Applied Intelligence
https://doi.org/10.1007/s10489-023-04702-5

Feature engineering of EEG applied to mental disorders: a systematic mapping study

Sandra García-Ponsoda[1,2] · Jorge García-Carrasco[1] · Miguel A. Teruel[1,3] · Alejandro Maté[1,3] · Juan Trujillo[1,3]

Accepted: 9 May 2023

Abstract

Around a third of the total population of Europe suffers from mental disorders. The use of electroencephalography (EEG) together with Machine Learning (ML) algorithms to diagnose mental disorders has recently been shown to be a prominent research area, as exposed by several reviews focused on the field. Nevertheless, previous to the application of ML algorithms, EEG data should be correctly preprocessed and prepared via Feature Engineering (FE). In fact, the choice of FE techniques can make the difference between an unusable ML model and a simple, effective model. In other words, it can be said that FE is crucial, especially when using complex, non-stationary data such as EEG. To this aim, in this paper we present a Systematic Mapping Study (SMS) focused on FE from EEG data used to identify mental disorders. Our SMS covers more than 900 papers, making it one of the most comprehensive to date, to the best of our knowledge. We gathered the mental disorder addressed, all the FE techniques used, and the Artificial Intelligence (AI) algorithm applied for classification from each paper. Our main contributions are: (i) we offer a starting point for new researchers on these topics, (ii) we extract the most used FE techniques to classify mental disorders, (iii) we show several graphical distributions of all used techniques, and (iv) we provide critical conclusions for detecting mental disorders. To provide a better overview of existing techniques, the FE process is divided into three parts: (i) signal transformation, (ii) feature extraction, and (iii) feature selection. Moreover, we classify and analyze the distribution of existing papers according to the mental disorder they treat, the FE processes used, and the ML techniques applied. As a result, we provide a valuable reference for the scientific community to identify which techniques have been proven and tested and where the gaps are located in the current state of the art.

Keywords Electroencephalogram (EEG) · Feature engineering · Feature extraction · Feature selection · Machine learning · Mental disorders

1 Introduction

During 2019, around 968 million people suffered from some form of mental disorder, that is 1 out of 8 people around the world. One year later, because of the COVID-19 pandemic, this percentage increased significantly, rising to 26% for Anxiety disorder and 28% for Major Depressive Disorder (MDD) [1]. One of the most common disorders suffered by the population is MDD. Indeed, suicide is the third cause of death among 15-29-year-olds, influenced by MDD. Regarding young people, around 1 out of 5 children and adolescents suffer from some mental health issue. In addition, people with severe mental disorders have more chances to die prematurely in comparison with neurotypicals, specifically 10 to 20 years earlier in high-income countries and up to 30 years earlier in low-income countries [1]. Therefore, it is paramount to make a reliable diagnosis as early as possible. To the best of our knowledge, there is still a gap between people needing to be diagnosed, and access to effective and low-cost healthcare. With this work, we aim to help create an accurate, reliable, and accessible diagnosis, through the collection of Feature Engineering (FE) techniques, as well as Artificial Intelligence (AI) models that other researchers have used mainly to diagnose mental disorders.

The Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5) acts as a standard reference for psychiatry and it includes more than 450 different definitions of mental disorders [2]. This fact highlights the impact that mental disorders have on individuals and society in general. According to [3], over a third of the total European popula-

✉ Sandra García-Ponsoda
sandra.gp@ua.es

Extended author information available on the last page of the article

tion suffers from mental disorders nowadays, and only a third of all cases receive some kind of treatment, concluding that the burden of mental disorders has been considerably underestimated. Therefore, having efficient methods for an early diagnosis of mental disorders would be extremely helpful.

There are several procedures to diagnose mental disorders such as neurological exams [4], neuropsychological assessments [5] and neuroimaging modalities [6]. Physicians specialized in mental disorders usually turn to neuroimaging approaches for help and to improve the efficacy of the treatments. Neuroimages can be divided into two categories, depending on what type of data they collect: functional and structural. Functional neuroimages show information about the activity of the brain, while structural neuroimages capture the interior structures of the brain. Some of the most used functional neuroimages are Magnetoencephalography [7], functional Magnetic Resonance Imaging (fMRI) [8], Electroencephalogram (EEG) [9], and Positron Emission Tomography (PET) [10]. On the other side, part of the most usual structural neuroimaging modalities is structural MRI (sMRI) [11], Diffusion Tensor Imaging (DTI) [12] and Computed Tomography (CT) [13]. This study assesses papers that deal with functional data, specifically with EEG modalities, for the reasons set out below.

Electroencephalography consists in recording brain activity by measuring voltage fluctuations of brain regions via the placement of small electrodes around the scalp. This record is called EEG and it is widely used in the study and diagnosis of brain disorders such as Epilepsy [14], Dementia [15], Schizophrenia [16], and Alzheimer's Disease (AD) [17]. The most remarkable advantages of EEG are the following. First, EEG devices are relatively portable, easy to set up and non-invasive. Second, these devices are characterized by their high temporal resolution, being capable of recording brain signals with up to 1 millisecond of resolution, even though their spatial resolution is worse compared with other methods such as MRI. Finally, EEG devices are relatively inexpensive, compared to other technology devices used to collect brain data, such as CT scanners or fMRI and PET devices. As a result, the use of EEG is a good candidate for an efficient and affordable diagnosis of mental disorders; especially since it is easy to use in underdeveloped countries, where quality healthcare is not fully accessible.

Traditionally, EEG was visually interpreted by highly specialized experts, and it was characterized as being a difficult and time-consuming task, as the volume of information that EEG data provides is considerably large. Because of this, the use of AI techniques has been proposed to automate the process and to aid in the diagnosis and study of mental disorders. Such techniques fall into two subsets of AI itself, defined as Machine Learning (ML), and a subset of ML, Deep Learning (DL). One of the most common tasks in the field of EEG and mental disorders is classification, i.e., an ML model takes several features derived from EEG data as input and outputs a prediction, e.g., whether a patient has a mental disorder or not. The input features are extracted from the raw EEG by applying FE. Extracting and choosing the right set of features for a given problem is one of the most relevant factors, as it can make the difference between an unusable ML model and a simple, effective model. In other words, it can be said that the FE is crucial, especially when using data such as EEG.

Indeed, properly applying FE on EEG data to train AI models related to brain disorders is still a challenging task, as there is no general FE pipeline that performs well on every task. For example, the authors of [18] showed that the beta band power was a relevant feature for detecting individuals with Insomnia, as they had significant and robust increases in that feature, whereas [19] showed that features such as the Variance, Energy, Nonlinear Energy and Shannon Entropy of the raw EEG signals were relevant for the task of epileptic seizure detection. In other words, the set of relevant features depends on the task and/or dataset, and properly applying FE remains a challenging task.

Given the importance of FE in the diagnosis of mental disorders by means of ML, it is clear that a secondary study that compiles the works in this area would foster the development of new techniques and lead to improvements in the diagnosis. Therefore, in this paper, we present a Systematic Mapping Study (SMS) with the purpose of clearly showing which FE techniques and ML models have been applied to each mental disorder in order to provide a way to easily find new research opportunities within the field. There are some secondary studies (reviews, surveys, and similar studies) on the topic of EEG and ML models applied to brain disorders, such as [20, 21]. Nevertheless, our work contains more significant contributions, as we can see in Table 1. Moreover, we will also share some insights and issues that we found after carefully analyzing the results of the SMS, as well as providing recommendations related to research directions. In order to help researchers to introduce new research opportunities discovered via this SMS, it is also included a brief description of other secondary studies that we found when collecting papers which can act as a starting point for future investigation gaps. It is worth noting that we do not report the efficiency achieved in each paper. That is mainly because it would not be correct to compare the accuracy obtained with different databases, since almost every paper selected uses a different one. In addition, as we present an SMS, we have only analyzed the abstract of each paper due to the number of works selected and we were not able to gather the databases used by reading only the abstract.

The rest of this paper is structured as follows: Sect. 2 presents the relevant concepts and background used in the diagnosis through the EEG field. Works related to the present study will be presented in Sect. 3. The methodology used in this work is thoroughly described in Sect. 4, including the

Table 1 Main contributions of our work in comparison with the closest related works [20, 21]

Characteristics	This work	de Almeida et al. [20]	Rivera et al. [21]
Writing language	English	Portuguese	English
Year of publication of selected papers	$\leq$ 2022	2013 – 2020	2016 – 2020
Number of papers reviewed	905	144	46
Number of brain disorders collected	15	9	9
Deals with transformation techniques	✓	×	×
Deals with feature extraction techniques	✓	✓	×
Deals with feature selection techniques	✓	✓	×
Deals with ML and DL algorithms	✓	×	✓

statement of research questions, how the search was conducted, the definition of inclusion and exclusion criteria, screening of papers, selection of keywords and the results of the whole process. In addition to the methodology, we also provide a description of reviews focused on this topic that could act as a starting point for new researchers in the field. Finally, a discussion of the results is presented in Sect. 5, followed by the conclusions of the study in Sect. 6.

2 Brain disorders, EEG, FE and ML

In this Section, we introduce the necessary concepts required to follow this paper. First, we briefly describe the brain disorders selected in Sect. 2.1. A description of EEG will be presented in Sect. 2.2, followed by an overview of FE techniques in Sect. 2.3 and a brief explanation of ML in Sect. 2.4.

2.1 Brain disorders

One of the objectives of this paper is to bring together all the FE techniques that have been used to classify, by means of AI, different brain disorders, as they affect society to a great extent. Along this subsection, we present a mental disorders overview and a brief description of the brain disorders selected.

According to [2], a mental disorder is described as a syndrome characterized by a clinically significant alteration in an individual's cognitive state, emotional regulation, or behavior that reflects dysfunction of the psychological, biological, or developmental processes underlying his or her mental function. Mental disorders can affect important areas of life, such as school or work performance, relationships with family and friends, and the ability to participate in the community. Fortunately, many mental disorders can be treated effectively at low cost, but there is still a gap between those who need care and those who have access to it. Greater investment is needed on all fronts, including increased awareness of mental health, access to effective treatments, and research to identify new treatments and diagnoses. The role of mental health in achieving global development goals is gaining recognition, proof of which is that mental health is included in the Sustainable Development Goals [22].

We conducted this study by choosing papers that study brain-related disorders. They are mostly mental, although we have also taken others that are not classified as such. We decided to include neurological disorders in addition to mental ones, as they are among the most studied cases and we were able to extract a high number of FE techniques, which is the main objective of this study.

In the following, we will present all the disorders we have taken for our study, indicating which family of brain disorders they belong to. We have based this classification on [2, 23]. All the disorders we mention are found in our study because they come from papers in which work is done with FE techniques together with one or more brain disorders. As explained in more detail in section 4.3, we have taken all disorders appearing in the abstracts of the selected papers and then grouped them according to the disorder family to which they belong. It should be noted that if the same disorder appears in a considerable number of papers, then we use it as a category on its own, without adding it to the family it belonged to.

- **Anxiety Disorders**. We have found papers that work diagnosing or classifying Anxiety Disorder such as [24, 25].
- **Bipolar and Related Disorders**. We have located papers studying Bipolar disorder such as [26, 27].
- **Depressive Disorders**. Within this category, we have picked studies working on MDD such as [28–30].
- **Neurodevelopmental Disorders**. We have found papers focused on Attention-Deficit/Hyperactivity Disorder (ADHD) [31, 32], and Autism Spectrum Disorder (ASD) such as [33, 34]. We also include papers related to concentration such as [35] and mental tasks like [36] in the ADHD category.

- **Neurological Disorders**. Within this group, there are brain disorders that are not considered mental disorders. We have identified papers that work with AD [37], Mild Cognitive Impairment (MCI) [38] and dementia in general [39], which are neurodegenerative disorders. We call this group dementia. We have also found papers studying Parkinson's Disease (PD) such as [40], which is another neurodegenerative disorder. On the other hand, we have located studies on Migraine such as [41]. Within this category, we have a brain disorder that appears in most of the papers we have selected for this study, namely Epilepsy. We have selected papers classifying, predicting, or diagnosing epilepsy or its seizures such as [42–44].
- **Obsessive-Compulsive and Related Disorders**. In this family of disorders, we have located papers studying Obsessive-Compulsive Disorder (OCD) such as [45].
- **Schizophrenia Spectrum and Other Psychotic Disorders**. Several papers were found under the Schizophrenia category such as [46–48].
- **Sleep-Wake Disorders**. We have identified papers that work with FE techniques and (i) Sleep Apnea [49], which is a Breathing-Related Sleep Disorder, (ii) Insomnia Disorder [50], and (iii) Non-Rapid Eye Movement Sleep Arousal Disorder [51], which is included in the Parasomnias family. We joined these three disorders in a general group that we call Sleep Disorders.
- **Substance-Related and Addictive Disorders**. Within these disorders, we have located papers dealing with FE techniques and Alcohol Use Disorder (AUD) such as [52], Drug-Related Disorder [53], and Non-Substance-Related Disorder such as Gaming addiction [54]. As we have done with the Sleep Disorders group, we have grouped these three disorders in a new group that we will call Addictions. It is worth noting that in the DSM-5, the word addiction is omitted from the official classification due to its potentially negative connotation. Despite this, it is a term commonly used in many countries to describe severe problems related to compulsive and habitual use of substances or behaviors.
- **Trauma- and Stressor-Related Disorders**. We have identified studies focused on Post-Traumatic Stress Disorder (PTSD) such as [55, 56], and Stress-Related tasks such as [57, 58], that we just call Stress.

2.2 EEG

The human brain consists of millions of neurons, each of which acts as an electrical dipole that varies its polarity depending on whether the goal of the neuron is to make an excitatory synapse or to make an inhibitory synapse. To register this bioelectric activity, a neurophysiological and non-invasive technique, called EEG, is used. To obtain an EEG it is only required to place small electrodes on the patient's scalp, with the help of a helmet and/or a conductive gel. This helmet can be made from 1 electrode to hundreds of them. To use an international system that facilitates reproducibility, it is recommended to place the electrodes following the guidelines of the American Clinical Neurophysiology [59] or the International Federation of Clinical Neurophysiology [60]. An EEG contains the sum of all electric changes or potentials produced among the closest neurons. Therefore, the brain activity captured by the EEG is a combination of the information movements that occur in our brain in a certain period of time. It would be similar to the noise that we receive in a big city with a lot of traffic and annoying sounds from time to time.

One of the targets of analyzing the EEG signal is to find the most relevant characteristics of each signal. All EEG signals have two essential measurable characteristics: amplitude and frequency. The amplitude is directly proportional to the number of neurons that emitted their charge at the same time, and frequency counts the number of oscillations that the signal has per second [61].

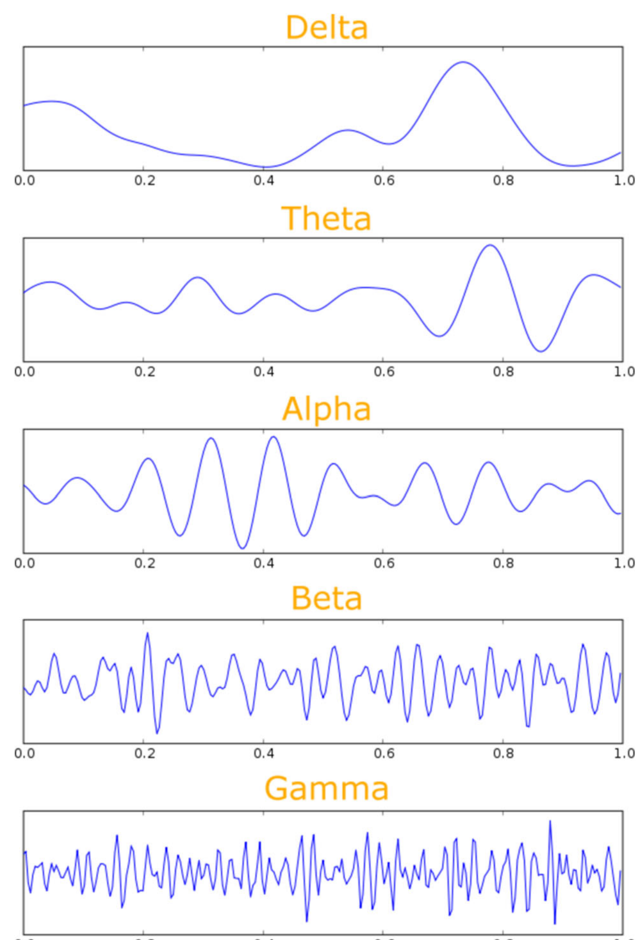


Fig. 1 Brain waves in EEG. Image extracted and modified from [65]

The signals received by the EEG come in a mitigated manner, so the electrodes located on the scalp have to be very powerful. Therefore, the EEG contains a lot of unwanted noises. That is why we have to do a cleaning process to select only the signal frequencies that we need. In the literature we can find, in a general manner, a classification of signal frequencies into bands. Below, we describe each band and relate them to different behaviors and mental states of the brain, based on [62–64]. Also, we provide these band waves in a graphical manner in Fig. 1.

- **Delta wave**. Every wave that has between 0.1 and 4 Hz is classified as a Delta wave. These types of waves are characterized by having the highest amplitude and by being the slowest waves. These waves are related to the grey matter of the brain and appear during the sleep stages. Delta waves are abnormal in adults when they are awake. When these waves are present, the growth hormone and the production of melatonin are stimulated. We could find them on the frontal side of the brain in adults, and on the posterior side in children.
- **Theta wave**. Theta waves vary between 4 and 8 Hz. We can see, through EEG, these types of waves in meditation and deep relaxation. These brainwaves are normal in children under their teens and abnormal for adults. These waves are in the thalamic region, a part of the brain located in the central area of the brain base, between the two hemispheres, involved in regulating the activity of the senses.
- **Alpha wave**. Waves between 8 and 13 *Hz* are called Alpha waves. These waves are related to the white matter of the brain. Alpha waves are present in all age groups, especially in relaxed and close-eyed adults. These waves are slightly higher in the non-dominant hemisphere. We can better gather Alpha waves from the occipital and parietal regions of the brain.
- **Beta wave**. Brainwaves between 13 and 30 *Hz* are called Beta waves. They are associated with behaviors and actions and these waves are related to our five senses. They appear when we talk, make decisions, solve problems, judge, and be on alert or focused. Beta waves are seen in the frontal and parietal lobes.
- **Gamma wave**. Waves that have more than 30 *Hz* are classified as Gamma waves. This type of brainwaves is characterized by having the smallest amplitude and by being the fastest waves. They are associated with perception and consciousness and appear during hyper alertness. Gamma waves induce the production of serotonin and endorphins. They are seen in the somatosensory cortex, located in the anterior part of the parietal lobe, and it is responsible for receiving and processing sensory information.

Unlike other techniques such as MRI or PET, acquiring an EEG device is much cheaper and it is possible to use it in other places less prepared than a hospital. In addition, EEG is a technique that gathers signals with very low frequencies without being an invasive technique. The characteristics of the EEG technique make this method capable of recording brain activity for hours or days, with the advantage of collecting information not only on brain activity but also the time related-information. As a result, the EEG gathers large amounts of data because of its precision and sophistication. Unfortunately, unwanted data is also collected, so a preprocess is needed to acquire only the necessary data. The EEG is a good choice if we want high temporal resolution. Although it does not have a spatial resolution as good as other techniques such as MRI or PET, there are some triangulation techniques that could improve the spatial resolution of EEG signals.

2.3 Feature Engineering

FE is the act of extracting features from raw data and transforming them into formats that are suitable for ML models [66]. Choosing the right set of features can make the difference between an unusable model and a simple, effective one. In other words, FE is a crucial step in the ML pipeline, especially when working with noisy, non-stationary data such as EEG.

