



# **Classificador BCI de 4 classes utilitzant BCI IIIa**

Memòria de Projecte Final de Grau/Màster  
**Màster Universitari en Enginyeria de Telecomunicació UOC-URL**  
Processament de Senyal

**Autor: Antea Herrera Alonso**

Professor: Jordi Solé Casals

09/06/2019

# Copyright



Aquesta obra està subjecta a una llicència de [Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

## Abstract

L'objectiu del TFM és la realització d'un programa que donada una extracció d'imatges cerebrals EEG d'una persona permeti determinar si aquesta persona estava pensant en moure el braç dret, el braç esquerra, el peu o la llengua.

Per assolir l'objectiu el programa disposa de 3 etapes; processat de dades, extracció de les característiques del senyal i classificació. Dins del processat de dades es realitzen les modificacions al senyal d'entrada oportunes per realitzar de manera òptima l'extracció de característiques. A l'extracció de les característiques del senyal s'extrauran les característiques del senyal necessàries per poder realitzar posteriorment la correcta classificació de les imatges. Finalment, la classificació del senyal permet obtenir la probabilitat de que l'individu estigui pensant en cadascuna de les 4 opcions que disposa. Per tant, s'obtidran 4 sortides contínues de la probabilitat de que l'individu estigui pensant en cadascuna de les possibilitats de manera única i la opció amb més probabilitat de les 4 serà la que es determina que esta pensant l'individu.

Paraules clau: EEG, BCI, Extracció de característiques.

## **Abstract (english version)**

The purpose of the TFM is the realization of a program that given a person's EEG brain image extraction allows to determine if that person was thinking of moving the right arm, the left arm, the foot or the tongue. To achieve the goal, the program has 3 stages; Data processing, extraction of signal characteristics and classification. Within the processing of data the modifications to the opportune signal of entrance are realized to optimally realize the extraction of characteristics. When extracting the characteristics of the signal, the signal characteristics necessary will be extracted so that the correct classification of the images can be subsequently performed. Finally, the classification of the signal allows to obtain the probability that the individual is thinking in each one of the 4 options available. Therefore, 4 continuous outputs of the probability that the individual is considering each one of the possibilities in a unique way and the most likely option of the 4 will be the one determined to be the individual's thinking.

Keywords: EEG, BCI, Feature extraction.

# Índex

1. Introducció .....	7
2. Objectius .....	8
3. Continguts.....	9
4. Metodologia .....	10
5. Planificació.....	11
6. Plataforma de desenvolupament.....	12
7. Procés de treball.....	13
8. Context .....	14
8.1. Senyal EEG.....	14
8.2. Dataset: BCI competition III Dataset IIIa.....	16
8.3. Biosignal tool.....	17
9. Arquitectura .....	21
9.1. Preprocessat .....	21
9.2. Extracció de característiques.....	21
9.3. Classificació .....	23
10. Projecció a futur.....	25
11. Conclusió .....	26
Annex 1. Lliurables del projecte .....	27
Annex 2. Codi font (extractes) .....	28
Annex 9. Bibliografia.....	30

# Figures i taules

## Índex de figures

Figura 1: Gantt .....	11
Figura 2: Exemple de les diferents ones dins d'un EEG. ....	15
Figura 3: Temps de la mostra .....	16
Figura 4: Posició dels elèctrodes .....	17

## Índex de taules

Taula 1: Camps de l'estructura de capçalera comuna a BioSignal. ....	19
Taula 2: Components de la llibreria BioSignal.....	19

# 1. Introducció

Una interfície cervell-ordinador, o Brain-Computer Interface (BCI), és un sistema d'interacció persona-ordinador que fa ús dels senyals electroencefalogràfics (EEG) per permetre el control d'un component extern. Durant el primer congrés internacional sobre tecnologies BCI celebrat al juny de 1999 a l'Institut de Rensselaerville a Nova York es va proposar la següent definició: "A braincomputer interfície is a communication system that does not depèn on the brain's normal output pathways of peripheral nerves and muscles "[1]. Aquesta definició, fa entendre que un BCI és una interfície que permet traduir els desitjos d'una persona a uns determinats comandaments sense utilitzar accions motores directes.

