40	$2^{-6}, 2^{-4}$	0.64, 0.64	0.46, 0.55
75202	30, 10, 50, 10	30, 20, 40, 40	$2^{-10}, 2^{-10}, 2^{-8}, 2^{-6}$	0.82, 0.64, 0.68, 0.54	0.75, 0.82, 0.74, 0.69
80702	30, 30, 50	20, 10, 30	$2^{-10}, 2^{-10}, 2^{-6}$	0.51, 0.63, 0.61	0.63, 0.45, 0.43
85202	10, 15	30, 40	$2^{6}, 2^{2}$	0.25, 0.28	0.79, 0.76
93402	20, 15	10, 10	$2^{-10}, 2^{8}$	0.64, 0.63	0.43, 0.48
93902	50, 50, 40	30, 30, 10	$2^{-10}, 2^{-10}, 2^{-10}$	0.71, 0.75, 0.46	0.51, 0.61, 0.68
94402	10, 10, 10, 10	10, 40, 40, 40	$2^{4}, 2^{-10}, 2^{-10}, 2^{-8}$	0.48, 0.71, 0.53, 0.57	0.62, 0.42, 0.51, 0.56
95202	10, 40, 40, 40	10, 40, 40, 40	$2^{-10}, 2^{6}, 2^{2}, 2^{2}$	0.77, 0.56, 0.59, 0.68	0.43, 0.43, 0.36, 0.42
96002	10, 20, 30, 30	40, 40, 30, 30	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.84, 0.46, 0.58, 0.56	0.56, 0.67, 0.57, 0.6
98102	25, 50	40, 40	$2^{8}, 2^{8}$	0.58, 0.69	0.32, 0.54
101702	10, 50	20, 30	$2^{-6}, 2^{-6}$	0.50, 0.46	0.57, 0.55
102202	45, 25, 50, 50	30, 40, 30, 10	$2^{0}, 2^{6}, 2^{4}, 2^{4}$	0.12, 0.25, 0.35, 0.24	0.88, 0.78, 0.59, 0.72
104602	20, 25	10, 10	$2^{-10}, 2^{-10}$	0.51, 0.54	0.58, 0.61
109502	30	10	2^{-6}	0.34	0.71
112802	10, 15, 15	10, 10, 10	$2^{-10}, 2^{8}, 2^{8}$	0.77, 0.49, 0.42	0.37, 0.42, 0.5
113902	15, 15, 15	20, 20, 20	$2^{-10}, 2^{-10}, 2^{-10}$	0.35, 0.44, 0.41	0.65, 0.75, 0.68
114702	50, 30, 30, 45, 50	30, 40, 40, 40, 40	$2^{-8}, 2^{4}, 2^{-10}, 2^{4}, 2^{4}$	0.46, 0.24, 0.28, 0.36, 0.44	0.41, 0.61, 0.64, 0.61, 0.57
114902	25, 35, 35, 35	10, 10, 10, 40	$2^{-10}, 2^{-10}, 2^{-10}, 2^{-10}$	0.7, 0.31, 0.31, 0.57	0.23, 0.6, 0.71, 0.66
123902	10, 10	40, 40	$2^{-6}, 2^{-10}$	0.92, 0.46	0.80, 0.88
Overall	-	-	-	0.56 ± 0.19	0.61 ± 0.14

Table C.5: Testing parameters and performance obtained for each patient with the Backwards-Landmark Window and the Chronological data partitioning and iterative re-training method.