FE comprises a vast set of techniques and methods, which can be divided into three subgroups: transformations, feature extraction and feature selection. The appropriate set of techniques depends largely on the problem, the type of data and the AI model to be used. Generally, the process of FE can be viewed as a pipeline composed of the aforementioned subgroups: (i) the raw data is transformed into a better format, then (ii) several features are extracted from said format, and finally (iii) the right set of features are extracted, ready to input to the ML or DL model (Fig. 2). A brief introduction to each part of the pipeline is presented below:

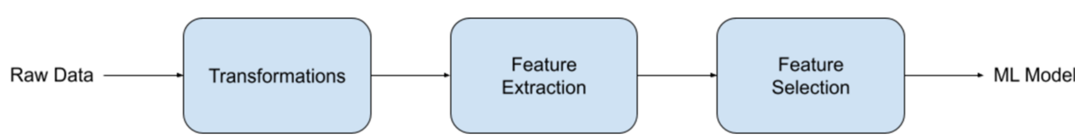


Fig. 2 The FE process can be thought as a pipeline: transformation of raw data, feature extraction and finally feature selection. Created by authors with *Google Drawings*

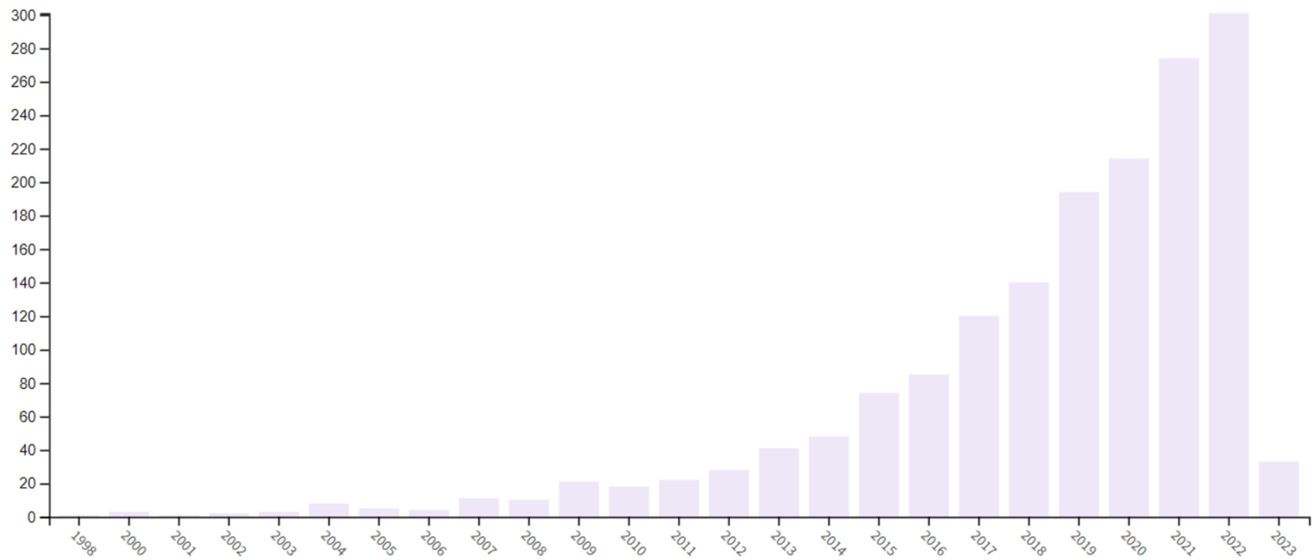


Fig. 3 Year-wise distribution of works related to the topic of signal processing of EEG for ML. Citation Report graphic is derived from Clarivate Web of Science [69], Copyright Clarivate 2023. All rights reserved

Transformations. It comprises any transformation that is applied directly to the raw data. In the case of EEG data, it is common to transform the signal from time domain to other domains, such as the frequency or time-frequency domains via Fourier Transform (FT) [67] or Wavelet Transform (WT) [68], respectively.

Feature extraction. It encloses any technique used to obtain hidden features from the transformed data. In the case of EEG data, features such as Entropy, Energy, or Fractal Dimension (FD) are frequently extracted.

Feature selection. Any method or metaheuristic that is used to select the most significant features for the current problem. This is specifically important in the case of EEG data, as the number of extracted features is typically large.

As can be seen in Fig. 3, there has been an increasing number of works focusing on the topic of EEG signal processing for ML. This upward trend is expected to continue as signal processing techniques play a key role in mental disorder detection by means of EEG. As previously stated, properly performing FE is crucial for the performance of an ML model. This becomes even more important when working with highly non-linear and non-stationary signals such as EEG. Given that the proper set of FE techniques depends on the task, the data and the model used, it remains a challenge to select the right set of FE techniques for mental disorder detection.

2.4 Machine Learning

According to [70], "*A ML algorithm is an algorithm that is able to learn from data*". ML has been successfully applied to a wide set of different problems, such as cancer detection [71] or credit risk assessment [72]. This ability to learn from data makes ML algorithms an excellent candidate for EEG data, as the volume of information it provides is large and difficult to be interpreted by humans. Next, we will provide a brief introduction of the concepts needed to follow this work.

As shown in Fig. 4, ML algorithms can be grouped by their learning method:

Supervised Learning. This group comprises any ML algorithm that learns from *labeled* data, i.e. the learned ML model tries to predict a given label. The model can be further

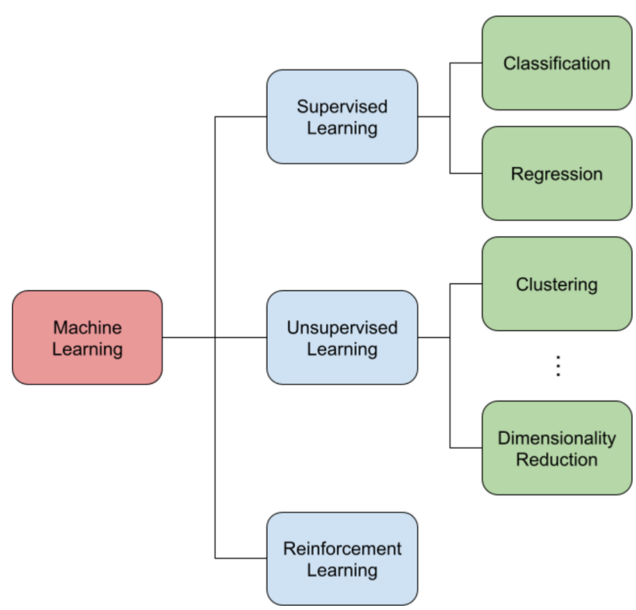


Fig. 4 ML algorithms can be grouped depending on its learning method. Created by authors with *Google Drawings*

subdivided by the type of label, if the label is categorical, it is a *classification* algorithm and if the label is numerical, then it is *regression*.

Unsupervised Learning. Opposite to the previous group, unsupervised learning includes any algorithms that learn from *unlabelled* data. This group can be divided into several subgroups depending on the main task of the algorithm. The best known within this type are (i) clustering, which refers to the task of grouping similar data into high-level clusters, and (ii) dimensionality reduction, which consists in reducing the number of features while minimizing information loss.

Reinforcement Learning. It refers to algorithms based on intelligent agents that learn by interacting with their environment.

It is important to remark that there is no clear line that separates these groups, and some ML algorithms may be difficult to classify into such groups. For example, there are ML algorithms that learn in a semi-supervised manner, that is, they learn on data that is partially labeled. However, it is still a valid classification for most of the problems.

DL is a subset of ML that emerged as a solution to some of the problems that classical ML algorithms struggled to solve. Specifically, ML algorithms may have difficulties in problems where the data is so complex that the manual engineering of features becomes unfeasible. As an example, the task of classifying a digit from an image is considerably difficult with ML. One could manually build detectors of multiple shapes, i.e., classify the image as 1 when a vertical line is detected. However, it would be costly and ineffective, since rotating or changing the typography of the digit would make the ML algorithm fail. On the other hand, DL algorithms are able to automatically learn representations from the data. Then, instead of having to manually craft the needed representations for classifying the image, the DL algorithm is able to automatically learn them, making it a preferable choice for this kind of problem. Therefore, DL frameworks could revolutionize the clinical applications for EEG-based diagnosis.

Figure 5 shows a year-wise comparison of the number of works related to the application of ML or DL on EEG data. It can be seen that ML is more prevalent than DL. However, since approximately 2012, the popularity of DL has been exponentially increasing and is expected to become prevalent if the current trends continue.

Recent developments in DL have led to promising results in the area of medical diagnostics, especially in the diagnosis of mental and neurological disorders. Regarding Depression, [73] has proposed an EEG-based DL framework that automatically discriminates depressed and healthy controls and provides the diagnosis, achieving 98.32% accuracy, using a model based on a Convolutional Neural Network (CNN). This high accuracy makes it possible to use this DL model as an automatic diagnostic model for Depression. Going further, [74] has developed a DL model based on the attention technique that classifies Twitter data and predicts the depressed and non-depressed users, reaching 99.86% accuracy. This shows that early detection of depression is possible simply by analyzing social media posts, which could improve or save people's lives. A similar case occurs with Dementia. [75] has used a CNN-based model to diagnose AD from neuroimages, achieving 95.73% accuracy, which could serve as a computer-aided system for physicians who need to make an early diagnosis. [76] has used a DL approach to predict brain age using MRI of brain grey matter, showing that the difference between the predicted and the chronological brain age serves as a biomarker for early-stage neurodegeneration. Another successful case is Epilepsy, where the early detection of seizures through non-invasive and wearable devices could improve the management of Epilepsy. [77] has developed a Long-Short Term Memory (LSTM)-based DL model to detect and classify epileptic seizures in an ambulatory and in-hospital environment using wearable devices. The proposed model has yielded a mean Area Under the Curve (AUC) of 0.97 and 0.98 for ambulatory and in-hospital patients respectively. Thus, it demonstrates that the detection of motor epileptic seizures is possible through the use of wearable devices.

To summarize, DL models are excellent candidates for the detection of brain disorders, and especially mental disorders,

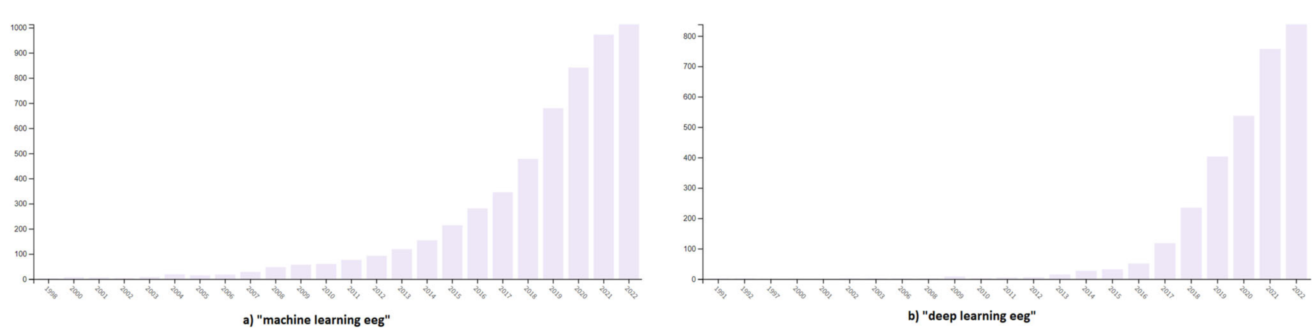


Fig. 5 Year-wise distribution of works related to a) the use of ML and b) the use of DL on EEG. Citation Report graphic is derived from Clarivate Web of Science [69], Copyright Clarivate 2023. All rights reserved

due to their automatic representation learning capabilities. On the other hand, traditional ML models do not have that capability and require their input features to be hand-crafted, but they are simpler to train and require less amount of data.

Despite its success, DL also has a disadvantage over ML, as DL algorithms are considerably less explainable than traditional ML algorithms. In ML, it is easier to explain or analyze why the model has made a certain prediction (e.g., which features have contributed the most to the prediction and how), whereas this becomes a complicated task in DL. The lack of explainability considerably hinders the application of DL to fields such as medicine [78, 79], where an erroneous decision can have critical consequences. In order to use a DL system to support decisions, the domain expert has to be able to understand the predictions of the model to avoid errors. However, great progress is being made in the field of eXplainable Artificial Intelligence (XAI) applied to the medical field [80, 81]. Finally, it is important to remark that, as mentioned above, DL models are able to automatically learn the adequate representations for each problem, thus reducing the need for FE. However, DL architectures can still greatly benefit from said techniques, as they can increase their performance and/or enable us to use a simpler architecture.

Since the aim of this work is to perform an SMS focused on FE, ML and EEG applied to mental disorders, the main task that will be encountered is to diagnose a mental disorder, i.e., to predict whether an individual has a mental disorder or not by analyzing their EEG data. Thus, it will be a supervised problem, since the main objective is to predict a label, and a classification problem, because the label is categorical (the individual has a mental disorder or not). Therefore, most of the algorithms found during the present work are expected to be classification algorithms. A brief description of some of the most common ML classification algorithms is presented below. Since presenting too extensive explanations would be out of the scope of this paper, we refer to [82, 83] as excellent resources to obtain in-depth descriptions of ML models.

- **Logistic Regression**. Logistic Regression is one of the most popular models due to its simplicity and interpretability. Logistic Regression models the log odds of the event as a linear function:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p \tag{1}$$

where p is the probability of an event or a specific class (e.g., a patient has ADHD), β_i are the coefficients learned by the model and x_i are the values of the input features. The coefficients of the model are typically learned by minimizing the negative log-likelihood. Once the parameters are learned, the probability p for a given sample can be predicted as

$$\begin{aligned} p &= \frac{1}{1+\exp-(\beta_0 + \beta_1 x_1 + ... + \beta_p x_p)} \\ &= \sigma(\beta_0 + \beta_1 x_1 + ... + \beta_p x_p) \end{aligned} \tag{2}$$

where σ is the sigmoid function. Despite the simplicity of the model, Logistic Regression models fail to detect more complex, nonlinear patterns, so they have limited predictive power.

- **Decision Tree (DT)**. The aim of DTs is to partition the data into smaller, more homogeneous groups. In other words, the data space is partitioned via a series of if-then statements, where each partition has an associated predicted label.
  DTs are mathematically simple models, which makes them easily interpretable. These models are also capable of handling different types of data, are robust to outliers and perform automatic feature selection. However, these models tend to be unstable, i.e. a slight difference in the data can drastically change the structure of the tree. Thus, interpreting the model becomes more difficult as the tree grows bigger. Moreover, the predictive performance may not be optimal because the trees partition the data into rectangular regions, which can hinder the detection of some patterns. Ensemble methods such as Random Forest, which combines many DTs into one, have been designed to combat these disadvantages.
- **Support Vector Machine (SVM)**. SVMs were introduced in the 90s for binary classification [84] and have been extended to regression and multilabel classification. SVMs were designed in the context of *robust regression*, i.e. building a model that is robust to outliers. Moreover, SVMs are powerful models which are able to extract nonlinear patterns from the data.
- **Naive Bayes**. The Naive Bayes is a probabilistic classifier that is based on applying the Bayes' rule. This rule answers the question "based on the predictors that we have observed, what is the probability that the outcome is class C_l?", which is mathematically described as [83]:

$$Pr[Y = C_l \mid X] = \frac{Pr[Y]\, Pr[X \mid Y = C_l]}{Pr[X]} \tag{3}$$

where X represents the input features and Y the class variable.

- **k-Nearest Neighbors (k-NN)**. The aim of this model is to classify a new sample based on the labels of the k-closest points in the training set given a distance metric. In the case of classification, we assign the most common class among the neighbors as the predicted class.

DL can also be used for classification, using architectures such as Multilayer Perceptron (MLP), CNN, and Recurrent Neural Networks (RNN). Refer to [70] for an excellent in-depth description of DL models. It is also expected to find algorithms beyond classification, such as clustering since it can be used to group individuals based on their similarities and extract insights.

3 Related works

To the best of our knowledge, there are only three works related to our study. The first one is an SMS focused on the diagnosis and prognosis of mental disorders using EEG and DL techniques [21]. This study is led by four research questions. The first one focuses on which mental disorder is diagnosed and prognosed by means of EEG and DL. The second research question identifies which DL techniques are applied. The third research question collects other biometric data used to diagnose and prognose mental disorders along with EEG signals. The last research question focuses on the source of the datasets used to carry out each reviewed paper. Afterwards, the authors elaborate several graphs with the distribution of the studies according to the answers to the research questions. It is important to say that, in this SMS, 46 out of 373 works were selected to do the mapping.

The second study that we found in the literature is a systematic review of ML algorithms used to analyze data from wearable devices and sensors [85]. 67 studies were selected out of the 1530 pre-selected works. To carry out that review, the authors propose four research questions that cover the whole processing of healthcare data. The first research question deals with the types of sensors or wearable devices used for gathering data. The second question consists in the use of one feature type for different kinds of data. The third research question refers to the use of ML algorithms to analyze healthcare data. The fourth question addresses how to combine, process, and analyze heterogeneous types of healthcare data. Subsequently, several graphs are displayed, showing the distributions of sensors used for monitoring symptoms, types of features extracted, ML and neural network algorithms used for the chosen analysis and evaluation criteria. In addition, they show the distribution of algorithms with the best performance.

Finally, we have found [20], which is the study most related to ours. Their SMS focuses on highlighting which neurological disorders, feature extraction techniques, feature selection methods and classifier algorithms have been studied the most in the past using EEG signals. In this research, 144 studies have been evaluated to compose the paper. Regarding the research questions, two are proposed: one is related to the feature extraction technique used and the other one is focused on the feature selection method applied. In addition, they add a section in which some quality criteria are proposed to give a quality grade to each reviewed paper. Graphically, they show the distribution of papers and mental disorders studied according to the year of publication and a bubble chart that displays the joint occurrence of feature extraction and selection techniques in the reviewed studies.

The first SMS, [21], is closely related to our work, but it focuses on DL techniques, whereas our study is focused on FE. On the other hand, although the second paper [85] includes EEG signals as e-health data, it focused on ML techniques and considered papers that study disorders like heart disease, diabetes, blood diseases or hypertension, which are not considered mental disorders. In addition, it selects studies that use wearable devices. Therefore, it is focused on ML algorithms that could provide meaningful results in a reasonable time period and with reasonable complexity to be used together with these wearable devices. Finally, the third SMS [20], although it is the most related to ours, is written in Portuguese, so it is not very accessible to the scientific community. Moreover, it excludes studies that are published before 2013 and we have no restriction in this regard. Whereas the other papers leave out a significant portion of the advances achieved, unlike ours which is more recent and covers a broader time span.

Therefore, after having presented the most related works to ours, we could say that our study is the first English-written SMS focused on FE of EEG signals used to identify mental disorders. Furthermore, the number of studies that we have gathered -6133- and the number of papers reviewed -905- are far superior to the other SMS that they could be compared with.

4 Methodology

SMS is a secondary study which aims to provide an overview of a research area by identifying the quantity, type of research, and results available within [86–88]. The difference between a primary and a secondary study is that the former presents direct advances in the research area, whereas the latter gathers data from such primary studies to extract insights. Other types of secondary studies, such as a Systematic Literature Review (SLR) [89], can also be used to provide such an overview. The main difference between them is that the SLR is focused on a detailed reading of a small number of papers, whereas the SMS is focused on a less detailed reading of a large number of papers. In this case, we decided that the SMS would provide better results since the research area of interest is quite broad and a large number of primary studies are expected.

This work will follow the methodology presented in [87]:

1. *Definition of Research Questions.* These questions will guide the whole process in order to achieve the desired goals.

2. *Conduct Search.* Searching for primary studies that could be related to the research questions by using string queries on several scientific databases.
3. *Screening of papers.* Several exclusion and inclusion criteria are defined in order to discard papers that are not related to the defined research questions.
4. *Keywording using Abstracts.* Reading the abstracts and looking for keywords that could characterize each paper. Then, the keywords are used to form higher-level groups.
5. *Data Extraction and Mapping Process.* The previous groups are used to answer the research questions via analysis and different visualizations. A frequency analysis will enable us to identify which topics have been exploited in the past, gaps in the literature, and future research directions.