Els senyals de control requerits en aquest tipus d'interfícies s'obtenen a partir dels potencials bioelèctrics que es generen en el cervell i que poden ser detectats a través d'elèctrodes aplicats en el cuir cabellut. Aquests potencials, amb una amplitud que varia al voltant de les desenes de microvolts, si bé poden ser pertorbades per una varietat d'artefactes, també poden ser correlades o associades a diferents estats mentals. Per exemple, l'activitat elèctrica cerebral pot veure's afectada per el moviment de braços i cames, a així com la visualització, la resolució de problemes o fins i tot només amb la imaginació. Les prestacions d'un BCI dependran en gran mesura de la capacitat que tingui un subjecte per controlar els seus patrons EEG.

Una de les aplicacions més importants dels sistemes BCI és en el camp de la medicina i més concretament en la rehabilitació, contribuint a establir un canal de comunicació i control per a aquells individus amb importants deficiències en les seves funcions motores però que no presenten trastorns a nivell cerebral.

## 2. Objectius

L'objectiu és la realització d'un programa que donada una extracció d'imatges cerebrals EEG d'una persona permeti determinar si aquesta persona estava pensant en moure el braç dret, el braç esquerra, el peu o la llengua.



## 3. Continguts

Dins d'aquest document es descriurà la metodologia i arquitectura utilitzada per a la creació d'un classificador BCI que donada una extracció d'imatges cerebrals EEG d'una persona permeti determinar si aquesta persona estava pensant en moure el braç dret, el braç esquerra, el peu o la llengua. Primerament, es descriuran totes les eines utilitzades i després es descriurà la solució realitzada explicant tots els passos realitzats. Finalment, es mostraran els resultats obtinguts i es donaran unes petites conclusions dels resultats.

## 4. Metodologia

Per aconseguir les diferents fites d'aquest projecte s'ha seguit la metodologia SCRUM. La metodologia SCRUM s'executa en cicles temporals curts i de durada fixa. Cada iteració, que en el cas d'aquest projecte rep el nom de PAC proporciona un resultat complet amb un increment del producte final que s'ha entregat al tutor. A continuació es detallen totes les iteracions realitzades:

- Iteració1: PAC1.
- Iteració2: PAC2.
- Iteració3: PAC3.
- Iteració4: PAC4 .

# 5. Planificació

El treball ha tingut 4 fases. Cadascuna d'elles va relacionada amb una fita a aconseguir i un document a entregar. Els entregables són:

- PAC1 (17/03/2019): Definició del projecte.
- PAC2 (26/05/2019): Entrega del programa incloent l'extracció de característiques.
- PAC3 (09/06/2019): Entrega de la memòria final incloent una comparativa dels diferents classificadors.
- PAC4 (16/06/2019): Entrega de la presentació del projecte.

El diagrama Gantt que representa aquestes fases i fites és el que es mostra a continuació:

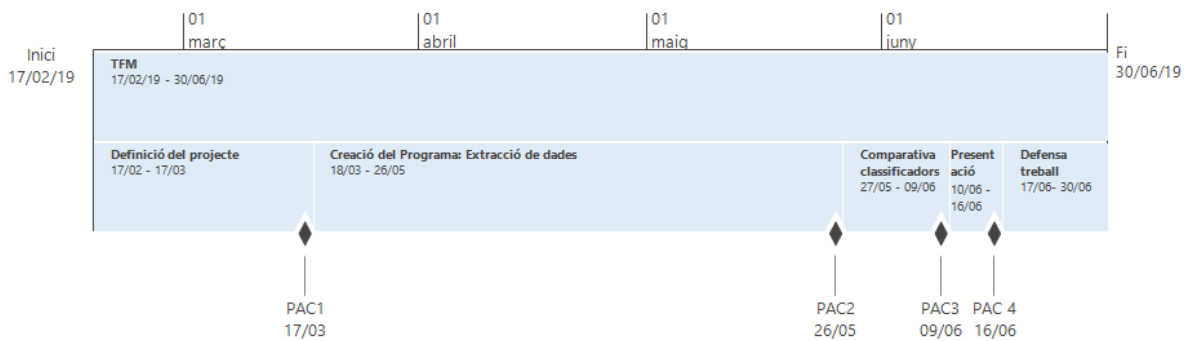


Figura 1: Gantt

## 6. Plataforma de desenvolupament

Per a la realització d'aquest treball han sigut necessaris els recursos tecnològics de software següents:

- Matlab
- Biosignal (mòdul compatible amb Matlab)
- Dataset IIIa

## 7. Procés de treball

El desenvolupament del projecte es pot estructurar en base als lliuraments, ja que han estat escollits en funció de les fases realitzades del projecte.