Chronological - Backwards-Landmark Window

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	40, 20	0.50	1.85	0.32 ± 0.24	0.00	•
8902	2	20, 30	0.50	0.19	0.12 ± 0.21	0.00	•
11002	1	10	1.00	0.00	0.00 ± 0.00	0.00	•
16202	4	15, 10, 50, 20	0.75	6.98	0.51 ± 0.21	0.00	•
21902	1	40	1.00	1.23	0.47 ± 0.50	0.00	•
23902	2	50, 10	1.00	2.74	0.58 ± 0.34	0.00	•
26102	1	50	1.00	2.88	0.67 ± 0.47	0.00	•
30802	5	10, 10, 50, 25, 50	0.20	1.51	0.27 ± 0.21	0.97	
32702	2	15, 25	0.50	1.67	0.17 ± 0.24	0.00	•
45402	1	15	1.00	0.23	0.13 ± 0.34	0.00	•
46702	2	30, 10	0.50	3.54	0.48 ± 0.38	0.41	
50802	2	15, 25	0.50	0.90	0.10 ± 0.24	0.00	•
53402	1	40	1.00	0.03	0.07 ± 0.25	0.00	•
55202	5	35, 10, 25, 10, 10	0.20	0.29	0.15 ± 0.14	0.02	•
56402	1	25	0.00	4.77	0.50 ± 0.50	1.00	
58602	3	10, 10, 10	0.33	0.71	0.06 ± 0.12	0.00	•
59102	2	20, 15	1.00	1.00	0.38 ± 0.21	0.00	•
60002	3	20, 20, 40	0.33	0.28	0.14 ± 0.21	0.00	•
64702	2	30, 20	0.50	0.52	0.35 ± 0.39	0.02	•
75202	4	30, 10, 45, 35	0.25	0.02	0.01 ± 0.04	0.00	•
80702	3	35, 50, 15	0.67	2.17	0.16 ± 0.17	0.00	•
85202	2	15, 40	0.50	9.62	0.37 ± 0.29	0.01	•
93402	2	15, 20	0.50	0.31	0.03 ± 0.12	0.00	•
93902	3	50, 50, 15	0.33	1.91	0.54 ± 0.25	1.00	
94402	4	10, 40, 15, 15	0.50	2.50	0.20 ± 0.22	0.00	•
95202	4	10, 35, 40, 10	0.25	1.10	0.18 ± 0.2	0.04	•
96002	4	25, 40, 35, 45	0.25	1.50	0.51 ± 0.19	1.00	
98102	2	35, 45	1.00	12.98	0.65 ± 0.37	0.00	•
101702	2	30, 50	0.50	0.07	0.05 ± 0.15	0.00	•
102202	4	50, 45, 50, 50	0.00	0.32	0.08 ± 0.11	1.00	
104602	2	25, 50	1.00	0.74	0.23 ± 0.25	0.00	•
109502	1	10	0.00	0.47	0.03 ± 0.18	0.84	
112802	3	10, 15, 15	0.33	1.09	0.28 ± 0.24	0.12	
113902	3	45, 15, 45	1.00	1.36	0.42 ± 0.21	0.00	•
114702	5	35, 15, 15, 50, 15	0.00	0.16	0.05 ± 0.09	1.00	
114902	4	25, 50, 50, 20	0.25	0.03	0.03 ± 0.08	0.00	•
123902	2	10, 10	0.50	0.14	0.03 ± 0.12	0.00	•
Overall	-	-	0.53 ± 0.33	1.83 ± 2.71	0.23 ± 0.48	-	28

Table C.6: Testing parameters and performance obtained for each patient with the Seizure-batch Regression and the Chronological data partitioning and iterative retraining method.

Chronological - Seizure-batch Regression

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	10, 15	0.00	9.38	0.25 ± 0.34	1.00	
8902	2	20, 20	0.50	0.31	0.22 ± 0.25	0.00	•
11002	1	10	0.00	11.96	0.37 ± 0.48	1.00	
16202	4	45, 15, 10, 30	0.00	0.40	0.13 ± 0.15	1.00	
21902	1	10	1.00	0.19	0.30 ± 0.46	0.00	•
23902	2	50, 45	1.00	1.17	0.33 ± 0.32	0.00	•
26102	1	50	1.00	0.05	0.10 ± 0.30	0.00	•
30802	5	50, 50, 50, 35, 35	0.80	0.61	0.39 ± 0.17	0.00	•
32702	2	15, 15	0.50	1.63	0.15 ± 0.23	0.00	•
45402	1	50	1.00	97.86	0.63 ± 0.48	0.00	•
46702	2	35, 25	0.50	0.52	0.22 ± 0.31	0.00	•
50802	2	15, 20	0.50	0.39	0.07 ± 0.17	0.00	•
53402	1	15	1.00	0.03	0.03 ± 0.18	0.00	•
55202	5	10, 10, 10, 10, 10	0.20	1.18	0.19 ± 0.18	0.35	
56402	1	25	1.00	1.55	0.27 ± 0.44	0.00	•
58602	3	10, 10, 20	0.67	0.22	0.03 ± 0.10	0.00	•
59102	2	15, 50	0.50	1.20	0.57 ± 0.33	0.85	
60002	3	25, 15, 15	0.33	0.64	0.19 ± 0.25	0.00	•
64702	2	50, 25	0.50	0.43	0.20 ± 0.24	0.00	•
75202	4	30, 10, 30, 40	0.00	0.10	0.05 ± 0.1	0.99	
80702	3	30, 30, 50	0.67	0.60	0.23 ± 0.20	0.00	•
85202	2	10, 15	0.50	0.50	0.08 ± 0.19	0.00	•
93402	2	50, 20	1.00	1.58	0.48 ± 0.33	0.00	•
93902	3	50, 50, 40	0.33	1.51	0.24 ± 0.23	0.02	•
94402	4	10, 50, 25, 10	0.25	5.11	0.35 ± 0.25	0.98	
95202	4	10, 30, 10, 40	0.25	0.41	0.06 ± 0.12	0.00	•
96002	4	10, 20, 30, 30	0.00	2.68	0.43 ± 0.23	1.00	
98102	2	25, 50	0.50	0.55	0.15 ± 0.29	0.00	•
101702	2	10, 40	0.50	0.24	0.20 ± 0.24	0.00	•
102202	4	45, 25, 50, 50	1.00	0.83	0.38 ± 0.19	0.00	•
104602	2	20, 30	0.50	0.82	0.28 ± 0.31	0.00	•
109502	1	10	0.00	0.00	0.00 ± 0.00	-	
112802	3	10, 15, 15	0.33	3.01	0.29 ± 0.21	0.13	
113902	3	15, 15, 10	1.00	2.39	0.72 ± 0.21	0.00	•
114702	5	35, 45, 10, 15, 15	0.00	0.17	0.07 ± 0.09	1.00	
114902	4	25, 35, 35, 35	0.00	0.21	0.13 ± 0.14	1.00	
123902	2	10, 20	0.50	0.18	0.05 ± 0.15	0.00	•
Overall	-	-	0.50 ± 0.36	4.07 ± 15.82	0.23 ± 0.49	-	24