As an additional step, we will briefly describe the secondary studies collected during the latter process, as this can act as a starting point for researchers that decide to take a certain research opportunity after analyzing the results of the SMS.

4.1 Definition of research questions

Since our study is focused on mental disorders and FE of EEG signals, we proposed the following research questions, that we will abbreviate as **RQ**s:

RQ1: Which mental disorders have been studied using ML, DL and FE of EEG signals?

RQ2: Which FE techniques have been used before feeding the data into ML algorithms?

- RQ2.1: Which transformation techniques have been used?
- RQ2.2: Which feature extraction techniques have been used?
- RQ2.3: Which feature selection techniques have been used?

RQ3: Which ML and DL techniques are used after applying FE techniques?

RQ4: Which secondary studies have been done in the research field?

The first, second and third **RQ**s have been selected according to the goal of this work, which is to gather all FE techniques used to mainly classify mental disorders, using AI models and EEG data. With **RQ1** we want to extract the mental or neurological disorder each work focuses on. It can be only one disorder or more than one in each paper collected. **RQ2** is the most important question of this work. We divided it into three different ones because several studies [90–92] consider that there is this number of steps (3) before feeding the AI model with data. **RQ2.1** deals with the data transformation techniques, in which the data change their domain, from raw EEG signal to 2D images, graphs networks, Wavelets or another domain. With **RQ2.2** we want to gather every feature extraction used to feed the AI models. It is essential to show all characteristics considered necessary to identify differences between control and brain disorder patients. **RQ2.3** collects all the algorithms used to select which features best capture the necessary information to classify the chosen brain disorder. In **RQ3** we want to gather all ML and DL techniques used as the last step to classify patients with control and brain disorders. It should be noted that we do not extract the accuracy achieved, nor the contributions and drawbacks of each study. This is mainly due to the fact that we have only read the abstract of each paper and we were seldom able to clearly find the contributions and drawbacks. There is also a reason not to collect the accuracy of each paper and it is because we consider it unfair to compare the performance of studies carried out with different databases. We have not collected the country or year of publication either, since the vast majority of the articles gathered did not provide us with this information. Lastly, **RQ4** is introduced as an addition to the whole methodology, and it will be answered by carefully selecting the secondary works retrieved by the search conducted in the section below. We thought that this addition would be helpful for researchers that decide to take a certain research opportunity after reading the results of this SMS. We decided not to include these studies in the related works section because these papers are not sufficiently related to our work.

4.2 Conducted search

In order to find relevant works related to the defined **RQ**s, the following terms were considered:

- EEG
- FE
- Feature extraction
- Feature selection
- Statistical parameters

These terms were composed into the query: *EEG AND (feature engineering OR feature extraction OR feature selection OR statistical parameters)*. The search was conducted in five popular public databases: *Scopus, IEEE Xplore, Web Of Science, ACM Digital Library* and *ScienceDirect*. The specific queries written on the syntax of each search engine are shown in Table 2.

It is important to remark that the *EEG* term was searched on the title, abstract and keywords, but the rest of the terms

Table 2 Queries used for each database

Database	Query
Scopus	TITLE-ABS-KEY((*eeg* OR electroencephalogra*)) AND (TITLE ("feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters") OR KEY("feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters"))
Web of Science	TS=((*eeg* OR electroencephalogra*)) AND (TI=("feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters") OR AK=("feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters") OR KP=("feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters"))
IEEE Xplore	("All Metadata": *eeg* OR electroencephalograp*) AND (("Document Title": "feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters") OR ("Index Terms": "feature engineering" OR "feature extraction" OR"feature selection" OR "statistical parameters"))
ScienceDirect	Title, abstract, keywords: eeg OR electroencephalogram OR electroencephalography Title: "feature engineering" OR "feature extraction" OR "feature selection" OR "statistical parameters"
ACM	((All: *eeg*) OR (All: electroencephalogra*)) AND ((Title: "feature engineering") OR (Title: "feature extraction") OR (Title: "feature selection") OR (Title: "statistical parameters")) AND ((Abstract: "feature engineering") OR (Abstract: "feature extraction") OR (Abstract: "feature selection") OR (Abstract: "statistical parameters"))

were only searched on the title and keywords. This is due to the fact that in some papers, DL is used directly on raw EEG signals, so they mention that there is no need for FE in the abstract. As no FE techniques are used, this will be an exclusion criterion, which will be defined in the next subsection.

4.3 Screening of papers

Once the search is conducted and the duplicated papers are removed, a set of inclusion and exclusion criteria is defined in order to refine the search results and only take the works that are relevant to the research questions of our study. The criteria defined to guide the filtering process are shown in Table 3.

We choose four inclusion criteria to gather the necessary papers with greater precision. With the aim of collecting as many FE techniques as possible, we have included conference papers (*i2*) as well as academic journal papers (*i3*). In addition, only primary studies were chosen (*i1*) due to the goal of this SMS. As we need quality papers, we added an English-written inclusion criterion (*i4*). It is worth noting that we have not added any criteria related to the date of publication, since we consider that there may be FE techniques, which were previously used without promising results, that now, with the ML and DL models, could be useful. We do not consider it necessary to discuss the three exclusion criteria (*e1*, *e2*, *e3*) because they are trivial.

As shown in Fig. 6, after removing duplicates, the inclusion criteria are applied in order to keep the papers that are relevant to our study and discard works such as book chapters, reviews, etc. This was done both automatically by using the search engines and manually by reading the titles and abstracts, to ensure that none of the selected papers bypassed the criteria. Then, the filtering process continues manually applying the different exclusion criteria by reading the title and abstract of the selected works. Once all the criteria have been applied, a total of 905 papers are left.

It is important to remark that, as previously stated, a total of 20 secondary studies were carefully retrieved in order to present them in a compilation after the whole SMS process. This will act as a starting point for researchers that decide to take a certain research direction. These 20 studies were retrieved out of the 72 studies that did not meet the inclusion criteria *i1* (Fig. 6).

4.4 Keywording of full text

Once the screening of papers is done and the final batch of papers is collected, it is time to classify each paper via the process denominated *Keywording*. As presented in [87],

Table 3 Inclusion and exclusion criteria defined to filter the obtained works

Criteria	
Inclusion	Exclusion
i1: Primary Study	*e1:* Does not treat any mental disorder
i2: Conference Paper	*e2:* Does not use FE techniques
i3: Academic Journal Paper	*e3:* Does not use EEG signals
i4: English-written	

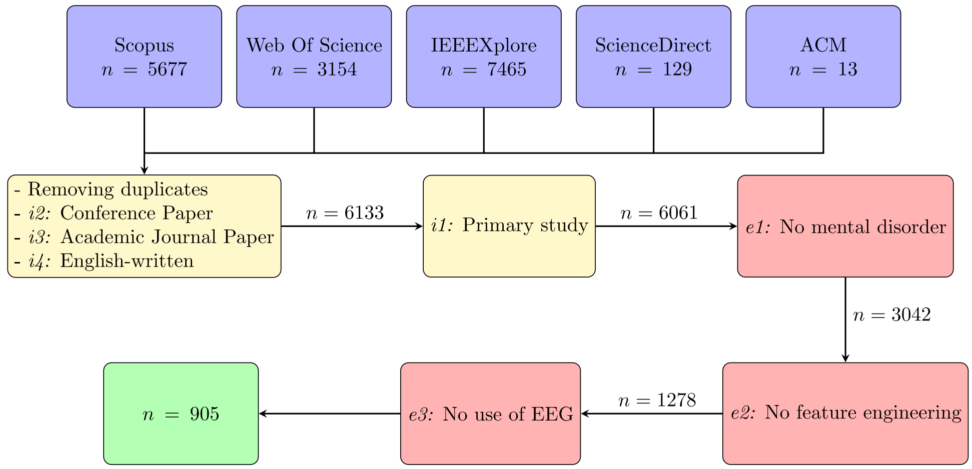


Fig. 6 Diagram of the filtering pipeline via the exclusion and inclusion criteria. Created by authors with *TikZ* package from LaTeX

keywording is a systematic process that ensures all papers are taken into account, while reducing the time needed in developing the classification scheme. Keywording is divided into two steps:

1. The authors read the abstracts and define keywords that characterize the work of each paper related to the different **RQ**s. This will result in a highly granular set of keywords, but it will help the authors to fully identify the context of the research. To illustrate this process, let us use an example. Suppose that we find a paper that has in its abstract the following: *Our goal is to distinguish ADHD from ASD subjects. To carry out this classification we use WT and extract nonlinear features to feed three AI models: SVM, k-NN and CNN-based model*. In this case, we would extract seven keywords. *ADHD* and *ASD* inside **RQ1**, *WT* for **RQ2.1**, nonlinear features in **RQ2.2** and *SVM*, *k-NN* and *CNN-based models* in **RQ3**, leaving **RQ2.3** without keywords. If more details are needed, all the tags for each reviewed work, together with its title and authors are in the supplementary information.
2. The keywords obtained are combined to form higher-level categories that will be used to answer the **RQ**s in the following subsection, except **RQ4** which will be answered separately in the last subsection. In this step, we have gathered all keywords and grouped them into more general groups. These groups were made following our criteria and observing the number of labels of each category. It should be noted that we have taken some labels out of the general categories if they were significant, i.e., if they appeared a considerable number of times. As an example, we have made a group called *Nonlinear features* in which we add all papers that do not specify which nonlinear features have been used, but we made other groups in which there are nonlinear features as well, such as *Chaotic features* group or *Complexity Measures*. Therefore, we know there are labels that can be in more than one group. If more details are needed, the Appendix 1 contains all the categories we have created along with the labels that make up each of them.

After performing the first step, a total of 634 keywords were obtained from all **RQ**s. Once the resulting keywords had been carefully analyzed, the opensource application *OpenRefine* [93] was used to combine them into general groups, resulting in 15, 14, 15, 8 and 14 categories associated to **RQ1**, **RQ2.1**, **RQ2.2**, **RQ2.3** and **RQ3** respectively. The categories defined for each **RQ** are briefly described below:

RQ1. *ADHD* is a neuropsychiatric disorder characterized by hyperactive-impulsive and/or inattentive behavior which has a high prevalence among young people, but can be carried onto adulthood [94]; *Addictions* are mental disorders characterized by the recurrent failure to control a compulsive behavior in order to obtain reward stimuli despite the negative consequences [95]; *Anxiety* is a feeling of worry and fear in a diffuse threat, which can be out of proportion and interfere with the daily lives of the affected [96]; *ASD* is a neurodevelopmental disorder characterized by deficits in social communication and the presence of restricted interests and repetitive behaviors [97]; *Dementia* is characterized by the deterioration of mental functioning in its cognitive, emotional and conative aspects [98]; *Depressive Disorders* are characterized by having a lowered mood, and the loss of interest and enjoyment during periods, among other symptoms [99]; *Dyslexia* occurs when an individual has significant difficulties with speed and accuracy of word decoding [100]; *Epilepsy* is a neurological disorder characterized by an enduring predisposition to generate epileptic seizures due to abnormal excessive neuronal activity in the brain [101]; *Migraine* is characterized by severe headache attacks, autonomic nervous system dysfunction, and in some patients, an aura involving neurological symptoms [102]; *OCD* is characterized by intrusive unwanted thoughts and/or images (obsessions) and ritualized repetitive behaviors (compulsions) [103]; *PD* is a neurodegenerative disorder that mainly affects motor function [104]; *PTSD* is a disorder that can be developed after a traumatic experience [105]; *Schizophrenia* is a severe psychiatric disorder mainly characterized by having delusions, hallucinations, psychotic episodes, marked alterations in cognition, and impaired functioning with high rates of disability [106]; *Sleep Disorders* include disorders related to sleep, namely *Sleep Apnea* [107], *Insomnia* [108] and *Sleep Arousals* [109]; *Stress* is a condition in which an individual is aroused and made anxious by an uncontrollable aversive challenge [110].

RQ2.1. *Common Spatial Patterns (CSP)* is a procedure used to decompose signals of two different groups into modes that are common to both groups and maximally suited to distinguish them [111]; *Complex Networks* refers to the methods used to transform the raw EEG signals into a graph, which can then be used to extract topological features [112]; *Discrete Cosine Transform (DCT)* is a frequency domain transformation which decomposes a signal into cosine waves with different frequencies [113]; *FT* represents a set of widely used techniques, which decompose a function depending on time into a set of functions that depend on frequency in order to obtain the so-called *spectrum* of frequencies [67]; *Frequency Transformation* comprises any transformation in the frequency domain that is not further specified; *Hilbert-Huang Transform (HHT)* enables us to obtain instantaneous frequency data, making this technique specially suitable for nonstationary and nonlinear data [114]; *Linear Transformation* comprises any linear transformation that is not further specified; *Mel-Frequency Cepstrum* is a transformation that can be described as a kind of "spectrum of a spectrum" [115]; *Mode Decomposition* is composed of all methods that decompose the signal without leaving the time domain, mainly *Empirical Mode Decomposition (EMD)* and variations [116]; *Nonlinear transformations* includes any nonlinear transformations that are not further specified; *Short Time Fourier Transform (STFT)* is a variant of the FT that decomposes the signal into time-frequency components, similar to the WT [117]; *Time-Frequency Transformation* includes transformations in the time-frequency domain that are not specified; similarly, *Time Transformation* comprises transformations in the time domain that are not specified; *WT* is similar to the FT so that it performs a decomposition into functions depending on frequency, but also on time, i.e. it gives *local frequency information* [118].

RQ2.2. *Autoencoder (AE)* [119] is a type of neural network used to extract features that efficiently represent the data. The AE is trained to map the input data to a smaller vector space, and then to reconstruct the original vector, i.e., it is trained to map the data into a smaller feature space while minimizing the loss of information. Therefore, it can be thought of as an automatic feature extraction technique; *Autoregressive Model (AR)* is used to describe signals with a set of parameters, and then those parameters can be used as features [120]; *Chaotic Features* refers to any variables that measure whether a system is chaotic or not, such as the *Lyapunov exponents* [121]; *Complexity Measures* can be used to estimate brain dynamics, which can be useful to study mental disorders by means of EEG signals. This group is composed of the *Kolmogorov Complexity (KC)*, *Lempel-Ziv Complexity (LZC)*, *epsilon-Compexity* [122] and fractal-related features, namely *FD*, *Correlation Dimension (CD)* and *Line Length* [123]; *Correlation Measures* can be used to measure the relation between two signals, in this case, two different channels. The main feature of this group is the *coherence* [124]; The *Energy* of a signal is defined as the area under its square magnitude. Nevertheless, energy is typically obtained via the summation of all frequency components of its spectra, thanks to the *Parseval theorem* [125]; *Entropy* can be used to measure the degree of randomness or unpredictability of the system, as well as being related to chaos. There exist many ways of computing entropy, such as *Sample Entropy (SampEn)*, *Approximate Entropy (ApEn)*, *Fuzzy Entropy (FuzzyEn)*, etc. [126]; *Frequency Domain Features* comprises any feature computed in the frequency domain that is not further specified; *Geometric Features* are extracted from 2D representations of the signal, for example by applying WT. Most features are focused on texture, such as *Local Binary Patterns* [127] and their variations and *Haralick Features* [128]; *Graph Features* describe the topology of the graphs obtained from the raw EEGs such as the *Degree of Centrality* [129] or the *Direct Transfer Function* [130]; *Nonlinear Features* is composed by any nonlinear feature that is not included in any of the other groups, and/or it is not specified; *Statistical Parameters* is comprised by any statistical measure of the signal that is not included in any of the other groups, such as the *variance*, *Hurst exponent* [131], *Hjorth parameters* [132], etc; *Tensor Decomposition* refers to any feature extraction technique that uses data in tensor (or matrix) form such as *singular value decomposition* [133], *Hermite decomposition* [134] or *Wishart distribution* [135]; Finally, the groups *Time Domain Features* and *Time-Frequency Domain Features* enclose any features that belong to the aforementioned domains and are not included in any of the previous groups.

RQ2.3. *Clustering* refers to the task of grouping similar data into high-level clusters. Once the clusters are formed, they can be used for feature selection by analyzing which features are most relevant to differentiate them [136]; *Distance Based Feature Selection* encloses all the methods that use the distance between data points to perform feature selection, for example, the Relief method [137]; *Genetic Algorithm* is referred to the group of feature selection metaheuristics based on the theory of evolution [138]; *L1 Regularization* is a regularization technique that can be used to perform feature selection, as it shrinks the coefficients of irrelevant features to zero [139]; *Metaheuristics* encloses all the feature selection metaheuristics that cannot be included in any of the presented groups. A metaheuristic is defined as a high-level procedure designed to solve complex optimization problems, in this case, a feature selection problem [140]; *Principal Component Analysis (PCA)* is a technique that consists of building successive uncorrelated variables called principal components that maximize variance[141]. It can be used either as a dimensionality reduction tool, by choosing the first k principal components, or as a feature selection tool. As the principal components will be linear combinations of the initial set of features, the coefficients (also called *loadings*) can be used to rank the importance of each feature; A *Statistical Test* is a procedure used to check whether there is enough evidence to reject a conjecture (called the *null hypothesis*) or not. This kind of tests can be used for feature selection by defining the null hypothesis as $H0$: The feature$X$ is relevant for the current model. If the null hypothesis is rejected by the test, then the feature is discarded; *Swarm Intelligence* refers to the set of optimization metaheuristics inspired by the behavior found in collective and decentralized biological systems [142].