- Definició del projecte. Inclòs al lliuraments PAC1.
- Realització de la etapa de preprocessat. Inclòs a la PAC2.
- Realització de la etapa de extracció de característiques. Inclòs a la PAC2.
- Realització d la etapa de classificació. Inclòs a la PAC3.
- Resultats i conclusions del projecte. Inclòs a la PAC3 i PAC4.
- Presentació del projecte. Inclòs a la PAC4.

# 8. Context

En aquest apartat es posarà en context el programa realitzat, això vol dir que es descriurà la senyal d'entrada al sistema, tant en les característiques tècniques del senyal fins com s'ha extret aquest senyal per obtenir la base de dades que s'ha utilitzat. També es descriurà el programari utilitzat per a aquest treball.

## 8.1. Senyal EEG

Per poder analitzar correctament les dades procedents del dataset, s'introduiran primerament uns coneixements bàsics sobre les ones que es poden trobar en un senyal EEG normal per una persona de 20 a 60 anys incloent també les particularitats de cada ona. Aquest coneixement es necessari per poder dissenyar de la millor manera el nostre sistema classificador.

Els diferents tipus d'ona són:

### 1. Les Ones Delta (1 a 3 Hz)

Són molt habituals en els nadons i en els infants més petits, de manera que a mesura que envellim, tendim a produir menys ones d'aquest tipus. Cal destacar que el son i la capacitat per a descansar es van perdent, poc a poc, a mesura que passen els anys.

- Un nivell adequat d'ones delta afavoreix i cuida del sistema immunitari, del nostre descans i de la capacitat d'aprenentatge.

### 2. Ones Theta (3,5 a 8 Hz)

El segon dels tipus d'ones cerebrals va dels 3,5 als 8 Hz i es relaciona sobretot amb les capacitats imaginatives, amb la reflexió i el son. Com a curiositat, les ones Theta acostumen a mostrar una elevada activitat quan s'experimenten emocions molt profundes.

- Un pic elevat d'ones Theta pot relacionar-se amb algun trastorn depressiu o falta d'atenció.
- Els pics baixos es relacionen amb ansietat, estrès i baixa auto-consciència emocional.

### 3. Ones Alfa (8 a 13 Hz)

Les Alfa apareixen en aquell crepuscle intermig on hi ha calma, però no son, on hi ha relaxació i un estat propici per a meditar. Es pot experimentar també quan s'està descansant al sofà veient la televisió o al llit descansant, però sense arribar al punt de quedar-se adormit.

- Un nivell elevat d'ones alfa ens impediria poder centrar l'atenció o sentir-nos amb molt poques forces per a realitzar una tasca.
- Un nivell baix es relaciona amb ansietat, estrès i insomni.

### 4. Ones Beta (12 a 33 Hz)

Espectre de freqüències més altes que apareixen com a resultat d'una activitat neuronal intensa. Es parla d'estats molt interessants, i tanmateix complexos. Estats que es relacionen amb aquelles activitats quotidianes on es posa tota l'atenció de l'individu, quan s'està en un estat d'alerta i quan al mateix temps un ha d'estar pendent de diversos estímuls. Activitats tan comunes com conduir, realitzar un examen, fer una exposició, estar en una reunió de treball presentant un projecte, etc., són moments de màxima activació.

- Un nivell baix d'ones Beta comportaria un estat massa relaxat i possiblement depressiu.
- Un nivell òptim d'aquestes ones ajuda a estar molt més receptiu, enfocat a millorar la capacitat de l'individu per resoldre problemes.

Cal mencionar també que quan les persones realitzen tasques que requereixen de la memòria de treball, el cervell utilitza aquesta classe d'ones de baixa freqüència per anar canviant entre les diferents parts de la informació. És a dir, actua com una comporta reguladora que decideix quan es llegeix o es descarta la informació emmagatzemada breument en la memòria de treball, i que permet poder pensar en una altra cosa.

### 5. Ones Gamma (25 a 100 Hz)

Resulta molt difícil captar-la en los electroencefalogrames. Es parla d'un tipus d'ona que s'origina en el tàlem i es mou des de la part posterior del cervell cap endavant i a una velocitat increïble.