Even though the sensitivity was above the one from the Surrogate, it was not considered as statistically validated as the FPR/h is extremely large (•).

Table C.7: Testing parameters and performance obtained for each patient with the Dynamic Weighted Ensemble and the Chronological data partitioning and iterative retraining method.

Chronological - Dynamic Weighted Ensemble

Patient	Evaluated seizures	SOP	SS	FPR/h	SS Surrogate	p-value	Above chance
402	2	10, 50	1.00	2.23	0.20 ± 0.24	0.00	•
8902	2	15, 20	0.50	0.54	0.18 ± 0.27	0.00	•
11002	1	10	0.00	14.37	0.33 ± 0.47	1.00	
16202	4	45, 15, 15, 50	0.50	0.77	0.21 ± 0.18	0.00	•
21902	1	40	0.00	1.50	0.40 ± 0.49	1.00	
23902	2	50, 45	1.00	2.59	0.50 ± 0.34	0.00	•
26102	1	50	1.00	0.88	0.47 ± 0.50	0.00	•
30802	5	35, 10, 10, 25, 50	0.80	1.45	0.32 ± 0.14	0.00	
32702	2	15, 15	0.50	1.86	0.27 ± 0.31	0.00	•
45402	1	50	1.00	0.00	0.00 ± 0.00	0.00	•
46702	2	30, 10	0.00	0.88	0.13 ± 0.22	1.00	
50802	2	15, 20	1.00	1.21	0.15 ± 0.23	0.00	•
53402	1	15	1.00	2.03	0.27 ± 0.44	0.00	•
55202	5	10, 10, 10, 10, 10	0.60	2.21	0.22 ± 0.16	0.00	•
56402	1	25	1.00	3.30	0.43 ± 0.50	0.00	
58602	3	10, 10, 20	0.33	0.48	0.08 ± 0.14	0.00	•
59102	2	50, 45	1.00	1.23	0.47 ± 0.31	0.00	•
60002	3	25, 15, 15	0.67	1.48	0.23 ± 0.21	0.00	•
64702	2	50, 25	1.00	2.14	0.53 ± 0.31	0.00	•
75202	4	30, 10, 50, 10	0.50	0.41	0.10 ± 0.17	0.00	•
80702	3	30, 30, 50	1.00	1.52	0.52 ± 0.19	0.00	•
85202	2	10, 15	0.00	0.64	0.10 ± 0.24	0.98	
93402	2	20, 15	0.50	2.30	0.45 ± 0.35	0.22	
93902	3	50, 50, 40	0.33	2.05	0.39 ± 0.27	0.86	
94402	4	10, 10, 10, 10	0.75	4.16	0.29 ± 0.21	0.00	•
95202	4	10, 40, 40, 40	1.00	1.99	0.49 ± 0.25	0.00	•
96002	4	10, 20, 30, 30	0.75	2.39	0.39 ± 0.21	0.00	•
98102	2	25, 50	1.00	0.90	0.20 ± 0.31	0.00	•
101702	2	10, 50	0.50	4.05	0.40 ± 0.37	0.08	
102202	4	45, 25, 50, 50	0.50	0.03	0.00 ± 0.00	0.00	•
104602	2	20, 25	0.00	0.79	0.10 ± 0.20	0.99	•
109502	1	30	0.00	0.00	0.00 ± 0.00	-	
112802	3	10, 15, 15	0.67	2.03	0.26 ± 0.22	0.00	•
113902	3	15, 15, 15	0.67	1.31	0.23 ± 0.21	0.00	•
114702	5	50, 30, 30, 45, 50	0.80	0.86	0.44 ± 0.2.0	0.00	•
114902	4	25, 35, 35, 35	0.75	0.48	0.26 ± 0.18	0.00	•
123902	2	10, 10	0.00	0.97	0.07 ± 0.17	0.98	•
Overall	-	-	0.61 ± 0.36	1.84 ± 2.31	0.25 ± 0.50	-	28