RQ3. *DT* arranges a set of basic decisions into a tree structure in order to make the final prediction. The advantage of this model is that it can be easily explained [143]; *Ensemble* refers to any model that combines a set of weak classifiers to get a better global solution, such as Random Forest [144] or XGBoost [145]; *Clustering* refers to the task of grouping similar data into high-level clusters. This group is composed of any use of clustering that is not feature selection. If clustering is used for feature selection, it is then included in **RQ2.3**; *Naive Bayes* is a family of classification algorithms based on applying Bayes' theorem [146]; *Fuzzy Classifier* is referred to any classifier that is based on fuzzy logic. In fuzzy logic, as opposed to Boolean logic, values can be any real number between 0 and 1 [147]; *Hidden Markov Model* models the system as a Markov process with hidden states. A Markov Process is a stochastic model where the next state depends solely on the current state [148]; *Logistic Regression* is a popular classification model that provides the probability of belonging to each class, while also being easily explainable [149]; *Linear Discriminant Analysis (LDA)* tries to find the hyperplane that minimizes interclass variance while maximizing the distance between classes. This can be used as a classifier, or as a dimensionality reduction tool [150]; *SVM* is a robust classifier characterized by being highly effective on high dimensional data [84]; *k-NN* predicts the class of a data point by looking at the classes of the neighboring points [151]; *MLP* is the most basic form of neural network, composed of an input layer, an output layer, and at least one hidden layer between these two. Each layer is composed of neurons, and the weights of the connections between the neurons of consecutive layers are trained via backpropagation [152]; *CNN* is able to automatically extract spatial

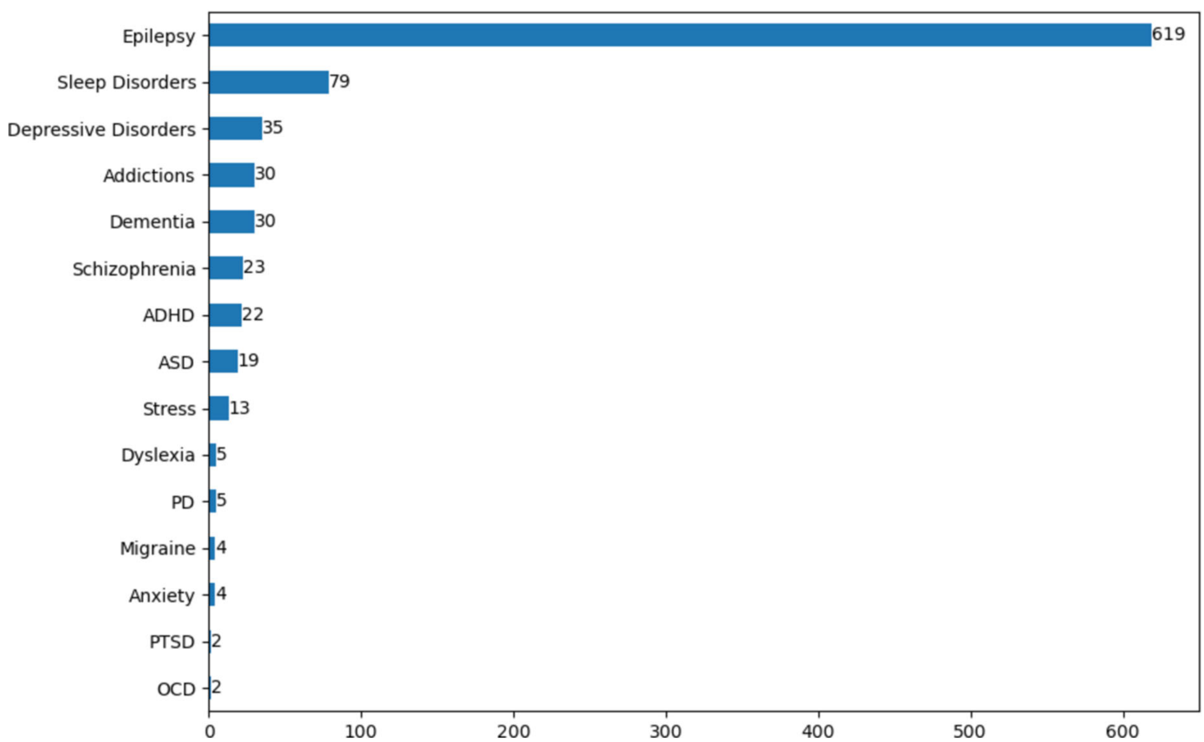


Fig. 7 Frequency of the defined brain disorder categories. Created by authors with *Matplotlib* package from *Python*

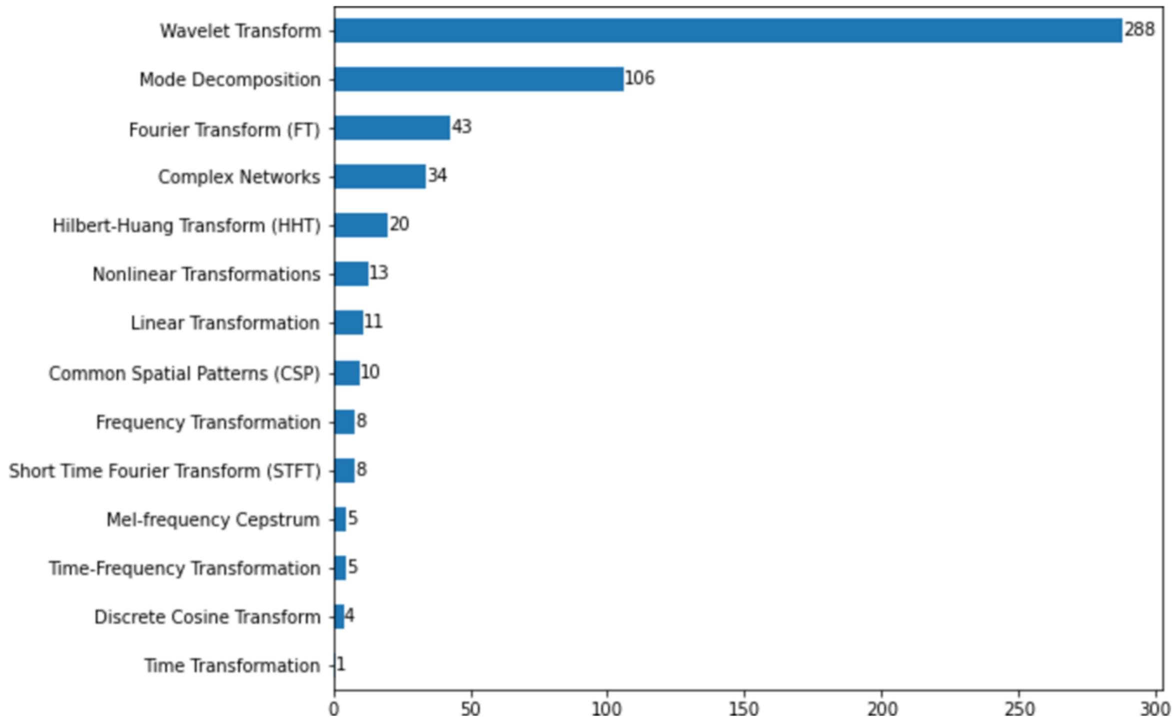


Fig. 8 Frequency of the defined groups for **RQ2.1**. Created by authors with *Matplotlib* package from *Python*

features via convolution operations, making them useful on signal or image data [153]; *RNN* encloses any neural network that has feedback loop connections, making them suitable to learn from sequential data [154]; *Extreme Learning Machine (ELM)* is a feedforward neural network whose nodes are randomly chosen and fixed, and its output weights are obtained analytically in a single step, resulting in a faster training process than in a conventional backpropagation neural network [155].

4.5 Data extraction and mapping of studies

Finally, the last step of the methodology is to extract insights from the obtained classification scheme in order to answer the **RQ**s proposed in Sect. 4.1. First, a frequency analysis will be performed for the categories of each **RQ**. It is important to notice that the sum of the frequencies in the following analysis might not be equal to the totality of the papers, as there are papers labeled with zero or more than one label of the same **RQ**. Additionally, the code used to create Figures 7-10, 12-20 and the supplementary information are available on GtLab, so that the reader can access, verify, and extract new insights from our work.

RQ1. A total of 15 brain disorder categories are present in the collected works (Fig. 7). Epilepsy is by far the most predominant one, representing 68.40% (619 papers) of the total, followed by Sleep Disorders with 8.73% (79 papers). There are less studied categories, with 20-40 papers each, namely Depressive Disorders (3.87%, 35 papers), Addictions (3.31%, 30 papers), Dementia (3.31%, 30 papers), Schizophrenia (2.54%, 23 papers), ADHD (2.10%, 22 papers), ASD (2.10%, 19 papers). Finally, the least studied categories are Stress (1.88%, 13 papers), Dyslexia (5 papers), PD (5 papers), Migraine (4 papers), Anxiety (4 papers), PTSD (2 papers) and OCD (2 papers).

RQ2.1. The frequency of the transformations can be seen in Fig. 8. The WT is the most used transformation by far, being used in roughly 31.82% of the papers (288 papers), followed by Mode Decomposition techniques (11.71%, 106 papers), FT (4.75%, 43 papers), Complex Networks (3.76%, 34 papers), HHT (2.21%, 20 papers), Nonlinear Transformations (1.44%, 13 papers), Linear Transformations (1.22%, 11 papers), CSP (1.10%, 10 papers), Frequency Transformation (0.88%, 8 papers), STFT (0.88%, 8 papers), Mel-frequency Cepstrum (0.55%, 5 papers), Time-frequency Transformation (0.55%, 5 papers), DCT (0.44%, 4 papers) and Time Transformation (0.11%, 1 paper). There are 417 studies (46.1% of the works) where no transformations are studied. It is also interesting to highlight that, in some papers, more than one transformation is applied. Table 4 shows the four most frequent pairs of transformations applied: 22 papers study the use of Mode Decomposition together with WT, 13 study FT together with WT, 6 study the use

Table 4 Most frequent pairs of transformations used in a single paper

Pair	Frequency
Mode Decomposition - Wavelet Transform	22
Fourier Transform - Wavelet Transform	13
Hilbert-Huang Transform - Wavelet Transform	6
Fourier Transform - Mode Decomposition	4

of HHT and WT, and 4 papers study the FT and Mode Decomposition.

RQ2.2. By analyzing the frequency chart for **RQ2.2** (Fig. 9), it can be seen that the most used features are Statistical Parameters (22.54%, 204 papers), Frequency Domain Features (18.01%, 163 papers) and Entropy (18.01%, 163 papers), followed by Time Domain Features (7.85%, 71 papers), Complexity Measures (7.07%, 64 papers), Energy (7.07%, 64 papers), Nonlinear Features (6.41%, 58 papers), Time-Frequency Domain Features (5.64%, 51 papers), Correlation Measures (4.86%, 44 papers), AR (3.31%, 30 papers), Chaotic Features (3.09%, 28 papers), Geometric Features (2.87%, 26 papers), AE Features (2.10%, 19 papers), Tensor Decomposition Features (2.10%, 19 papers) and Graph Features (1.66%, 15 papers). Out of the compiled works, 261 of them (28.8%) do not study any feature extraction-related techniques.

RQ2.3. As for the frequency of feature selection techniques, it can be seen in Fig. 10 that PCA is the most studied one, being present in roughly 5% of the works (49 papers), followed by Genetic Algorithms (38 papers, 4.20%), Statistical Tests (29 papers, 3.20%), Swarm Intelligence (23 papers, 2.54%), Distance Based Feature Selection (15 papers, 16.57%), Clustering (12 papers, 13.26%), Metaheuristics (12 papers, 13.26%) and L1 Regularization (6 papers, 6.63%). A total of 735 studies (81.2% of the works) do not take into account any feature selection techniques.

It is also worth noting that a single study can use techniques corresponding to different sub-questions of the **RQ2**, i.e., it could use the WT, which is a transformation (**RQ2.1**), and a Genetic Algorithm, which is a feature selection technique (**RQ2.3**). It is worth mentioning that the same study may have different labels of the same **RQ**, e.g. if the study uses various feature extraction techniques such as Entropy, Energy and Nonlinear Features, then it will have these three labels associated to the **RQ2.2**. A Venn diagram (Fig. 11) enables us to easily visualize which papers study each sub-question of the **RQ2**: there are 321 papers which focus only on using and studying feature extraction-related techniques, 242 papers which study a combination of transformations and feature extraction, 172 which focus only on transformations, etc. It can also be seen that the total number of papers adds up to 905 works, as expected.

RQ3. Regarding the techniques used in the works to classify brain disorders (Fig. 12), the SVM is the most applied technique, used in 33.04% of the works (299 papers), followed by the MLP (18.56%, 168 papers), k-NN (14.70%, 133 papers), Ensemble models (8.40%, 76 papers), CNN

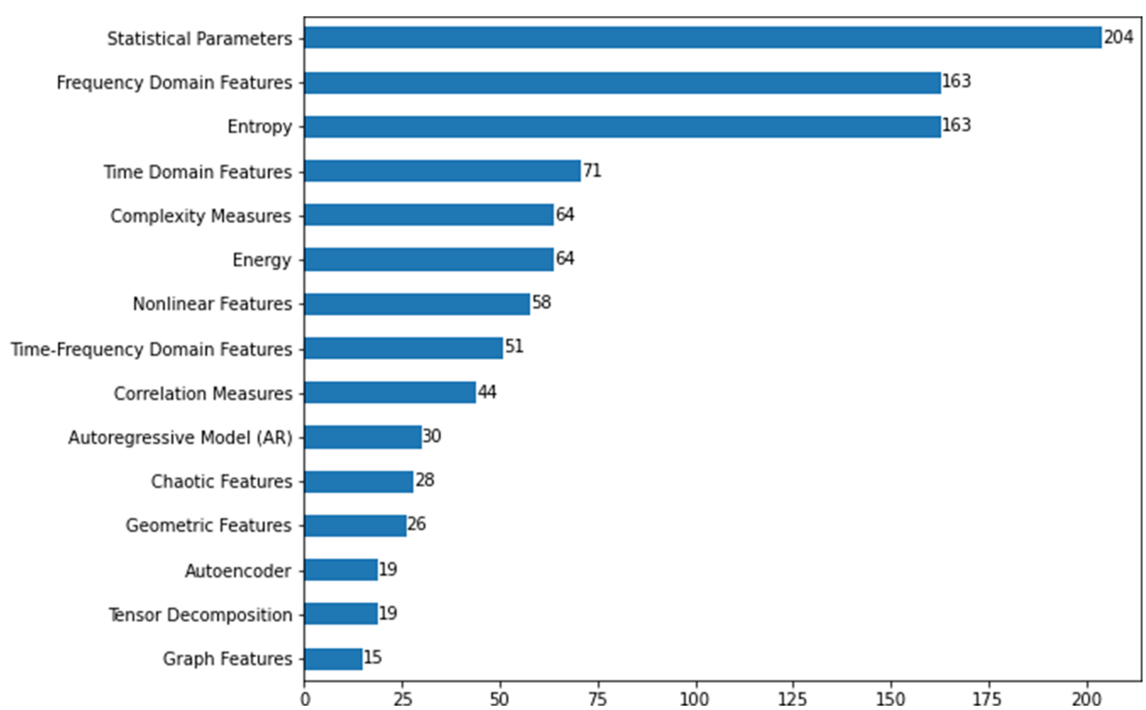


Fig. 9 Frequency of the defined groups for **RQ2.2**. Created by authors with *Matplotlib* package from *Python*

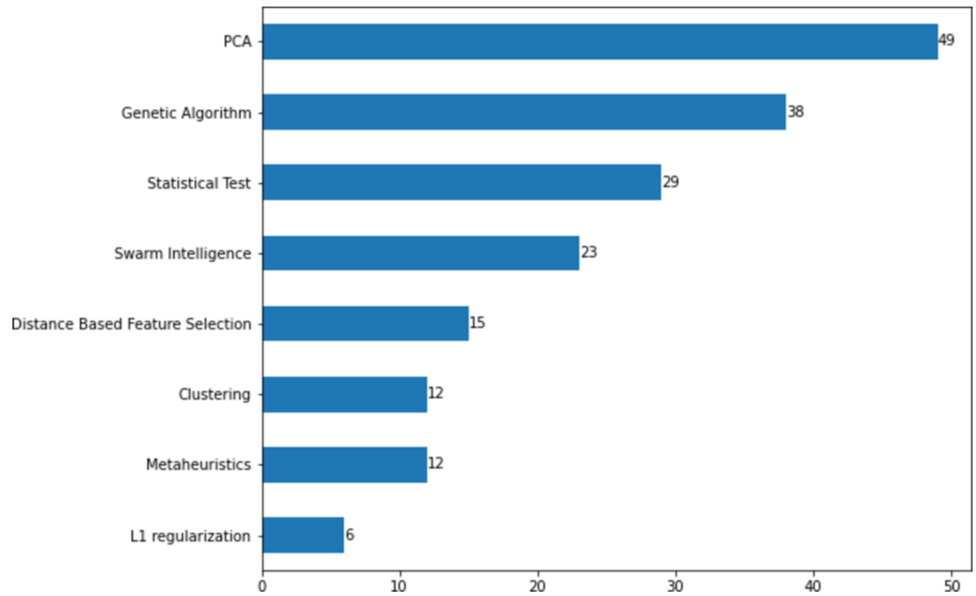


Fig. 10 Frequency of the defined groups for **RQ2.3**. Created by authors with *Matplotlib* package from *Python*

(5.52%, 50 papers), LDA (5.08%, 46 papers), Naive Bayes (3.65%, 33 papers), DT (3.31%, 30 papers), Logistic Regression (2.76%, 25 papers), RNN (2.65%, 24 papers), Fuzzy Classifier (2.21%, 20 papers), Clustering (1.99%, 18 papers), ELM (1.99%, 18 papers) and Hidden Markov Model (0.66%, 6 papers).

It is also interesting to extend the frequency analysis by relating different **RQ**s via bubble maps, as proposed in [87]. This kind of visualization enables us to observe which categories have been already extensively studied, and where the possible research gaps for future directions are. It should be noted that there are two bubble plots missing: mapping

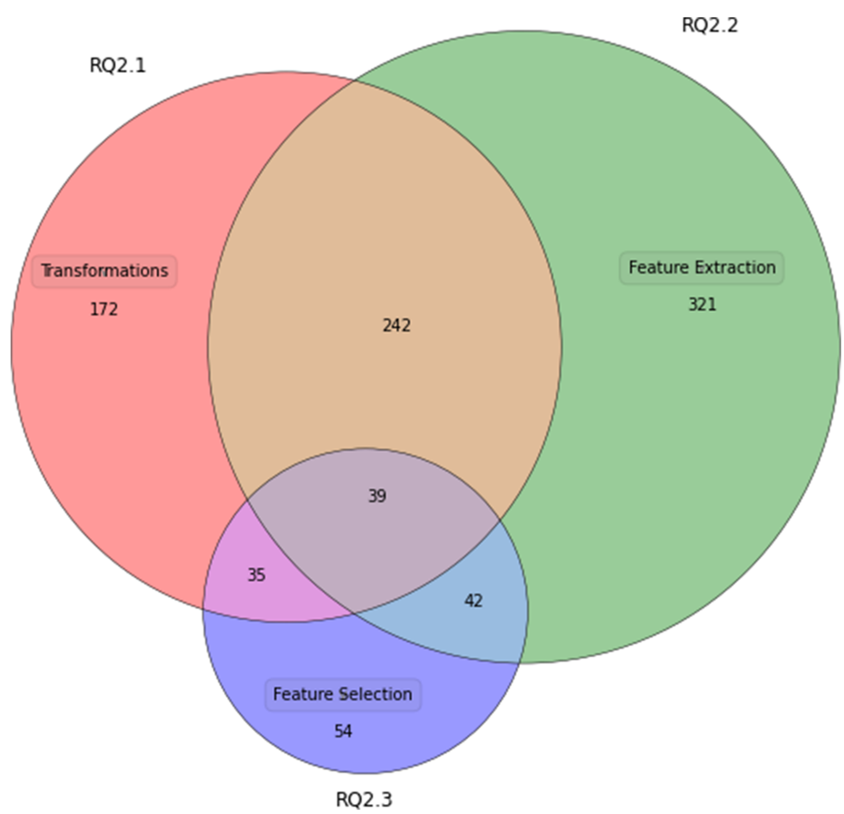


Fig. 11 Venn diagram representing the number of papers that study each subquestion of **RQ2**. Created by authors with *Matplotlib* package from *Python*

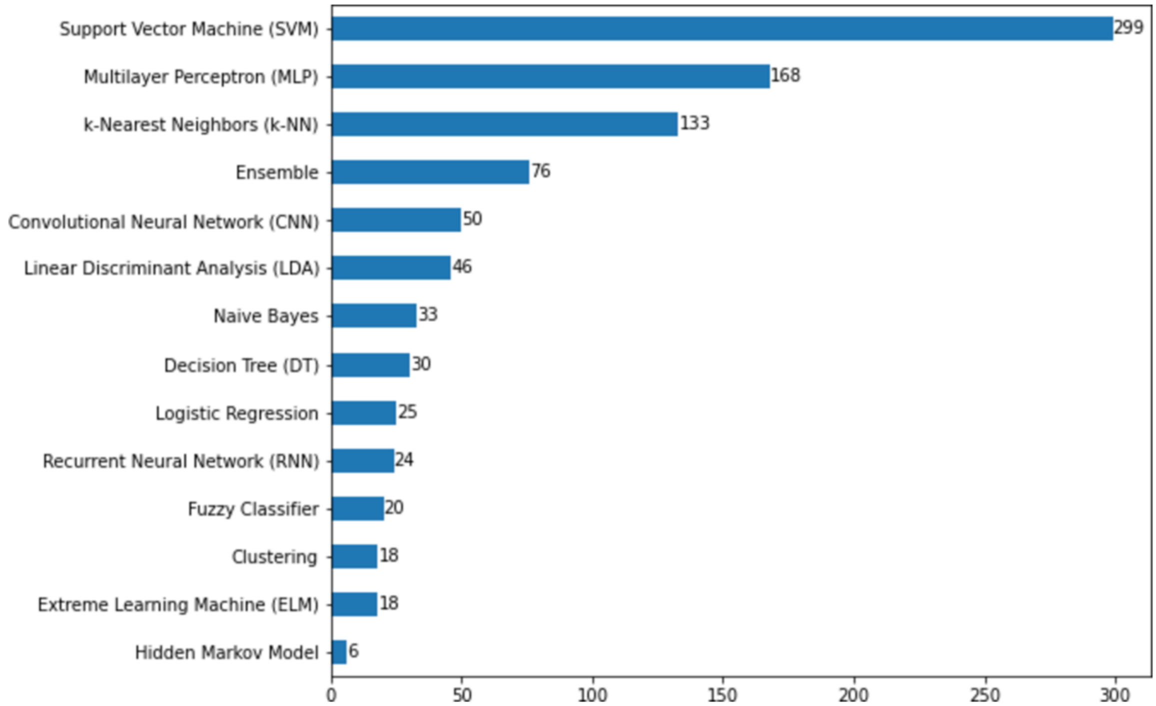


Fig. 12 Frequency of the defined ML and DL categories. Created by authors with *Matplotlib* package from *Python*

Table 5 Three most common combinations in each bubble map (Figs. 13-20)

RQ1 vs. **RQ2.1** (Fig. 13)	Epilepsy - Wavelet Transform (383 papers) Epilepsy - Mode Decomposition (136 papers) Sleep Disorders - Wavelet Transform (56 papers)
RQ1 vs. **RQ2.2** (Fig. 14)	Epilepsy - Statistical Parameters (190 papers) Epilepsy - Entropy (178 papers) Epilepsy - Frequency Domain Features (117 papers)
RQ1 vs. **RQ2.3** (Fig. 15)	Epilepsy - PCA (71 papers) Epilepsy - Genetic Algorithm (41 papers) Addictions - Statistical Test (33 papers)
RQ1 vs. **RQ3** (Fig. 16)	Epilepsy - SVM (310 papers) Epilepsy - MLP (178 papers) Epilepsy - kNN (114 papers)
RQ2.1 vs. **RQ2.2** (Fig. 17)	Wavelet Transform - Statistical Parameters (99 papers) Wavelet Transform - Entropy (98 papers) Wavelet Transform - Energy (42 papers)
RQ2.1 vs. **RQ3** (Fig. 18)	Wavelet Transform - SVM (172 papers) Wavelet Transform - MLP (93 papers) Wavelet Transform - kNN (72 papers)
RQ2.2 vs. **RQ3** (Fig. 19)	Statistical Parameters - SVM (89 papers) Entropy - SVM (82 papers) Frequency Domain Features - SVM (53 papers)
RQ2.3 vs. **RQ3** (Fig. 20)	Statistical test - SVM (28 papers) Genetic Algorithm - SVM (26 papers) PCA - SVM (22 papers)

between **RQ2.1** and **RQ2.3**, and mapping between **RQ2.2** and **RQ2.3**. This is due to the limited information they provide, as there are few articles in **RQ2.3**, in this manner, we reduce the length of this paper as well. In addition, the three most common combinations for each of the pairs of **RQ**s studied are presented in Table 5.