- Es relaciona amb tasques d'un alt processament cognitiu.
- Té a veure amb l'estil d'aprenentatge, amb la capacitat d'assentar informació nova i també amb els sentits i percepcions.
- Es coneix, per exemple, que les persones amb problemes mentals o d'aprenentatge tendeixen a tenir una activitat a la ona gamma menor que la mitjana.
- Els estats de felicitat evidencien també pics elevats en aquest tipus d'ona.
- La fase REM del son també sol caracteritzar-se per una alta activitat d'aquest rang de freqüències.

Per observar una representació de les diferents ones es pot observar la següent imatge:

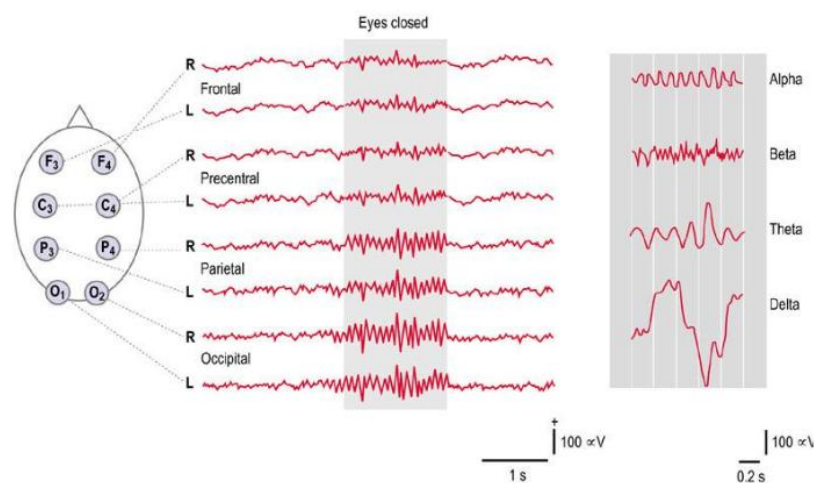


Figura 2: Exemple de les diferents ones dins d'un EEG.

## 8.2. Dataset: BCI competition III Dataset IIIa

Aquest conjunt de dades són d'imatges motoritzades (multicèl·lula) amb 4 classes (mà esquerra, mà dreta, peu, llengua) proporcionades pel Laboratori d'Interfícies Brain-Computer (BCI-Lab) de la Universitat de Tecnologia de Graz (Gert Pfurtscheller), Alois Schlögl).

### 8.2.1. Desenvolupament de l'experiment

Per obtenir el conjunt de dades es va realitzar un experiment sobre 3 subjectes. L'objectiu era captar imatges que reflecteixin els moviments mentals de moure la mà esquerra, la mà dreta, el peu o de la llengua segons una indicació, amb ordre aleatori. Durant l'experiment el subjecte havia de seure a una cadira amb reposa braços.

L'experiment consistia en la realització de diverses tirades (mínim 6) amb 40 proves cadascuna. Un cop que comença la prova, els primers 2 segons eren de "relax", sense cap estímul extern. Als 2 segons, un estímul acústic indicava el començament de la prova i es mostrava una creu "+". A  $t = 3$  es mostrava una fletxa cap a l'esquerra, dreta, amunt o avall durant 1s i al mateix temps se li demanava al subjecte d'imaginar un moviment específic segons la direcció de la fletxa (mà esquerra, mà dreta, llengua o peu, respectivament) fins que desaparegués la creu a  $t = 7$ s.

Cadascuna de les 4 possibilitats es mostrava 10 vegades en cada execució en un ordre aleatori (fent un total de 40 proves, tal com s'ha especificat abans).

El resum gràfic d'un únic assaig es mostra a la figura següent.

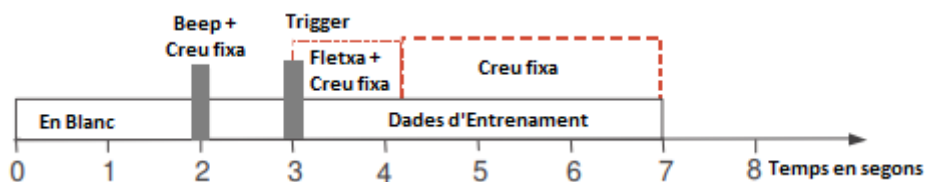


Figura 3: Temps de la mostra

### 8.2.2. Format del dataset

Les dades es guarden en format GDF i poden ser carregades a Matlab o Octave amb Biosigtoolbox (versió 0.81 o posteriors) utilitzant la instrucció `[s,HDR] = load(filename)`. Les dades poden contenir NaNs; aquests NaNs indiquen que hi ha hagut algun problema en la transformació analògic a digital de la senyal.

El principi de cada prova ( $t = 0$ s d'acord amb la Fig. 1) pot obtenir-se a partir de HDR.TRIG.

Les etiquetes de cada classe s'emmagatzemen a HDR.Classlabel. HDR.Classlabel pot contenir els valors '1','2','3','4', i 'NaN'. Els valors '1','2','3','4' indiquen les etiquetes de cada prova, NaN indica els intents dels tests.

### 8.2.3. Informació tècnica de la presa de dades

La presa de dades es va fer amb un amplificador EEG de 64 canals des del neuroescàner tal com es mostra a la figura 2, utilitzant el mastoide esquerre com a referència i el mastoide dret com a terra. L'EEG està mostrejat a 250 Hz i filtrat entre 1 i 50Hz amb Notchfilter. Les dades de totes les execucions estan concatenades i convertides a format GDF.



Cada conjunt de dades té un període d'entrenament durant el qual el subjecte realitza les imatges del motor previstes.

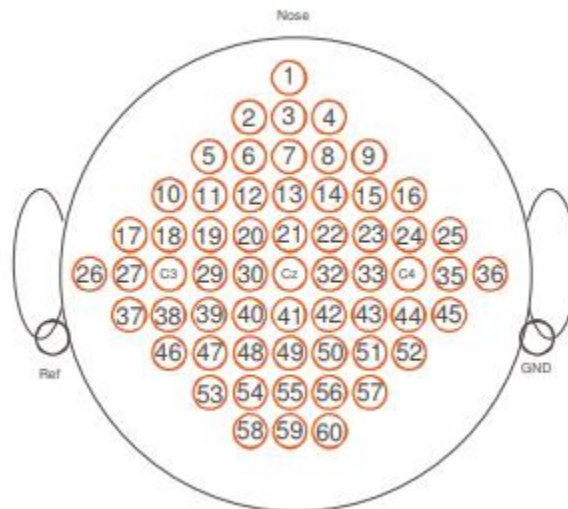


Figura 4: Posició dels elèctrodes

### 8.3. Biosignal tool

Biosig va ser creat al 2003 i és un software de codi lliure i gratuït que conté una llibreria de processament de senyals biomèdics. Els primers llançaments es basaven exclusivament en el codi M per a Matlab i Octave, encara que posteriorment es va desenvolupar una versió per C/C++.

Actualment, BioSig proporciona una interfície comuna compatible amb uns 50 formats de dades diferents. Els formats de dades més habituals en la investigació BCI són BCI2000, el format de dades BioSemi (BDF), el format de dades general (GDF), BrainVision i diversos tipus de fitxers Matlab.

Per facilitar el suport de diferents formats de dades, BioSig defineix una estructura de dades comuna per emmagatzemar la informació de capçalera, del fitxer i defineix codis estandarditzats per a esdeveniments. Aquesta estructura es defineix a la taula següent:

General Information-Fixed Header	
TYPE	Format del fitxer
VERSION	Número de versió
FileName	Nom del fitxer
FILE.OPEN	0: tancat, 1: mode lectura, 2: mode escriptura
FILE.POS	Position of file pointer in data segments or samples
T0	Date and time of start of recording
NS	Número de senyals (channels)
SampleRate	Taxa de mostreig
NRec	Number of records (data blocks)
SPR	Mostres per registre