In addition, we consider interesting to focus on brain disorders. We have also created a table (Table 6) showing the most used FE techniques for each disorder.

4.6 Research question 4

During the previous process, a considerable number of related secondary studies were collected. A brief summary of these works is presented in this section as an additional step of the methodology, as it can act as a starting point for researchers that decide to take a new research opportunity. This section will be structured as follows: first, we present secondary studies related to each phase of the **RQ2** (transformation, feature extraction, feature selection), followed by a subdivision of brain disorders.

4.6.1 Signal transform related works

This paper [156] focuses on the applications of sparse representation in brain signal processing. In addition, it deals with Blind Source Separation (BSS), EEG inverse imaging, components extraction, feature selection, and classification. Then, all these techniques are applied to EEG and fMRI data.

[157] presents a comprehensive description of time-frequency and time-scale representations of non-stationary signals such as EEG. They conduct an in-depth review of the principles and design of time-frequency and time-scale methods. Then, time-frequency and time-scale features are presented before comparing the classification efficiency of each one.

[158] compares 19 studies focused on WT and EMD to transform the EEG signal for diagnosing Epilepsy. They gather the type of wavelet and EMD used, the performance obtained, the number of folds used for the cross-validation technique and classification techniques applied for each performed study.

4.6.2 Feature extraction related works

There are other studies which focus on feature extraction methodology. [159] presents FD features and composes a review with more than fifty papers that have applied FD and multi-fractal geometries to extract information from Electrocardiogram (ECG) and EEG signals, brain imaging, mammography and/or bone imaging.

Another paper that focuses on a single technique of feature extraction is [160]. They work with Pattern Recognition techniques to extract information from image features obtained from the transformed EEG signals. In addition, they make a deep description of the most used Pattern Recognition techniques.

4.6.3 Feature selection related works

The work from [161] is only focused on feature selection. They make a review of the most used techniques of feature selection, from 2015 to 2019, in the field of medicine. After all techniques are presented and described, they are applied to different types of data, like medical images (X-rays, CT scans, MRI, retinographies and ultrasound images), biometric signals (EEG, ECG and Electromyography (EMG)) and DNA microarray. Finally, they show an experimental study to compare those described feature selection techniques.

4.6.4 Classification techniques related works

[162] presents a review of the research on the automated diagnosis of 5 neurological disorders in the last 20 years using AI. Those disorders are Epilepsy, PD, AD, Multiple Sclerosis and Ischemic Brain Stroke. The reviewed papers work with physiological signals and images. In the review, they collect the methodology applied or features extracted, the classifier used, and the performance obtained. They make seven summaries of Computer Aided Diagnosis (CAD) systems for: Epilepsy using EEG (112 studies), PD using measurable indicators (9 studies), PD using brain images (8 studies), PD using physiological signals (38 studies), AD using MRI (37 studies), Ischemic Brain Stroke using MRI (23 studies) and Multiple Sclerosis using MRI (8 studies).

[163] elaborates a review of DL techniques used to detect epileptic seizures from intracranial electroencephalography (iEEG) or EEG data from humans or animals. At the beginning of this work, they summarize the most important characteristics of popular and available EEG databases for epileptic seizure detection. Then, they present and describe the most promising DL techniques and list several studies that apply those techniques, gathering the name of the DL-technique used, the number of layers, the type of final classifier, and the accuracy obtained. They reviewed 26 studies focused on 2D-CNN, 24 using 1D-CNN, 15 using RNN, 17 using AEs, 9 using convolutional recurrent neural network (CNN-RNN) and 5 using convolutional autoencoders (CNN-AEs). They also collect 8 studies that have used non-EEG-based data, like sMRI, fMRI and PET scans.

4.6.5 Brain disorders related works

In this subsection we show papers focused on the classification of patients with a mental disorder by means of EEG. Those papers review what other studies have done, gathering

Fig. 13 Mapping between **RQ1** and **RQ2.1**. Created by authors with *Matplotlib* package from *Python*

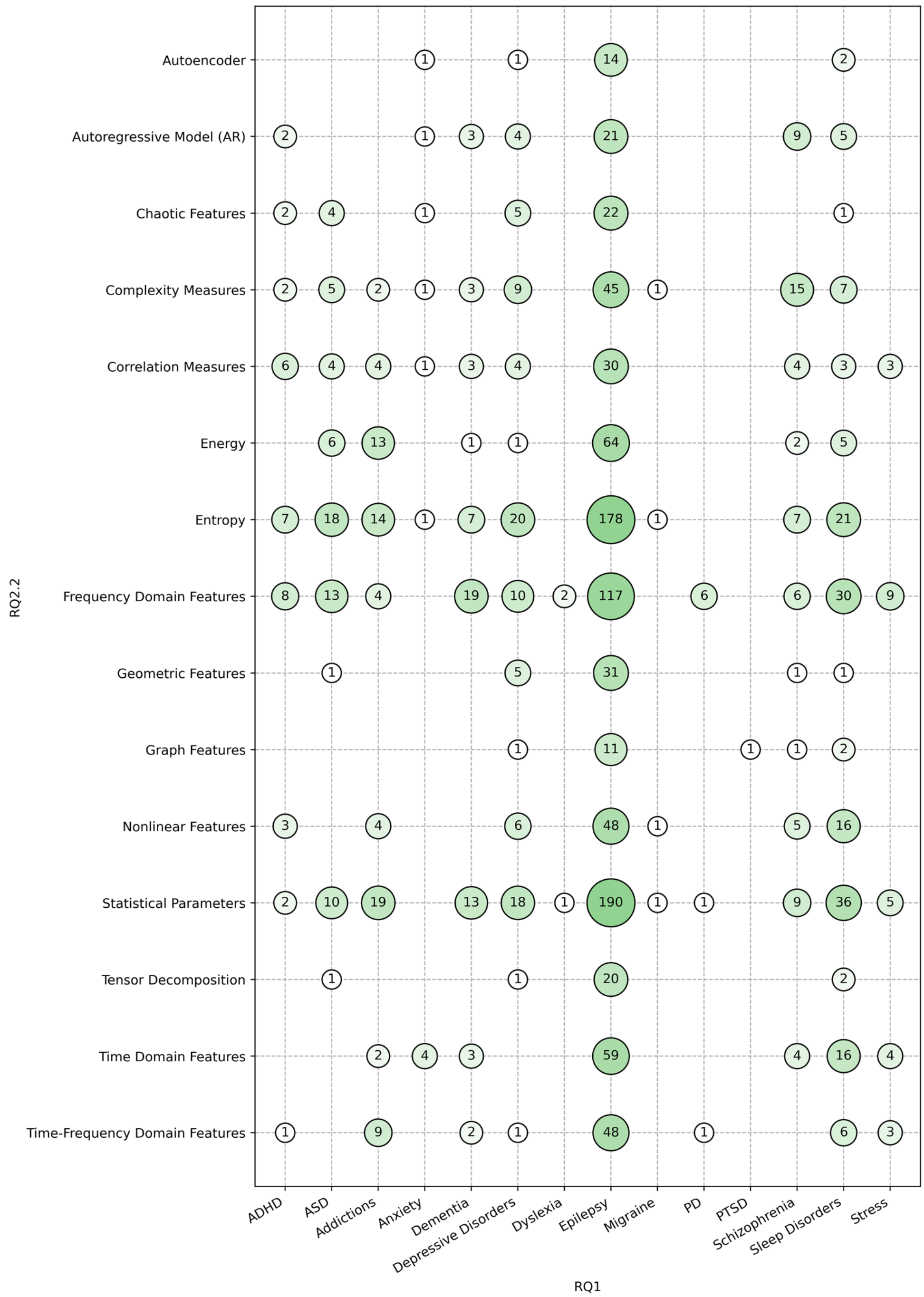
the signal transformations used, the features extracted and selected, and the classification techniques used. Below we present this subsection divided into the disorders that these works deal with:

Mild Cognitive Impairment. [164] (172 studies in the review) extract 234 studies from 172 reviewed works, which are related with AD and MCI. From each study, they collect the dataset used, the time of prediction, the data type (EEG, MEG or fMRI), the classification algorithm, the number of folds used in the cross-validation technique and some performance metrics like AUC, accuracy, sensitivity and specificity.

[165] composes a systematic review with 82 studies that are focused on Dementia. They present, in a visual manner, the distribution of papers according to the sampling frequency applied, number of study subjects, number of electrodes used, recording time, tools used to process the signal and classification techniques.

[166] uses brain imaging techniques applied to EEG, MEG, MRI and fMRI data to diagnose AD. They elaborate

Fig. 14 Mapping between **RQ1** and **RQ2.2**. Created by authors with *Matplotlib* package from *Python*



an interesting table with the advantages and disadvantages of some classification and artifacts removing algorithms.

Epilepsy. [167] makes an extended review of the most used signal transformations and feature extraction techniques to detect epileptic seizures and diagnose Epilepsy. In addition, two tables are elaborated with the most relevant data of the studies reviewed: name of features extracted, name of the classifier used and accuracy obtained. The first table contains 21 previous works for the automated detection of normal and epileptic classes. The second table has 17 previous works for the automated detection of normal, interictal and epileptic classes.

[168] describes briefly the most used features in the literature of EEG seizure detection dividing them into time domain, frequency domain and time-frequency domain. Previously, they had summarized 55 studies by writing down in a table the type of features extracted, the transformation signal method and the performance obtained according to

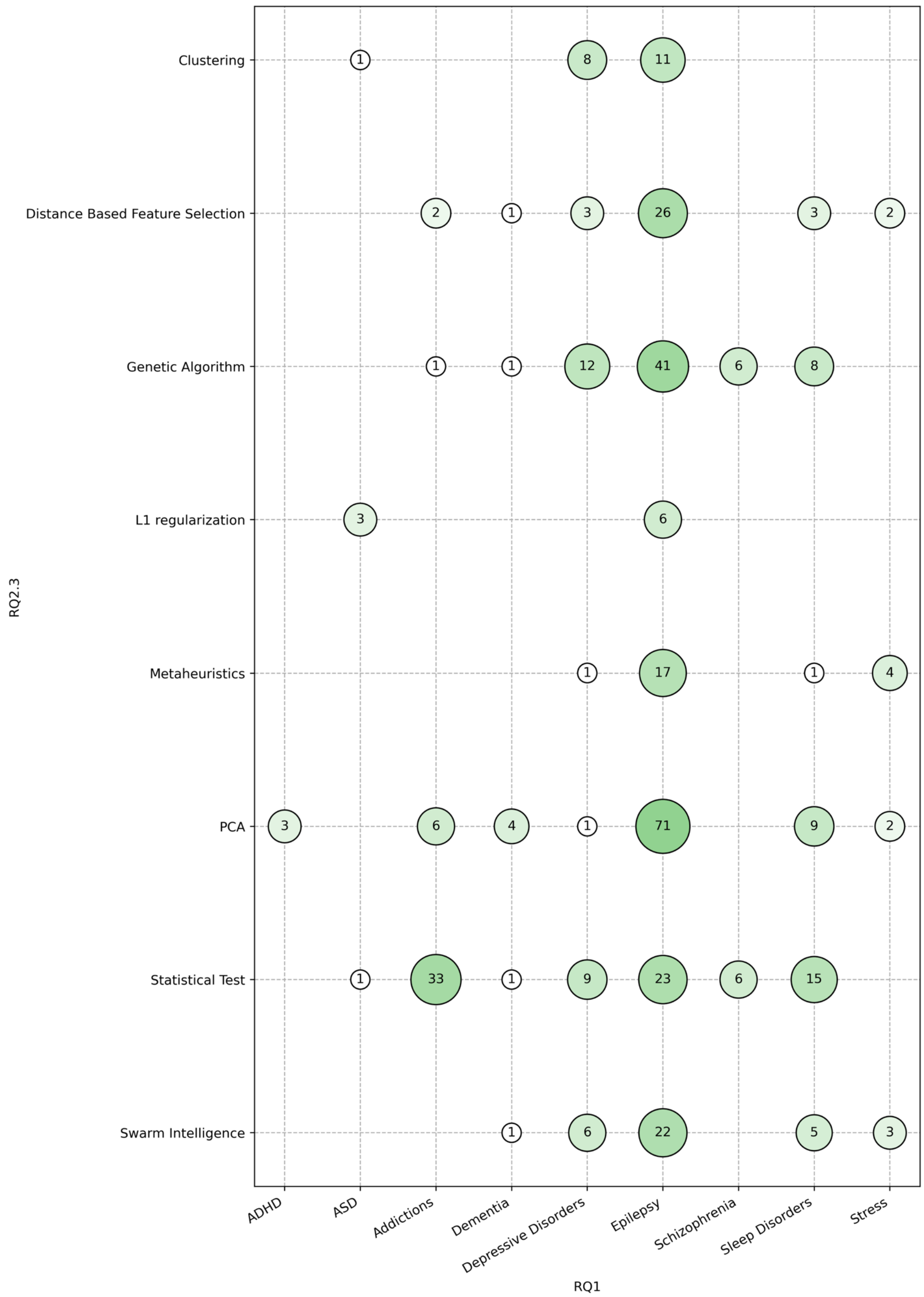


Fig. 15 Mapping between **RQ1** and **RQ2.3**. Created by authors with *Matplotlib* package from *Python*

the database used (CHB-MIT scalp EEG database and Bonn University database). Eventually, an experiment is performed using the best results of reviewed works.

[169] reviews 87 studies made between 2010 and 2020. This work gathers 58 studies using conventional feature extraction techniques together with ML classifiers. They also review 29 studies that use DL techniques, without hand-crafted feature extraction.

[170] does a review divided in two, depending on the type of subjects taken to make the experiment. They take 36 studies that have worked with human subjects and 5 studies that have worked with animal subjects. This work collects information about the goal of the study, the database and the methodology used, the time of prediction and the performance obtained.

[171] offers a comprehensive review of signal processing techniques like preprocessing, feature extraction, feature selection and classification schemes for Epileptic Seizure Prediction (ESP). This work has a summary with recent ESP surveys, from 2016 to 2021 and it collects other works that

Fig. 16 Mapping between **RQ1** and **RQ3**. Created by authors with *Matplotlib* package from *Python*

use the feature selection methods exposed in the review. In addition, the manuscript includes some tables with detailed information about the architecture of artificial neural networks that other studies have used. Moreover, it offers interesting sections with trends and emerging classification techniques, as well as another section focused on the limitations and challenges of ESP.

Sleep disorders. Regarding sleep disorders, [172] focuses on sleep stages and presents several tables that contain information about many previous studies. Those tables collect the feature extraction and feature selection techniques, the classification method, the sleep stages and sleep disorders classified, the number of subjects, the database used, the channels chosen, and the performance obtained for each study. It also carries out a practical experiment by comparing several classifiers.

[173] collects from each study reviewed, the features extracted, the preictal time, the database used, the year the

Fig. 17 Mapping between **RQ2.1** and **RQ2.2**. Created by authors with *Matplotlib* package from *Python*

study was conducted, the number of patients, the recording type (iEEG or sclap), the sensitivity, the false prediction rate obtained and the statistical validation used.

4.6.6 Other related works

[174] makes a brief summary of 11 studies focused on diagnosing MDD, taking into account the applied methodology, the extracted features, the classification method used and the accuracy obtained from each one.

[175] does a brief review of studies that investigate how to classify Dyslexia. We could find the test group size and control group size, the age range of the subjects, the number, and the name of picked channels and the signal transformation used in each study reviewed.

Fig. 18 Mapping between **RQ2.1** and **RQ3**. Created by authors with *Matplotlib* package from *Python*

5 Discussion

Epilepsy is by far the most studied neurological disorder by means of EEG (Fig. 7). In general, every FE technique taken into account in this work has been applied to the study of Epilepsy (Figs. 13, 14, 15) as well as every ML algorithm (Fig. 16). Compared to Epilepsy, the other brain disorders have been considerably understudied by means of EEG, so we strongly recommend considering them as potential research options. In our opinion, Epilepsy might be the most studied because (i) it affects a considerably large portion of the population, and quickly predicting a seizure could greatly prevent injuries (ii) a seizure is characterized by abnormal excessive activity in some regions of the brain [101], thus it might be easier to predict such periods of high activity than predicting other patterns related to other brain disorders such as ADHD, that could be more difficult to differentiate from the patterns of a neurotypical individual.

Regarding Epilepsy, even though it has been extensively studied, more work could be done applying less-studied FE

Fig. 19 Mapping between **RQ2.2** and **RQ3**. Created by authors with *Matplotlib* package from *Python*

techniques. For example, (i) by applying any other transformation besides the WT or Mode Decomposition, (ii) using less-studied features such as chaotic, complexity, tensor or graph-related features, and (iii) focusing more on feature selection techniques in general. However, we should be especially careful when choosing new research gaps in the bubble plots, as there may be combinations that are not studied simply because it does not make sense to combine them.

The WT is the most used because of its effectiveness in nonstationary signal analysis. The ability to perform time-frequency decomposition into Wavelet components facilitates the analysis of certain localized patterns that can occur in EEGs. This justifies its use over other transforms such as the FT, or the STFT. Despite its popularity, we can see in Fig. 13 that it has mostly been studied in relation to Epilepsy, so there are research gaps regarding other brain disorders, as it is not trivial to directly apply the same Wavelet analysis in every case. For example, the Daubechies-4 Wavelet is commonly used to detect epileptic spikes [176], but it may not be suitable to study another brain disorder. It is also com-

Fig. 20 Mapping between **RQ2.3** and **RQ3**. Created by authors with *Matplotlib* package from *Python*

mon to apply several transformations together. For example, as presented in Table 4, there are 22 works where a Mode Decomposition technique is used to decompose the signal into different meaningful components, and then the WT is used to extract useful features in the time-frequency domain.