Dur	Duration of each record
<b>Subject-Specific Information</b>	
Identification	Identificació única del pacient
VERSION	Número de versió
FileName	Nom del fitxer
FILE.OPEN	0: tancat, 1: mode lectura, 2: mode escriptura
FILE.POS	Position of file pointer in data segments or samples
T0	Data i temps d'inici
NS	Número de senyals (channels)
SampleRate	Taxa de mostreig
NRec	Número de mostres (data blocks)
SPR	Mostres per registre
Dur	Durada de cada mostra
Sex	Genere
Birthday	data d'aniversari
Handedness	0:desconegut, 1: dret, 2: esquerra, 3: igual
Medication	Nom de la medicació
Height	Alçada en cm.
Weight	Pes en kg
...	Altra informació del pacient
<b>Channel-specific header</b>	
Cal	Escala, guany, factor de calibració
Off	Offset
DigMin, DigMax	Rang de mesura (digital domain)
PhysMin, PhysMax	Rang de mesura (physical domain) PhysMin = DigMin*Cal + Off PhysMax = DigMax*Cal + Off
Label	Descripció del canal
LeadIdCode	Codi estàndard per descriure el canal (ISO11073:10101/IEEE 1073)
PhysDim	Dimesió física (text)
PhysDimCode	Dimesió física codificada (ISO11073:10101/IEEE 1073)
Lowpass	Upper edge frequency
Highpass	Lower edge frequency
Notch	Filtre Notch Off/50Hz/60Hz
GDFTYP	Tipus de dada (3: int16, 5: int32, 16: float32, 17: float64)
SPR	Número de mostres per registre
<b>Event Table</b>	
POS	posició inicial d'un esdeveniment

TYP	Tipus d'esdeveniment segons la traula predefinida (1-255 és pels esdeveniment definits per l'usuari)
DUR	Durada de l'esdeveniment (opcional)
CHN	0: Cap canal específic, altre valor: esdeveniment relacionat amb el canal específic
CodeDesc	Descripció de la llista específica de codis d'esdeveniment de l'usuari

Taula 1: Camps de l'estructura de capçalera comuna a BioSignal.

L'eina BioSignal disposa d'un seguit de mètodes de processat de senyal aplicables a senyals biomèdiques, es per aquest motiu que es tan útil. Tots els components que inclou l'eina estan a la taula següent:

Component	Key features
Common interface	<ul style="list-style-type: none"> <li>load loads whole data files</li> <li>sopen, sread, and sclose load segments of data</li> <li>Common header structure</li> </ul>
Data preprocessing	<ul style="list-style-type: none"> <li>Triggering, partitioning of data</li> <li>Artifact processing</li> <li>Quality check of data through histogram analysis</li> <li>Spatial filters</li> <li>Detection of EMG artifacts</li> <li>Common spatial patterns</li> </ul>
Feature extraction	<ul style="list-style-type: none"> <li>Band power</li> <li>Adaptive autoregressive parameters</li> <li>Adaptive multivariate autoregressive parameters</li> <li>(Adaptive) Hjorth</li> <li>(Adaptive) Barlow</li> <li>(Adaptive) Wackermann</li> <li>(Adaptive) time-domain parameters</li> <li>Adaptive brain rate, spectral edge frequency</li> </ul>
Feature classification	<ul style="list-style-type: none"> <li>Linear discriminant analysis (LDA)</li> <li>Quadratic discriminant analysis (QDA)</li> <li>Support vector machines</li> <li>Naïve Bayesian classifier (NBC)</li> <li>Augmented NBC</li> <li>Sparse LDA</li> <li>Generalized discriminant analysis</li> </ul>
Evaluation criteria	<ul style="list-style-type: none"> <li>Classification accuracy</li> <li>Cohen's kappa coefficient</li> <li>Receiver operating characteristics (ROC)</li> <li>Area under the ROC curve</li> <li>Mutual information, information transfer rate</li> <li>Correlation coefficient</li> </ul>
Metafunctions	<ul style="list-style-type: none"> <li>findclassifier, cross-validation, standardized analysis</li> <li>demo2 (example of a standardized offline analysis)</li> </ul>
Real-time data processing	<ul style="list-style-type: none"> <li>Based on Matlab/Simulink and the Real-Time Workshop</li> <li>Provides standardized online BCI experiment</li> </ul>
SigViewer	<ul style="list-style-type: none"> <li>Stand-alone viewing and scoring software</li> <li>C++ and Qt</li> </ul>

Taula 2: Components de la llibreria BioSignal

Com es pot veure a la taula, existeixen tant funcions de lectura de fitxers per obtenir l'estructura descrita amb anterioritat com funcions de processat, oferint, per exemple, diverses eines detectar artefactes en les gravacions EEG. També es disposen de diferents eines per extracció de característiques com per exemple AAR i mètodes de classificació per característiques com diferents versions de LDA i MDA i SVM.

A més de oferir funcions de processat i classificació també disposa de mètodes de avaluació com per exemple el càlcul del coeficient kappa.

# 9. Arquitectura

El sistema disposa d'un mòdul de preprocessat del senyal procedent del dataset, una etapa d'extracció de característiques i una etapa de classificació.