The most common features used in the models are Statistical Parameters, such as the mean or standard deviation of the amplitude or frequency. Features related to the Entropy, Energy, complexity or chaos are also frequently used because of the good results produced. In general, it seems to be a trend for trying to quantify the degree of disorder of certain regions of the brain. Moreover, features related to correlation and graph connectivity are used to assess how different parts of the brain are interacting.

The SVM is a versatile classifier that is highly effective in high dimensional spaces, i.e., a high number of different features. Typically, the application of FE techniques such as WT or FT to EEG data yields a large number of features, justifying why it might be the main choice in almost every brain disorder (Figs. 12, 16). However, we found that there

Table 6 Most common techniques used with brain disorders. The number in brackets indicates how many papers this technique appears in

Most common techniques				
Brain disorder	Transformation	Feature extracted	Feature selection	Classification model
ADHD	WT (8)	Frequency Domain Features (8)	PCA (3)	SVM (7)
ASD	WT (22)	Entropy (18)	L1 Regularization (3)	k-NN (23)
Addictions	WT (25)	Statistical Parameters (19)	Statistical Test (33)	SVM (24)
Anxiety	(0)	Time Domain Features (4)	(0)	MLP (1) Logistic Regression (1) Ensemble (1)
Dementia	WT(12)	Frequency Domain Features (19)	PCA (4)	SVM (17)
Depressive Disorders	WT(14)	Entropy (20)	Genetic Algorithm (12)	SVM (23)
Dyslexia	WT(3)	Frequency Domain Features (2)	(0)	SVM (1)
Epilepsy	WT(383)	Statistical Parameters (190)	PCA (71)	SVM (310)
Migraine	WT (3)	Complexity Measures (1) Entropy(1) Nonlinear Features (1) Statistical Parameters (1)	(0)	Ensemble (2)
OCD	FT (3)	(0)	(0)	MLP (2)
PD	Time-Frequency Transformation (1) Mode Decomposition(1) Mel-frequency Cepstrum (1)	Frequency Domain Features (6)	(0)	SVM (2)
PTSD	Complex Networks (2)	Graph Features (2)	(0)	(0)
Schizophrenia	WT (10)	Complexity Measures (15)	Statistical Test (6) Genetic Algorithm (6)	SVM (16)
Sleep Disorders	WT (56)	Statistical Parameters (36)	Statistical Test (15)	SVM (49)
Stress	WT (5)	Frequency Domain Features (9)	Metaheuristics (4)	SVM (14)

is an ascending trend of using DL in the latest years, which might be justified by the following advantages over ML:

- Even though classical ML models can perform classification on temporal data via techniques such as moving window, using Hidden Markov Models or the use of FT and WT, they may not be able to fully process the sequential or temporal nature of EEG data, as most ML algorithms have no notion of order within their input features. On the other hand, there are DL models such as RNN, LSTM, or CNN that could exploit such sequential information [177].
- DL models are able to automatically learn optimal representations of the input data [70], reducing the burden of manually engineering some of the input features. Despite this, FE is still important to DL, as carefully applying the correct set of FE techniques can considerably reduce the complexity of the model.

Thus, we suggest experimenting with different DL models while also taking into account FE techniques, in order to reduce its complexity. Specifically, we think that there is a lot to be done regarding feature selection techniques applied to DL. As we can see in Fig. 11, there are only 170 out of 905 studies that use any feature selection techniques. Moreover, Fig. 20 shows that only a small subset of the papers use feature selection techniques for DL models.

During the keywording process, we also tried to collect data related to the evaluation of the ML pipelines in order to present insights about the techniques and models that perform better for each brain disorder. However, it was difficult to make an objective comparison, mainly because of (i) differences in the evaluation process across the works and (ii) the lack of a public dataset of reference for some brain disorders. Regarding the evaluation process, we found that, in some works, inter-subject evaluation was performed. In other words, the data of the same subject was used for training and for evaluation of the algorithm, resulting in an artificially higher performance than the real one. Moreover, to be able to compare studies, it is essential to perform the evaluation in the same dataset. To the best of our knowledge, there are public datasets for only 6 out of the 15 brain disorders taken into consideration in this work (epilepsy, sleep disorder, depression, schizophrenia, ADHD and stress). Therefore, it is of crucial importance to create quality public datasets in order to correctly perform comparisons among works. Moreover, creating large enough datasets would greatly contribute to the use of DL architectures, as their performance typically increase with the amount of training data. In summary, we strongly suggest the creation of quality public datasets as a future research opportunity that will be extremely useful for the community.

It is also interesting to compare the frequency of studies of each brain disorder against its prevalence. [3] performed an estimation of the number of individuals in Europe suffering from any of the most 12-month prevalent brain disorders, which were the following: Anxiety Disorders (61.5m), Unipolar Depression (30.3m), Insomnia (29.1m), Somatoform Disorders (20.4m), Alcohol Dependence (14.6m), Sleep Apnea (12.5m), PTSD (7.7m), Dementia (6.3m), ADHD (3.3m) and Epilepsy (2.64m). By comparing the previous list with the frequency of the studies, it can clearly be seen that there is a lot of research to be done related to FE of EEG for the most prevalent brain disorders.

6 Conclusions and future works

In this section we present our findings and conclusions, then we discuss the limitations of this work and finally, we suggest future works based on our results.

6.1 Conclusions

The main goal of this SMS has been to provide a clear overview of the research area of FE of EEG applied to mental disorders. Four **RQ**s were proposed in order to guide such an overview. A total of 6133 initial works were retrieved by searching in five well-known different databases (Scopus, Web of Science, IEEEXplore, ScienceDirect, ACM) through the synthesis of adequate queries. After applying the defined set of exclusion and inclusion criteria, the final set of relevant works resulted in 905 papers, with labels corresponding to each **RQ**. Finally, the labels were used to extract insights.

The results corresponding to each **RQ** are as follows:

- A total of 15 brain disorder groups have been studied by using FE of EEG signals, being Epilepsy the most predominant by far, present in 68.40% of the studied papers.
- 14 transformation groups, being WT the most applied in 31.82% of the papers.
- 15 groups related to feature extraction, with Statistical Parameters being the most used in 22.54% of the works.
- 8 feature selection techniques, PCA being the most used in 5% of the papers.
- 14 ML and DL techniques, SVM being the most used, being present in 33.04% of the studies.

6.2 Limitations

As limitations of this work, we would like to remark that it is focused on FE techniques and it has ignored the preprocessing stage, i.e. eliminating noise, applying filters on the signal, removing artifacts such as blinking, heart rate, etc. Even though preprocessing is crucial, it is out of the scope of this paper as the main goal of the SMS is to provide an overview of the current FE techniques used on EEG data for the diagnosis of mental disorders, and how it is combined with different ML algorithms. Preprocessing should therefore be the goal of another SMS itself. It should be noted that we do not extract the accuracy achieved, nor the contributions and drawbacks of each study. This is mainly because we have only read the abstract of each paper, and we were seldom able to clearly find the contributions and drawbacks. There is also a reason not to collect the accuracy of each paper and it is because we consider it unfair to compare performances among studies carried out with different databases. In addition, we were not able to gather the databases used by reading only the abstract. Moreover, this SMS did not collect information about the year and country of publication, which could have been useful to visualize trends, because most of the papers lacked this data during its fetching and it was an unfeasible task to perform at hand.

6.3 Future work

The analysis was extended further by combining different **RQ**s via bubble plots, which provided a clear visual overview of the research area, useful for researchers and practitioners in deciding on future research options. After carefully analyzing the previous results, insights and recommendations were proposed in the discussion section:

- It is worth noting that there an are plenty of research gaps concerning an all brain disorders except Epilepsy.
- As we can see in Fig. 11, feature an selection is by far the less studied technique and its an application is crucial to reduce the complexity of ML models as well as to provide explainability to DL models.
- The SVM is the most used ML model in an almost every brain disorder since it is efficient when working with an high-dimensional spaces, but we recommend focussing on experimenting with models that can take advantage of the sequential nature of EEG data, such as RNN or CNN-based architectures.

We also discussed the difficulty of an objectively comparing ML pipelines, as the evaluation process differed among works. We suggest an standardizing the evaluation process an by selecting a standard set of metrics, using an the same dataset and ensuring that there is no data an of the same subject in both the training and test set, as this artificially an improves the metrics, thus giving a sense of fake an performance. This also shows the importance of having quality datasets available to the community, an and we strongly recommend the synthesis of quality datasets of EEG data for any mental disorder as a future research direction.

By analyzing the most prevalent brain disorders in Europe, it was discovered that little work, if any, was done in terms of FE of EEGs related to these disorders. Therefore, we would encourage others to keep on researching and advancing on these topics, namely Anxiety Disorders, Unipolar Depression, Insomnia, Somatoform Disorders, Alcohol Dependence, Sleep Apnea, PTSD, Dementia, and ADHD.

Given the current trends, we expect the use of DL architectures to predominate over traditional ML for the automated diagnosis of mental disorders in the near future, as they have more predictive capabilities and benefit from large datasets. As DL models require large amounts of data to be trained, collecting large-scale quality datasets of EEG recordings will be even more important. We also expect a rise in the use of techniques such as transfer learning [178], or domain adaptation [179] to leverage the performance of DL models. Additionally, we also expect future work on XAI techniques applied to EEG, as explainability is crucial for the application of DL into clinical diagnosis tools. FE will have a critical role in both aspects: FE will be required to fully take advantage of the future large-scale datasets, whereas FE could merge with XAI techniques to explain DL models in terms of features that are understandable to the domain expert.

Acknowledgements This work has been co-funded by the BALLADEER (PROMETEO/2021/088) project, a Big Data analytical platform for the diagnosis and treatment of Attention Deficit Hyperactivity Disorder (ADHD) featuring extended reality, funded by the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital (Generalitat Valenciana) and the AETHER-UA project (PID2020-112540RB-C43), a smart data holistic approach for context-aware data analytics: smarter machine learning for business modelling and analytics, funded by Spanish Ministry of Science and Innovation. Sandra García-Ponsoda holds a predoctoral contract granted by *ValgrAI - Valencian Graduate School and Research Network of Artificial Intelligence* and the *Generalitat Valenciana*, and co-funded by the *European Union*. Jorge García-Carrasco holds a predoctoral contract (CIACIF/2021/454) granted by the *Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital (Generalitat Valenciana)*.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Declarations

Competing interests The authors declare that they have no conflict of interest.

Appendix A Labels and categories

In this section we present the labels extracted during the keywording process. As the final number of labels was extremely large, they were grouped into different, more general, categories. The obtained labels and its respective category are presented in the tables below. On the fist column of the table there are the labels used to classify the read abstract of every selected paper. On the second column there are the general labels that we used to decrease the number of labels used to perform this work. Table 7 contains the labels and categories corresponding to **RQ1**, Tables 8-9 to **RQ2.1**, Tables 10-13 to **RQ2.2**, Tables 14-15 to **RQ2.3** and Table 16 to **RQ3**.

Table 7 Labels and categories for **RQ1**

Research Question 1	
Label	Category
ADHD	ADHD
Mental tasks	
Concentration	
Alcohol Use Disorder (AUD)	Addictions
Drug-Related Addiction	
Gaming Addiction	
Anxiety	Anxiety
Autism Spectrum Disorder (ASD)	Autism Spectrum Disorder (ASD)
Alzheimer's Disease (AD)	Dementia
Dementia	
Mild Cognitive Impairment (MCI)	
Depression	Depressive Disorders
Major Depressive Disorder (MDD)	
Bipolar Disorder	
Dyslexia	Dyslexia
Epilepsy	Epilepsy
Seizure	
Migraine	Migraine
Obsessive-Compulsive Disorder (OCD)	Obsessive-Compulsive Disorder (OCD)
Parkinson's Disease (PD)	Parkinson's Disease (PD)
Post-Traumatic Stress Disorder (PTSD)	Post-Traumatic Stress Disorder (PTSD)
Schizophrenia	Schizophrenia
Sleep disorder	Sleep disorders
Apnea	
Insomnia	
Sleep Arousal	
Stress	Stress

Table 8 Labels and categories for **RQ2.1**

Research Question 2.1	
Label	Category
GARCH modeling of Wavelet coefficients	Wavelet Transform
Wavelet Transform	
Flexible Analytical Wavelets Transform (FAWT)	
Symplectic Geometric Decomposition	Mode Decomposition
Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN)	
Modified Empirical Mode Decomposition (EMD)	
ICA	
Intrinsic Time-scale Decomposition (ITD)	
Variational Mode Decomposition	
Intrinsic Mode Functions (IMFs)	
Hermite Transform	
Empirical Mode Decomposition (EMD)	
Jacobi Polynomial Transform	
Nonlinear Mode Decomposition	
Local Mean Decomposition (LMD)	
Hilbert-Huang Transform (HHT)	Hilbert-Huang Transform (HHT)
Hilbert Vibration Decomposition	
Nonlinear Transformations	Nonlinear Transformations
Short Time Fourier Transform (STFT)	Short Time Fourier Transform (STFT)
Fourier Transform	Fourier Transform (FT)
Fourier	
Fast Fourier Transform (FFT)	
Thompson Multi Taper Method	Frequency Transformation
Walsh Hadamard Transform	
Prony Method	
Periodogram	
Hadamard Transform	

Table 8 continued

Research Question 2.1	
Label	Category
Weighted Complex Network	Complex Networks
Functional Brain Network	
Large-scale Brain Network	
Functional Networks	
Complex Networks	

Table 9 Labels and categories for **RQ2.1**

Research Question 2.1	
Label	Category
Mel-Frequency Cepstrum	Mel-frequency Cepstrum
Linear Predictive Cepstral Coefficients	
Cepstral Coefficients	
Graph-regularized Non-negative Matrix Factorization (GNMF)	Tensor Decomposition
Wishart Distributed Matrices	
Hermite Decomposition	
Tensor Decomposition	
Tensor Factorization	
Singular Value Decomposition	
Non-Negative Matrix Factorization (NNMF)	
Matrix Determinant	
Common Spatial Patterns (CSP)	Common Spatial Patterns (CSP)
Linear Transformation	Linear Transformation
Choi-Williams Distribution (CWD)	Time-Frequency Transformation
Synchrosqueezing Transform	
Smoothed Pseudo Wigner-Ville Distribution	
Matching Pursuit	Time Transformation
Discrete Cosine Transform	Discrete Cosine Transform
Discrete Sine Transform	

Table 10 Labels and categories for **RQ2.2**

Research Question 2.2	
Label	Category
Fast-Slow-Wave Energy Ratio (FSR)	Energy
Smoothed non-Linear Energy Operator	
Frequency-Specific Energy Levels	
Energy	
Multivariate Multi-Scale Sample entropy (MMSE)	Entropy
Fuzzy Entropy (fuzzyEn)	
Entropy	
Refined Composite Multi-Scale Dispersion Entropy (RCMDE)	
Statistical Parameters	Statistical Parameters
Harmonic Parameters	
Brain Symmetry Index	
Statistical Distributions	
Standard Deviation (STD)	
Hjorth Parameter	
Hurst Exponent (HE)	
Asymmetry	
Variance Inflation Factor (VIF)	
Higher Order Statistics	
Volatility Index	
Delay Vector Variance	
Fluctuation Index	
Hjorth's Mobility	
Coefficient of Variation	
Variance	
Successive Decomposition Index	
Synchronization Methods	
Convex Optimized Delay Vector Variance (CDVV)	
Variability Coefficient	
Autoencoder	Autoencoder
Feature Encoding with Autoencoder	
non-Linear Features	Nonlinear Features
Time-Spectral Domain Features	Time-Frequency Domain Features
Time-Frequency Domain Features	

Table 11 Labels and categories for **RQ2.2**

Research Question 2.2	
Label	Category
Spatial Pattern of Network	Geometric Features
Quad Binary Pattern (QBP)	
One-Dimensional Local Binary Pattern	
Local Neighbor Gradient Pattern (LNGP)	
Neighborhood Gradient Pattern	
Geometric Features	
Local Binary Pattern	
Relaxed Local Neighbor Difference Pattern (RLNDiP)	
Volterra Features	
Haralick Features	
Symmetrically Weighted Local Neighbor Gradient Pattern (SWLNGP)	
One-Dimensional Ternary Patterns	
Quadruple Symmetric Pattern (QSP)	
Power Frequency	Frequency Domain Features
Power Spectrum	
Alpha Peak Frequency (iAPF)	
Mean Amplitude	
Frequencies	
Individualized Alpha Absolute Power (iABP)	
Higher-Order Spectra (HOS)	
Power	
Multi-Scale Spectral Features (MSSFs)	
Frequency Relive Power (RP)	
Spectral Domain Features	
Average Frequency	
Power Spectra	
Power Spectral Density	
Band Power	
Auto-Correlation Frequencies	
Frequency Domain Features	
Power Band	
Power Levels of Power Spectral Density	

Table 12 Labels and categories for **RQ2.2**

Research Question 2.2	
Label	Category
Time Domain Features	Time Domain Features
Relative Amplitude	
Local Maxima and Minima	
RMS Amplitude	
Correlation Analysis	Correlation Measures
Phase Amplitude Coupling	
Cross-Correlation	
Dissimilarity Measure	
Phase Synchronization	
Coherence	
Correlation Dimension (CD)	
Cross-Spectral Coherence	
Correlation between Channels	
Phase Lagging Index	
Cross-Frequency Coupling	
Centered Correntropy (CC)	
Itakura Distance	
Partial Directed Coherence (PDC)	
Correlation	
Synchronization Likelihood	
Mean Phase Coherence	
Phase Correlation	
Covariance Matrix	
Mutual Information Evaluation Function	
Nonlinear Interdependence (NI)	
Phase Locking Value	
Kolmogorov Complexity (KC)	Complexity Measures
Lempel-Ziv Complexity (LZC)	
Fractal Dimension (FD)	
Local Fractal Spectrum	
Capacity Dimension	
Line Length	
Epsilon-Complexity	
Fractal Analysis	

Table 13 Labels and categories for **RQ2.2**

Research Question 2.2	
Label	Category
Phase Space Features	Chaotic Features
Phase Space Representation	
Lyapunov Exponents	
Recurrence Quantification Analysis	
Chaotic Features	
Reconstructed Phase Space	
Phase Space Reconstruction (PSR)	
Detrended Fluctuation Analysis	
Autoregressive Model Coefficients	Autoregressive Model (AR)
Absolute Value of the Highest Slope of Autoregressive Coefficients (AVLSAC)	
Autoregressive (AR) Parameters	
Horizontal Visibility Graph (HVG)	Graph Features
Graph Features	
Topographic Features	
Direct Transfer Function	
Degree Centrality	
Local Graph Structure (LGS)	
Graph Theory	

Table 14 Labels and categories for **RQ2.3**

Research Question 2.3	
Label	Category
PCA	PCA
Functional PCA	
Genetic Programming	Genetic Algorithm
Immune Clonal Algorithm	
Genetic Algorithm	
Differential Evolution	
Binary Differential Evolution	
Fisher Score	Statistical Test
Gini Impurity Score	
Pearson Correlation Matrix	
T-test	
Kruskal-Wallis Test	
Pearson Coefficient	
Shapley Value	
F-Score	
Hypothesis Tests	
Fast Correlation-Based Feature Selection	
Improved Correlation-based Feature Selection (ICFS)	
Mann-Whitney Test	
Lambda of Wilks Criterion	
Water Cycle Algorithm	Metaheuristics
Harmony Search	
Maximum Relevance Minimum Redundancy	
Fuzzy Based Cuckoo Search	
ReliefF	Distance Based Feature Selection
Conditional Mutual Information Maximization	
Neighborhood Component Analysis	
Mahalanobis Distance	
Binary Bat Algorithm	Swarm Intelligence
Artificial Bee Colony	
Hunting Search	
Firefly Algorithm	
Particle Swarm Optimization	
Grey Wolf Optimization	
Ant Colony Optimization (ACO)	