## 9.1. Preprocessat

Els senyals obtinguts són generalment molt sorollosos i contenen molts defectes incloent el soroll d'alta freqüència degut a interferències elèctriques i defectes fisiològics. A part del soroll, hi ha el problema de la conducció de volum a l'EEG a causa de la distància entre el cuir cabellut i les neurones, cosa que dificulta fixar el punt exacte on es duu a terme l'activació. Per tant, el preprocessat d'aquests senyals és un pas necessari.

Com a preprocessat s'aplica un filtre passabanda i la reducció espacial CSP.

### 9.1.1. Filtre passabanda

Un filtre passabanda és un dispositiu o circuit electrònic que permet passar senyals entre dues freqüències específiques i bloqueja senyals en altres freqüències. Se sap que, la majoria de les activitats del cervell relacionades amb les imatges motor es troben dins de la banda de freqüències de 8-33 Hz. Així doncs, s'utilitza un filtre passabanda pel rang de freqüències 8-33 Hz. El filtre passabanda filtra la major part del soroll d'alta freqüència. El filtre es realitza amb dues sub-bandes de 8-13Hz i de 12-33Hz. L'elecció de les sub-bandes es realitza pel fet que les ones cerebrals Alfa i Beta es troben dins d'aquestes bandes de freqüència.

### 9.1.2. Common Spatial Patterns

Els dispositius EEG son dispositius multicanal, però per a la classificació no són necessaris tots els canals. Per reduir els canals s'utilitza un filtre espacial.

El filtre espacial redueix el nombre de canals de manera que els senyals resultants siguin fàcils de classificar. Els Common Spatial Patterns (CSP) són un dels filtres espacials més utilitzats en la construcció de BCI per dues classes ja que maximitza la variància d'una classe mentre la minimitza per a l'altra classe. Els filtres CSP permeten representar als diferents canals per veure les activacions de diverses regions del cervell.

Degut a que es disposa d'un escenari amb 4 classes, s'utilitzen 4 filtres CSP.

## 9.2. Extracció de característiques

Es descriuran totes les característiques que s'extrauran dels senyals EEG per tal de reduir les dimensions de les dades i poder utilitzar aquestes característiques per tal de realitzar la classificació posterior.

### 9.2.1. Característiques en el domini freqüencial

Les característiques del domini de la freqüència capturen la informació basada en la freqüència dels ritmes cerebrals durant la presa d'imatges. Hi ha ritmes cerebrals en bandes de freqüències específiques on es produeixen la majoria de les activitats neuronals.

Per analitzar les característiques de les senyals EEG es pot realitzar la densitat espectral de potència (PSD). El PSD es calcula mitjançant la transformada de Fourier dels senyals de domini del temps en un domini de freqüència mitjançant mètodes no paramètrics. Un d'aquests mètodes es coneix com el mètode de Welch. Per a cada canal es calcula la potència de banda del senyal transformat de Fourier:

$$bandpower_{ch=i} = \log(1 + X^2)$$

on X és la senyal del domini de freqüència ch = i representa el canal específic on es troba la banda de potència.

Per obtenir aquesta característica s'ha utilitzat el codi següent:

```
bands = [8,13;12,33]; % definició de les bandes de freqüència
win = 1; % longitud en segons de la finestra de suavitzat (smoothing window)
s1 = s;
s1(isnan(s1))=0;
f3 = bandpower(s1, HDR.SampleRate, bands, win);
```

En aquest cas, el filtre utilitzat per realitzar la banda de potència és un filtre Butterworth passabanda. On f3 es una característica per introduir al classificador.

### 9.2.2. Coeficients Autoregressius (AR)

El mètode Autoregressiu (AR) modela el senyal en un moment donat com a una suma ponderada de senyals en un temps anterior i amb alguns sorolls.

L'algorisme de Burg s'utilitza per trobar els coeficients autoregressius. Els coeficients es van utilitzar com a característiques de l'anàlisi. Matemàticament es pot formular com:

$$X(t) = a_1X(t_1) + a_2X(t_2) + \dots + a_pX(t_p) + E_t$$

On, X(t) es la mesura del senyal en temps t, Et és el soroll, a1 fins a ap són els paràmetres autoregressius i p és l'ordre del model. Els coeficients autoregressius del model proporcionen les característiques d'aquest paradigma.