Table 15 Labels and categories for **RQ2.3**

Research Question 2.3	
Label	Category
Clustering	Clustering
Fuzzy Clustering	
Hierarchical Clustering	
Gaussian Mixture Model	
Multi-Cluster Feature Selection	
L1 Regularization	L1 Regularization
LASSO	
Adaptive LASSO	

Table 16 Labels and categories for **RQ3**

Research Question 3	
Label	Category
Ensemble	Ensemble
Neural Network (NN)	Multilayer Perceptron (MLP)
Extreme Learning Machine (ELM)	Extreme Learning Machine (ELM)
Support Vector Machine (SVM)	Support Vector Machine (SVM)
Decision Tree (DT)	Decision Tree (DT)
Linear Discriminant Analysis (LDA)	Linear Discriminant Analysis (LDA)
Clustering	Clustering
knn	KNN
Bayesian Classifier	Naive Bayes
Logistic Regression	Logistic Regression
CNN	Convolutional Neural Network (CNN)
Fuzzy Classifier	Fuzzy Classifier
Recurrent Neural Network (RNN)	Recurrent Neural Network (RNN)
Hidden Markov Model	Hidden Markov Model

References

1. Organization WH, et al (2022) Mental health and COVID-19: early evidence of the pandemic's impact: scientific brief, 2 March 2022. In: Mental health and COVID-19: early evidence of the pandemic's impact: scientific brief, 2 March 2022. World Health Organization 93:85–117
2. Association American Psychiatric et al (2013) Diagnostic and statistical manual of mental disorders: DSM-5, vol 5. American psychiatric association Washington, DC
3. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jönsson B, Olesen J, Allgulander C, Alonso J, Faravelli C et al (2011) The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol 21(9):655–679
4. Uchida A, Pillai JA, Bermel R, Jones SE, Fernandez H, Leverenz JB, Srivastava SK, Ehlers JP (2020) Correlation between brain volume and retinal photoreceptor outer segment volume in normal aging and neurodegenerative diseases. PLoS ONE 15(9):0237078
5. Guo N, Koerts J, Tucha L, Fetter I, Biela C, König M, Bossert M, Diener C, Aschenbrenner S, Weisbrod M et al (2022) Stability of attention performance of adults with ADHD over time: Evidence from repeated neuropsychological assessments in one-month intervals. Int J Environ Res Public Health 19(22): 15234
6. Shoeibi A, Khodatars M, Jafari M, Ghassemi N, Moridian P, Alizadesani R, Ling SH, Khosravi A, Alinejad-Rokny H, Lam H, et al (2022) Diagnosis of brain diseases in fusion of neuroimaging modalities using deep learning: A review. Inf Fusion
7. Wu HM, Hsiao FJ, Chen RS, Shan DE, Hsu WY, Chiang MC, Lin YY (2019) Attenuated NoGo-related beta desynchronisation and synchronisation in Parkinson's disease revealed by magnetoencephalographic recording. Sci Rep 9(1):1–12
8. Hong J, Hwang J, Lee JH (2023) General psychopathology factor (p-factor) prediction using resting-state functional connectivity and a scanner-generalization neural network. J Psychiatr Res 158:114–125
9. Liu S, Duan M, Sun Y, Wang L, An L, Ming D (2023) Neural responses to social decision-making in suicide attempters with mental disorders. BMC Psychiatry 23(1):19
10. Cervenka S, Frick A, Bodén R, Lubberink M (2022) Application of positron emission tomography in psychiatry–methodological developments and future directions. Transl Psychiatry 12(1): 248
11. Gan J, Liu W, Fan J, Yi J, Tan C, Zhu X (2023) Correlates of poor insight: A comparative fMRI and sMRI study in obsessive-compulsive disorder and schizo-obsessive disorder. J Affect Disord 321:66–73
12. Metzak PD, Shakeel MK, Long X, Lasby M, Souza R, Bray S, Goldstein BI, MacQueen G, Wang J, Kennedy SH et al (2022) Brain connectomes in youth at risk for serious mental illness: an exploratory analysis. BMC Psychiatry 22(1):1–18
13. Acharya R, Kafle S, Shrestha DB, Sedhai YR, Ghimire M, Khanal K, Malla QB, Nepal U, Shrestha R, Giri B (2022) Use of computed tomography of the head in patients with acute atraumatic altered mental status: A systematic review and meta-analysis. JAMA Netw Open 5(11):2242805
14. Xin Q, Hu S, Liu S, Zhao L, Zhang YD (2022) An attention-based wavelet convolution neural network for epilepsy EEG classification. IEEE Trans Neural Syst Rehabil Eng 30:957–966
15. Wilson S, Ardle RM, Tolley C, Slight S (2022) Usability and acceptability of wearable technology in the early detection of dementia. Alzheimers Dement 18(e059):820
16. Hassan F, Hussain SF, Qaisar SM (2023) Fusion of multivariate EEG signals for schizophrenia detection using CNN and machine learning techniques. Inf Fusion 92:466–478
17. Puri DV, Nalbalwar SL, Nandgaonkar AB, Gawande JP, Wagh A (2023) Automatic detection of Alzheimer's disease from EEG signals using low-complexity orthogonal wavelet filter banks. Biomed Signal Proc Control 81(104):439
18. Zhao W, Van Someren EJ, Li C, Chen X, Gui W, Tian Y, Liu Y, Lei X (2021) EEG spectral analysis in insomnia disorder: A systematic review and meta-analysis. Sleep Med Rev 59: 101457
19. Boonyakitanont P, Lek-Uthai A, Chomtho K, Songsiri J (2020) A review of feature extraction and performance evaluation in epileptic seizure detection using EEG. Biomed Sig Process Control 57:101702
20. de Almeida WF, de Moraes Lima CA, Peres SM (2021) A systematic mapping of feature extraction and feature selection methods

of electroencephalogram signals for neurological diseases diagnostic assistance. IEEE Lat Am Trans 19(5):735–745
21. Rivera MJ, Teruel MA, Maté A, Trujillo J (2021) Diagnosis and prognosis of mental disorders by means of EEG and deep learning: a systematic mapping study. Artif Intell Rev 1–43, 55:1209–1251
22. Organization WH (2023) Mental health - World Health Organization (WHO). https://www.who.int/health-topics/mental-health#tab=tab_1. Accessed 9 Feb 2023
23. Vos T, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, Abdulkader RS, Abdulle AM, Abebo TA, Abera SF et al (2017) Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the global burden of disease study 2016. Lancet 390(10100):1211–1259
24. Arsalan A, Majid M (2022) A study on multi-class anxiety detection using wearable EEG headband. J Ambient Intell Humanized Comput 13(12):5739–5749
25. Baghdadi A, Aribi Y, Fourati R, Halouani N, Siarry P, Alimi A (2021) Psychological stimulation for anxious states detection based on EEG-related features. J Ambient Intell Humanized Comput 12:8519–8533
26. Alimardani F, Boostani R (2018) DB-FFR: a modified feature selection algorithm to improve discrimination rate between bipolar mood disorder (BMD) and schizophrenic patients. Iran J Sci Technol Trans Electr Eng 42:251–260
27. Khaleghi A, Sheikhani A, Mohammadi MR, Nasrabadi AM, Vand SR, Zarafshan H, Moeini M (2015) EEG classification of adolescents with type I and type II of bipolar disorder. Australas Phys Eng Sci Med 38:551–559
28. Zhang B, Yan G, Yang Z, Su Y, Wang J, Lei T (2020) Brain functional networks based on resting-state EEG data for major depressive disorder analysis and classification. IEEE Trans Neural Syst Rehabil Eng 29:215–229
29. Mahato S, Paul S (2020) Classification of depression patients and normal subjects based on electroencephalogram (EEG) signal using alpha power and theta asymmetry. J Med Syst 44:1–8
30. Akbari H, Sadiq MT, Payan M, Esmaili SS, Baghri H, Bagheri H (2021) Depression detection based on geometrical features extracted from SODP shape of EEG signals and binary PSO. Traitement Sign 38(1):13–26
31. Moghaddari M, Lighvan MZ, Danishvar S (2020) Diagnose ADHD disorder in children using convolutional neural network based on continuous mental task EEG. Comput Methods Prog Biomed 197:105738
32. Bashiri A, Shahmoradi L, Beigy H, Savareh BA, Nosratabadi M, N Kalhori SR, Ghazisaeedi M (2018) Quantitative EEG features selection in the classification of attention and response control in the children and adolescents with attention deficit hyperactivity disorder. Futur Sci OA 4(5):FSO292
33. Abdolzadegan D, Moattar MH, Ghoshuni M (2020) A robust method for early diagnosis of autism spectrum disorder from EEG signals based on feature selection and DBSCAN method. Biocybernetics Biomed Eng 40(1):482–493
34. Baygin M, Dogan S, Tuncer T, Barua PD, Faust O, Arunkumar N, Abdulhay EW, Palmer EE, Acharya UR (2021) Automated ASD detection using hybrid deep lightweight features extracted from EEG signals. Comput Biol Med 134(104):548
35. Kanoga S, Mitsukura Y (2017) A study of pattern recognition in children using single-channel electroencephalogram for specialized electroencephalographic devices. Electron Commun Jpn 100(11):43–53
36. Uyulan C, Erguzel TT (2017) Analysis of time-frequency EEG feature extraction methods for mental task classification. Int J Comput Intell Syst 10(1):1280
37. Amini M, Pedram MM, Moradi A, Ouchani M (2021) Diagnosis of Alzheimer's disease by time-dependent power spectrum descriptors and convolutional neural network using EEG signal. Comput Math Methods Med 2021, p 17. https://doi.org/10.1155/2021/5511922,
38. Wen D, Cheng Z, Li J, Zheng X, Yao W, Dong X, Saripan MI, Li X, Yin S, Zhou Y (2021) Classification of ERP signal from amnestic mild cognitive impairment with type 2 diabetes mellitus using single-scale multi-input convolution neural network. J Neurosci Methods 363:109353
39. Sharma N, Kolekar MH, Jha K (2021) EEG based dementia diagnosis using multi-class support vector machine with motor speed cognitive test. Biomed Signal Process Control 63(102):102
40. Khare SK, Bajaj V, Acharya UR (2021) PDCNNet: An automatic framework for the detection of Parkinson's disease using EEG signals. IEEE Sensors J 21(15):17017–17024
41. Aslan Z (2021) Migraine detection from EEG signals using tunable Q-factor wavelet transform and ensemble learning techniques. Phys Eng Sci Med 44(4):1201–1212
42. Craley J, Johnson E, Jouny C, Venkataraman A (2021) Automated inter-patient seizure detection using multichannel convolutional and recurrent neural networks. Biomed Signal Proc Control 64(102):360
43. Tuncer T, Dogan S, Ertam F, Subasi A (2020) A novel ensemble local graph structure based feature extraction network for EEG signal analysis. Biomed Signal Proc Control 61:102006
44. Chakraborty M, Mitra D et al (2021) Automated detection of epileptic seizures using multiscale and refined composite multiscale dispersion entropy. Chaos, Solitons Fractals 146:110939
45. Erguzel TT, Ozekes S, Sayar GH, Tan O, Tarhan N (2015) A hybrid artificial intelligence method to classify trichotillomania and obsessive compulsive disorder. Neurocomputing 161:220–228
46. Hussain S, Pirzada N, Saba E, Panhwar MA, Ahmed T (2021) Evaluating domain knowledge and time series features for automated detection of schizophrenia from EEG signals. Int J Adv Comput Sci Appl 12(11):530–535
47. Das K, Pachori RB (2021) Schizophrenia detection technique using multivariate iterative filtering and multichannel EEG signals. Biomed Signal Process Control 67(102):525
48. Baygin M (2021) An accurate automated schizophrenia detection using TQWT and statistical moment based feature extraction. Biomed Signal Process Control 68(102):777
49. Gutiérrez-Tobal GC, Alvarez D, Del Campo F, Hornero R (2015) Utility of AdaBoost to detect sleep apnea-hypopnea syndrome from single-channel airflow. IEEE Trans Biomed Eng 63(3):636–646
50. McCloskey S, Jeffries B, Koprinska I, Miller CB, Grunstein RR (2019) Data-driven cluster analysis of insomnia disorder with physiology-based qEEG variables. Knowl-Based Syst 183:104863
51. Erdamar A, Aksahin MF (2020) Quantitative sleep EEG synchronization analysis for automatic arousals detection. Biomed Signal Process Control 59:101895
52. Anuragi A, Sisodia DS, Pachori RB (2020) Automated alcoholism detection using Fourier-Bessel series expansion based empirical wavelet transform. IEEE Sensors J 20(9):4914–4924
53. Xiong Y, Gao J, Zhang J (2019) Detection methamphetamine patients using ERP features. In: 2019 6th International Conference on Information Science and Control Engineering (ICISCE), IEEE, pp 259–262
54. Hafeez M, Idrees MD, Kim JY (2017) Development of a diagnostic algorithm to identify psycho-physiological game addiction attributes using statistical parameters. IEEE Access 5:22443–22452
55. Zhang Y, Wu W, Toll RT, Naparstek S, Maron-Katz A, Watts M, Gordon J, Jeong J, Astolfi L, Shpigel E et al (2021) Identifica-

tion of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography. Nat Biomed Eng 5(4):309–323
56. Ghanbari Z, Moradi MH, Moradi A, Mirzaei J (2020) Resting state functional connectivity in PTSD veterans: an EEG study. J Med Biol Eng 40:505–516
57. Ahirwal MK (2020) Analysis and identification of EEG features for mental stress. In: Evolution in Computational Intelligence: Frontiers in Intelligent Computing: Theory and Applications (FICTA 2020), vol 1. Springer, pp 201–209
58. Hag A, Handayani D, Altalhi M, Pillai T, Mantoro T, Kit MH, Al-Shargie F (2021) Enhancing EEG-based mental stress state recognition using an improved hybrid feature selection algorithm. Sensors 21(24):8370
59. Acharya JN, Hani AJ, Thirumala P, Tsuchida TN (2016) American clinical neurophysiology society guideline 3: a proposal for standard montages to be used in clinical eeg. Neurodiagnostic J 56(4):253–260
60. Babiloni C, Barry RJ, Başar E, Blinowska KJ, Cichocki A, Drinkenburg WH, Klimesch W, Knight RT, da Silva FL, Nunez P, Oostenveld R, Jeong J, Pascual-Marqui R, Valdes-Sosa P, Hallett M (2020) International federation of clinical neurophysiology (IFCN) – EEG research workgroup: Recommendations on frequency and topographic analysis of resting state EEG rhythms. part 1: Applications in clinical research studies. Clin Neurophysiol 131(1):285–307. https://doi.org/10.1016/j.clinph.2019.06.234
61. Sörnmo L, Laguna P (2005) Bioelectrical signal processing in cardiac and neurological applications, vol 8. Academic Press
62. Kumar JS, Bhuvaneswari P (2012) Analysis of Electroencephalography (EEG) signals and its categorization-a study. Procedia Eng 38:2525–2536
63. Birbaumer N, Elbert T, Canavan AG, Rockstroh B (1990) Slow potentials of the cerebral cortex and behavior. Physiol Rev 70(1):1–41
64. Yamada T, Meng E (2012) Practical guide for clinical neurophysiologic testing: EEG. Lippincott Williams & Wilkins
65. Bermudez D, Steyrl D, Müller-Putz G, Pock T (2020) Implementation of machine learning algorithm to exploit information from multimodal fMRI/EEG fused image data. https://doi.org/10.13140/RG.2.2.28622.82240
66. Zheng A, Casari A (2018) Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media, Inc
67. Bracewell RN, Bracewell RN (1986) The Fourier transform and its applications, vol 31999. McGraw-hill, New York
68. Bentley PM, McDonnell J (1994) Wavelet transforms: an introduction. Electron Commun Eng J 6(4):175–186
69. Clarivate Web of Science. https://www.webofscience.com/. Accessed 13 Feb 2023
70. Goodfellow I, Bengio Y, Courville A (2016) Deep Learning. MIT Press. http://www.deeplearningbook.org
71. Wang W, Charkborty G (2021) Automatic prognosis of lung cancer using heterogeneous deep learning models for nodule detection and eliciting its morphological features. Appl Intell 51(4):2471–2484
72. Liu W, Fan H, Xia M (2022) Multi-grained and multi-layered gradient boosting decision tree for credit scoring. Appl Intell 52(5):5325–5341
73. Mumtaz W, Qayyum A (2019) A deep learning framework for automatic diagnosis of unipolar depression. Int J Med Inform 132(103):983
74. Khafaga DS, Auvdaiappan M, Deepa K, Abouhawwash M, Karim FK (2023) Deep learning for depression detection using Twitter data. Intel Autom Soft Comput 36(2):1301–1313
75. Jain R, Jain N, Aggarwal A, Hemanth DJ (2019) Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. Cogn Syst Res 57:147–159
76. Wang J, Knol MJ, Tiulpin A, Dubost F, de Bruijne M, Vernooij MW, Adams HH, Ikram MA, Niessen WJ, Roshchupkin GV (2019) Gray matter age prediction as a biomarker for risk of dementia. Proc Natl Acad Sci 116(42):21213–21218
77. Nasseri M, Attia TP, Joseph B, Gregg NM, Nurse ES, Viana PF, Schulze-Bonhage A, Dümpelmann M, Worrell G, Freestone DR et al (2021) Non-invasive wearable seizure detection using long-short-term memory networks with transfer learning. J Neural Eng 18(5):056017
78. Gulum MA, Trombley CM, Kantardzic M (2021) A review of explainable deep learning cancer detection models in medical imaging. Appl Sci 11(10):4573
79. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow PM, Zietz M, Hoffman MM et al (2018) Opportunities and obstacles for deep learning in biology and medicine. J R Soc Interface 15(141):20170387
80. Gandolfi M, Galazzo IB, Pavan RG, Cruciani F, Valè N, Picelli A, Storti SF, Smania N, Menegaz G (2022) eXplainable AI allows predicting upper limb rehabilitation outcomes in sub-acute stroke patients. IEEE J Biomed Health Inform 27(1):263–273
81. Gimeno M, Villar S, Agirre X, Prosper F, Rubio A, Carazo F, et al (2022) Explainable artificial intelligence for precision medicine in acute myeloid leukemia. Front Immunol p 13, https://doi.org/10.3389/fimmu.2022.977358
82. Bishop CM, Nasrabadi NM (2006) Pattern recognition and machine learning, vol 4. Springer
83. Kuhn M, Johnson K et al (2013) Applied predictive modeling, vol 26. Springer
84. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
85. Site A, Nurmi J, Lohan ES (2021) Systematic review on machine-learning algorithms used in wearable-based eHealth data analysis, 9:112221–112235. IEEE Access
86. Budgen D, Turner M, Brereton P, Kitchenham B (2008) Using mapping studies in software engineering. Proceedings of PPIG 2008:2
87. Petersen K, Feldt R, Mujtaba S, Mattsson M (2008) Systematic mapping studies in software engineering. In: 12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12, pp 1–10. Electronic Workshops in Computing (eWiC), England, Wales and Scotland
88. Kitchenham B, Budgen D, Brereton P (2011) Using mapping studies as the basis for further research - a participant-observer case study. Inf Softw Technol 53:638–651. https://doi.org/10.1016/j.infsof.2010.12.011
89. Kitchenham B, Charters S (2007) Guidelines for performing systematic literature reviews in software engineering 2. Gaskell
90. Seshadri NG, Agrawal S, Singh BK, Geethanjali B, Mahesh V, Pachori RB (2023) EEG based classification of children with learning disabilities using shallow and deep neural network. Biomed Signal Process Control 82:104553
91. Aliyu I, Lim CG (2021) Selection of optimal wavelet features for epileptic EEG signal classification with LSTM. Neural Comput Applic, 35:1–21
92. Sharma Y, Singh BK (2023) Attention deficit hyperactivity disorder detection in children using multivariate empirical EEG decomposition approaches: A comprehensive analytical study. Expert Syst Appl 213(119):219
93. Ham K (2013) Openrefine (version 2.5). http://openrefine.org.free, open-source tool for cleaning and transforming data. J Med Libr Assoc JMLA 101(3):233
94. Faraone SV, Banaschewski T, Coghill D, Zheng Y, Biederman J, Bellgrove MA, Newcorn JH, Gignac M, Al Saud NM, Manor I,

Rohde LA, Yang L, Cortese S, Almagor D, Stein MA, Albatti TH, Aljoudi HF, Alqahtani MM, Asherson P, Atwoli L, Bölte S, Buitelaar JK, Crunelle CL, Daley D, Dalsgaard S, Döpfner M, Espinet (on behalf of CADDRA) S, Fitzgerald M, Franke B, Gerlach M, Haavik J, Hartman CA, Hartung CM, Hinshaw SP, Hoekstra PJ, Hollis C, Kollins SH, Sandra Kooij J, Kuntsi J, Larsson H, Li T, Liu J, Merzon E, Mattingly G, Mattos P, McCarthy S, Mikami AY, Molina BS, Nigg JT, Purper-Ouakil D, Omigbodun OO, Polanczyk GV, Pollak Y, Poulton AS, Rajkumar RP, Reding A, Reif A, Rubia K, Rucklidge J, Romanos M, Ramos-Quiroga JA, Schellekens A, Scheres A, Schoeman R, Schweitzer JB, Shah H, Solanto MV, Sonuga-Barke E, Soutullo C, Steinhausen HC, Swanson JM, Thapar A, Tripp G, van de Glind G, van den Brink W, Van der Oord S, Venter A, Vitiello B, Walitza S, Wang Y, (2021) The world federation of ADHD international consensus statement: 208 evidence-based conclusions about the disorder. Neurosci Biobehav Rev 128:789–818. https://doi.org/10.1016/j.neubiorev.2021.01.022
95. Goodman A (1990) Addiction: definition and implications. Br J Addict 85:1403–8
96. Arsalan A, Majid M, Anwar SM (2020) Electroencephalography based machine learning framework for anxiety classification. In: Intelligent Technologies and Applications: Second International Conference, INTAP 2019, Bahawalpur, Pakistan, November 6–8, 2019, Revised Selected Papers 2, Springer, pp 187–197
97. Hodges H, Fealko C, Soares N (2020) Autism spectrum disorder: definition, epidemiology, causes, and clinical evaluation. Transl Pediatr 9(Suppl 1):S55
98. Gustafson L (1996) What is dementia? Acta Neurol Scand 94:22–24
99. World Health Organization, et al (1992) The ICD-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. World Health Organization
100. Siegel LS (2006) Perspectives on dyslexia. Paediatr Child Health 11(9):581–587
101. Fisher RS, Acevedo C, Arzimanoglou A, Bogacz A, Cross JH, Elger CE, Engel J Jr, Forsgren L, French JA, Glynn M et al (2014) ILAE official report: a practical clinical definition of epilepsy. Epilepsia 55(4):475–482
102. Goadsby PJ, Lipton RB, Ferrari MD (2002) Migraine–current understanding and treatment. N Engl J Med 346(4):257–270
103. Pauls DL (2008) The genetics of obsessive compulsive disorder: a review of the evidence. American Journal of Medical Genetics Part C: Seminars in Medical Genetics, Wiley Online Library 148:133–139
104. Poewe W, Seppi K, Tanner CM, Halliday GM, Brundin P, Volkmann J, Schrag AE, Lang AE (2017) Parkinson disease. Nat Rev Dis Prim 3(1):1–21
105. National Collaborating Centre for Mental Health (UK, et al) (2005) Post-traumatic stress disorder: The management of PTSD in adults and children in primary and secondary care. Gaskell, Leicester (UK)
106. Charlson FJ, Ferrari AJ, Santomauro DF, Diminic S, Stockings E, Scott JG, McGrath JJ, Whiteford HA (2018) Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. Schizophr Bull 44(6):1195–1203
107. White DP (2006) Sleep apnea. Proc Am Thorac Soc 3(1):124–128
108. Roth T (2007) Insomnia: definition, prevalence, etiology, and consequences. J Clin Sleep Med 3(5 suppl):S7–S10
109. Halász P, Terzano M, Parrino L, Bódizs R (2004) The nature of arousal in sleep. J Sleep Res 13(1):1–23
110. Fink G (2016) Stress, definitions, mechanisms, and effects outlined: Lessons from anxiety. In: Stress: Concepts, cognition, emotion, and behavior, Elsevier, pp 3–11
111. Müller-Gerking J, Pfurtscheller G, Flyvbjerg H (1999) Designing optimal spatial filters for single-trial EEG classification in a movement task. Clin Neurophysiol 110(5):787–798
112. Supriya S, Siuly S, Wang H, Zhang Y (2021) Epilepsy detection from EEG using complex network techniques: A review. IEEE Rev Biomed Eng, 16:292–306
113. Ahmed N, Natarajan T, Rao KR (1974) Discrete cosine transform. IEEE Trans Comput 100(1):90–93
114. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, Yen NC, Tung CC, Liu HH (1998) The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc R Soc Lond Ser A Math Phys Eng Sci 454(1971):903–995
115. Mermelstein P (1976) Distance measures for speech recognition, psychological and instrumental. Pattern Recognit Artif Intell 116:374–388
116. Rilling G, Flandrin P, Goncalves P et al (2003) On empirical mode decomposition and its algorithms. In: IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, vol 3. Citeseer, pp 8–11
117. Sejdić E, Djurović I, Jiang J (2009) Time-frequency feature representation using energy concentration: An overview of recent advances. Digit Signal Process 19(1):153–183
118. Sifuzzaman M, Islam MR, Ali M (2009) Application of wavelet transform and its advantages compared to Fourier transform. J Phys Sci, 13:121–134.
119. Kingma DP, Welling M et al (2019) An introduction to variational autoencoders. Found Trends Mach Learn 12(4):307–392
120. Zhang Y, Liu B, Ji X, Huang D (2017) Classification of EEG signals based on autoregressive model and wavelet packet decomposition. Neural Process Lett 45(2):365–378
121. Übeyli ED (2010) Lyapunov exponents/probabilistic neural networks for analysis of EEG signals. Expert Syst Appl 37(2):985–992
122. Akar SA, Kara S, Latifoğlu F, Bilgiç V (2016) Analysis of the complexity measures in the EEG of schizophrenia patients. Int J Neural Syst 26(02):1650008
123. Woyshville M, Calabrese J (1994) Quantification of occipital EEG changes in Alzheimer's disease utilizing a new metric: The fractal dimension. Biol Psychiatry 35(6):381–387. https://doi.org/10.1016/0006-3223(94)90004-3. cited By 66
124. Guevara MA, Corsi-Cabrera M (1996) EEG coherence or EEG correlation? Int J Psychophysiol 23(3):145–153
125. Puthankattil SD, Joseph PK (2012) Classification of EEG signals in normal and depression conditions by ANN using RWE and signal entropy. J Mech Med Biol 12(04):1240019
126. Liang Z, Wang Y, Sun X, Li D, Voss LJ, Sleigh JW, Hagihira S, Li X (2015) EEG entropy measures in anesthesia. Front Comput Neurosci 9:16
127. Kaya Y, Uyar M, Tekin R, Yıldırım S (2014) 1D-local binary pattern based feature extraction for classification of epileptic EEG signals. Appl Math Comput 243:209–219
128. Sameer M, Gupta AK, Chakraborty C, Gupta B (2020) ROC analysis for detection of epileptical seizures using Haralick features of gamma band. In: 2020 National Conference on Communications (NCC), IEEE, pp 1–5
129. Adkinson JA, Karumuri B, Hutson TN, Liu R, Alamoudi O, Vlachos I, Iasemidis L (2018) Connectivity and centrality characteristics of the epileptogenic focus using directed network analysis. IEEE Trans Neural Syst Rehabil Eng 27(1):22–30
130. Kaminski MJ, Blinowska KJ (1991) A new method of the description of the information flow in the brain structures. Biol Cybern 65(3):203–210
131. Hurst HE (1951) Long-term storage capacity of reservoirs. Trans Am Soc Civ Eng 116(1):770–799

132. Hjorth B (1970) EEG analysis based on time domain properties. Electroencephalogr Clin Neurophysiol 29(3):306–310
133. Shahid A, Kamel N, Malik AS, Jatoi MA (2013) Epileptic seizure detection using the singular values of EEG signals. In: 2013 ICME international conference on complex medical engineering, IEEE, pp 652–655
134. Taran S, Bajaj V, Sharma D (2017) Robust Hermite decomposition algorithm for classification of sleep apnea EEG signals. Electron Lett 53(17):1182–1184
135. Dlask M, Kukal J (2021) Alzheimer disease diagnostics from EEG via Wishart distribution of fractional processes. SIViP 15(7):1435–1442
136. Gangurde HD (2014) Feature selection using clustering approach for big data. Int J Comput Appl 975:1–3
137. Kira K, Rendell LA (1992) A practical approach to feature selection. In: Machine learning proceedings 1992, Elsevier, pp 249–256
138. Tan F, Fu X, Zhang Y, Bourgeois AG (2008) A genetic algorithm-based method for feature subset selection. Soft Comput 12(2):111–120
139. Ng AY (2004) Feature selection, L1 vs. L2 regularization, and rotational invariance. In: Proceedings of the twenty-first international conference on Machine learning, p 78. Association for Computing Machinery, New York, NY, United States
140. Bianchi L, Dorigo M, Gambardella LM, Gutjahr WJ (2009) A survey on metaheuristics for stochastic combinatorial optimization. Nat Comput 8(2):239–287
141. Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Phil Trans R Soc A Math Phys Eng Sci 374(2065):20150202
142. Beni G, Wang J (1993) Swarm intelligence in cellular robotic systems. In: Robots and biological systems: towards a new bionics? Springer, pp 703–712
143. Swain PH, Hauska H (1977) The decision tree classifier: Design and potential. IEEE Trans Geosci Electron 15(3):142–147. https://doi.org/10.1109/TGE.1977.6498972
144. Biau G, Scornet E (2016) A random forest guided tour. Test 25(2):197–227
145. Chen T, He T, Benesty M, Khotilovich V, Tang Y, Cho H, Chen K, et al (2015) XGBoost: extreme gradient boosting. R Packag Version 04-2 1(4):1–4
146. Friedman N, Geiger D, Goldszmidt M (1997) Bayesian network classifiers. Mach Learn 29(2):131–163
147. Kuncheva L (2000) Fuzzy classifier design, vol 49. Springer Science & Business Media
148. Eddy SR (2004) What is a hidden Markov model? Nat Biotechnol 22(10):1315–1316
149. Kleinbaum DG, Dietz K, Gail M, Klein M, Klein M (2002) Logistic regression. Springer
150. Xanthopoulos P, Pardalos PM, Trafalis TB (2013) Linear Discriminant Analysis, Springer New York, New York, NY, pp 27–33. https://doi.org/10.1007/978-1-4419-9878-1_4
151. Fix E, Hodges JL (1989) Discriminatory analysis. Nonparametric discrimination: Consistency properties. Int Stat Rev/Rev Int Stat 57(3):238–247
152. Noriega L (2005) Multilayer perceptron tutorial. School of Computing Staffordshire University
153. LeCun Y, Bengio Y, et al (1995) Convolutional networks for images, speech, and time series. Handb Brain Theory Neural Netw 3361(10):1995
154. Medsker L, Jain LC (1999) Recurrent neural networks: design and applications. CRC Press
155. Huang GB, Zhu QY, Siew CK (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1–3):489–501
156. Li Y, Yu ZL, Bi N, Xu Y, Gu Z, Amari Si (2014) Sparse representation for brain signal processing: a tutorial on methods and applications. IEEE Signal Proc Mag 31(3):96–106
157. Boashash B, Ouelha S (2018) Designing high-resolution time-frequency and time-scale distributions for the analysis and classification of non-stationary signals: a tutorial review with a comparison of features performance. Digit Signal Process 77:120–152
158. Thangarajoo RG, Reaz MBI, Srivastava G, Haque F, Ali SHM, Bakar AAA, Bhuiyan MAS (2021) Machine learning-based epileptic seizure detection methods using wavelet and EMD-based decomposition techniques: A review. Sensors 21(24): 8485
159. Lopes R, Betrouni N (2009) Fractal and multifractal analysis: a review. Med Image Anal 13(4):634–649
160. Boashash B, Khan NA, Ben-Jabeur T (2015) Time-frequency features for pattern recognition using high-resolution TFDs: A tutorial review. Digit Signal Proc 40:1–30
161. Remeseiro B, Bolon-Canedo V (2019) A review of feature selection methods in medical applications. Computers in biology and medicine 112:103375
162. Raghavendra U, Acharya UR, Adeli H (2019) Artificial intelligence techniques for automated diagnosis of neurological disorders. Eur Neurol 82(1–3):41–64
163. Shoeibi A, Khodatars M, Ghassemi N, Jafari M, Moridian P, Alizadehsani R, Panahiazar M, Khozeimeh F, Zare A, Hosseini-Nejad H et al (2021) Epileptic seizures detection using deep learning techniques: a review. Int J Environ Res Public Health 18(11):5780
164. Ansart M, Epelbaum S, Bassignana G, Bône A, Bottani S, Cattai T, Couronné R, Faouzi J, Koval I, Louis M et al (2021) Predicting the progression of mild cognitive impairment using machine learning: a systematic, quantitative and critical review. Med Image Anal 67(101):848
165. Sánchez-Reyes LM, Rodríguez-Reséndiz J, Avecilla-Ramírez GN, García-Gomar ML, Robles-Ocampo JB (2021) Impact of EEG parameters detecting dementia diseases: A systematic review. IEEE Access
166. Annakutty AA, Aponso AC (2016) Review of brain imaging techniques, feature extraction and classification algorithms to identify Alzheimer's disease. Int J Pharma Med Biol Sci 5(3):178–183
167. Acharya UR, Sree SV, Swapna G, Martis RJ, Suri JS (2013) Automated EEG analysis of epilepsy: a review. Knowl-Based Syst 45:147–165
168. Boonyakitanont P, Lek-Uthai A, Chomtho K, Songsiri J (2020) A review of feature extraction and performance evaluation in epileptic seizure detection using EEG. Biomed Signal Process Control 57:101702
169. Saminu S, Xu G, Shuai Z, Abd El Kader I, Jabire AH, Ahmed YK, Karaye IA, Ahmad IS (2021) A recent investigation on detection and classification of epileptic seizure techniques using EEG signal. Brain Sci 11(5):668
170. Acharya UR, Hagiwara Y, Adeli H (2018) Automated seizure prediction. Epilepsy Behav 88:251–261
171. Xu Y, Yang J, Sawan M (2021) Trends and challenges of processing measurements from wearable devices intended for epileptic seizure prediction. J Signal Proc Syst 1–16
172. Aboalayon KAI, Faezipour M, Almuhammadi WS, Moslehpour S (2016) Sleep stage classification using EEG signal analysis: a comprehensive survey and new investigation. Entropy 18(9):272
173. Assi EB, Nguyen DK, Rihana S, Sawan M (2017) Towards accurate prediction of epileptic seizures: A review. Biomed Signal Proc Control 34:144–157
174. Mahato S, Paul S (2019) Electroencephalogram (EEG) signal analysis for diagnosis of major depressive disorder (MDD): a review. Nanoelectron Circ Commun Syst, 511:323–335

175. Perera H, Shiratuddin MF, Wong KW (2018) Review of EEG-based pattern classification frameworks for dyslexia. Brain Inform 5(2):1–14
176. Adeli H, Zhou Z, Dadmehr N (2003) Analysis of EEG records in an epileptic patient using wavelet transform. J Neurosci Methods 123(1):69–87
177. Moskovitch R (2022) Multivariate temporal data analysis-a review. Wiley Interdiscip Rev Data Min Knowl Disc 12(1):1430
178. Wan Z, Yang R, Huang M, Zeng N, Liu X (2021) A review on transfer learning in EEG signal analysis. Neurocomputing 421:1–14
179. Farahani A, Voghoei S, Rasheed K, Arabnia HR (2021) A brief review of domain adaptation. Advances in Data Science and Information Engineering: Proceedings from ICDATA 2020 and IKE 2020, pp 877–894. Springer, Cham, Switzerland

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Sandra García-Ponsoda holds a Mathematics degree from the University of Alicante since 2019, where she also obtained a M.Sc. in Data Science in 2021. Currently, she is pursuing a PhD in Computer Science, holding a predoctoral scholarship granted by ValgrAI – Valencian Graduate School and Research Network of Artificial Intelligence and the Generalitat Valenciana, and co-founded by the European Union. She is also researching in the Lucentia Research group. Her research involves feature engineering, Deep Learning, eHealth, and EEG signal processing.

Jorge García-Carrasco holds a Physics degree from the University of Alicante since 2020, where he also obtained a M.Sc. in Data Science in 2021. Currently he is pursuing a PhD, holding a predoctoral scholarship granted by the Conselleria de Innovación, Universidades, Ciencia y Sociedad Digital (Generalitat Valenciana), and researching in the Lucentia Research group. His research involves the areas of AI Interpretability and Explainability, Deep Learning, Computer Vision and AI applied to medical data.

Miguel A. Teruel is an Assistant Professor in the Department of Department of Software and Computing Systems and Director of the Master's Degree in Development of Applications and Web Services at the University of Alicante. Before becoming an Assistant Professor, he was a senior IT consultant at various companies, as well as CEO of MTMinfo. He held a doctoral scholarship from the Spanish Ministry of Education, Culture and Sport from 2021 to 2016, with an international stay at the VU Universiteit in Amsterdam. During his research stage, he has published papers in relevant international conferences and JCR journals. His current research interests are Post-WIMP environments, eHealth, Big Data Analytics, Gamification and Blockchain.

Alejandro Maté holds a Computer Science Engineering degree from the University of Alicante since 2009, where he also obtained a M.Sc. in Computer Science Technology in 2010 and a Ph.D. in 2013. Since 2019, he has been Associate Professor at the University of Alicante. He has published over 50 papers in international conferences (e.g. ER, CAiSE, RE) and JCR journals (e.g. Information Systems, Future Generations, Information & Software Technology). His research involves conceptual modelling, data warehouses, model driven development, and requirements engineering.

Juan Trujillo is a Full Professor at the University of Alicante, Dept. of Software and Computing Systems. Since he got his Ph.D. in 2001, he has been leading the Business Intelligence and Big Data research in the department and has also been the founder and director of The Lucentia Research Group since 2008. His main research topics include Business Intelligence, Big Data, Data Warehouses, Data Analytics, Decision Support System, Artificial Intelligence, Machine Learning and health. He is author of more than 200 conference paper, such as ER, UML, DAWAK or CAiSE, and more than 70 JCR papers, such as DKE, DSS, ISOFT, IS, or InfSci.

Authors and Affiliations

Sandra García-Ponsoda[1,2] · **Jorge García-Carrasco**[1] · **Miguel A. Teruel**[1,3] · **Alejandro Maté**[1,3] · **Juan Trujillo**[1,3]

Jorge García-Carrasco
jorge.g@ua.es

Miguel A. Teruel
materuel@dlsi.ua.es

Alejandro Maté
amate@dlsi.ua.es

Juan Trujillo
jtrujillo@dlsi.ua.es

1 Lucentia Research Group - Department of Software and Computing Systems, University of Alicante, Carretera de San Vicente del Raspeig, s/n, San Vicente del Raspeig 03690, Spain

2 ValgrAI - Valencian Graduate School and Research Network for Artificial Intelligence, Camí de Vera s/n, Valencia 46022, Spain

3 Alicante Institute for Health and Biomedical Research (ISABIAL), Alicante, Spain