Per obtenir aquesta característica s'ha utilitzat el codi següent:

```
p = 9; % Ordre del model AAR
MODE.MOP = [0,p,0]; % Ordre del model AAR
MODE.UC = 0.0085; % coeficient del model AAR

fprintf(1, '\3a: TDP (log(Hjorth)).\n');
f1 = tdp(s(:, eegchan), p, MODE.UC);

fprintf(1, '\3b: Adaptive Autoregressive parameters (AAR).\n');
f2 = [];
for ch = 1:length(eegchan),
    X = tvaar(s(:, eegchan(ch)), MODE.MOP, MODE.UC);
    X = tvaar(s(:, eegchan(ch)), X); % estimació AAR
    f2 = [f2, X.AAR, log(X.PE)];
```

```

end;

fprintf(1, '\3c: Differential Adaptive Autoregressive parameters (AAR).\n');
f4 = [];
ds = diff(s);
for ch = 1:length(eegchan),
    X = tvaar(ds(:, eegchan(ch)), MODE.MOP, MODE.UC);
    X = tvaar(ds(:, eegchan(ch)), X); % estimació AAR
    f4 = [f4, X.AAR, log(X.PE)];
end;

```

On es calculen 3 paràmetres autoregressius: TDP, AAR y el DAAR que seran 3 característiques més per classificar el senyal.

### 9.3. Classificació

Un classificador és un algoritme utilitzat per a assignar un element entrant no etiquetat en una categoria concreta coneguda. Aquest algoritme, permet ordenar o disposar per classes elements entrants, a partir de certa informació característica d'aquests.

Una manera d'implementar un classificador és seleccionar un conjunt d'exemples etiquetats i intentar definir una regla que permeti assignar una etiqueta a qualsevol altra dada d'entrada.

S'han utilitzat 2 classificadors: SVM i LDA.

Per realitzar la classificació s'ha utilitzat el codi següent, únicament definint inicialment el mètode de classificació utilitzat:

```

TYPE.TYPE % es defineix el tipus de classificador

CC1 = findclassifier(f1, HDR.TRIG, [HDR.Classlabel,NG], MODE, [], TYPE);
CC2 = findclassifier(f2, HDR.TRIG, [HDR.Classlabel,NG], MODE, [], TYPE);
CC3 = findclassifier(f3, HDR.TRIG, [HDR.Classlabel,NG], MODE, [], TYPE);
CC4 = findclassifier(f4, HDR.TRIG, [HDR.Classlabel,NG], MODE, [], TYPE);

d1 = [ones(size(f1,1),1), f1]*CC1.weights; % s'obté el resultat
d2 = [ones(size(f2,1),1), f2]*CC2.weights; % s'obté el resultat
d3 = [ones(size(f3,1),1), f3]*CC3.weights; % s'obté el resultat
d4 = [ones(size(f4,1),1), f4]*CC4.weights; % s'obté el resultat

```

#### 9.3.1. SVM

Les màquines de suport vectorial, màquines de vectors de suport o màquines de vector suport (Support Vector Machines, SVMs) són un conjunt d'algoritmes d'aprenentatge supervisat.

Aquests mètodes estan relacionats amb problemes de classificació i regressió. Donat un conjunt de mostres d'entrenament es poden etiquetar les classes i entrenar una SVM per a construir un model que predigui la classe d'una nova mostra. Intuïtivament, una SVM és un model que representa els punts de mostra en l'espai, separant les classes a 2 espais el més amplis possibles mitjançant un hiperplà de separació definit com el vector entre els 2 punts, de les 2 classes, més properes al que s'anomena vector suport. Quan les noves mostres es posin en correspondència amb el model en qüestió, en funció dels espais als que pertanyi, poden ser classificades a una o altra classe.

Dit d'una altra manera, una SVM construeix un hiperplà o conjunt d'hiperplans en un espai de dimensionalitat molt alta (o fins i tot infinita) que pot ser utilitzat en problemes de classificació o regressió. Una bona separació entre les classes permetrà una classificació correcta.

Per obtenir aquesta classificació s'ha utilitzat:

```
TYPE.TYPE = 'SVM';
```

### **9.3.1. LDA**

L'Anàlisi Discriminant Lineal o Linear Discriminant Analysis (LDA) és un mètode de classificació supervisat de variables qualitatives en el que dos o més grups són coneguts a priori i noves observacions es classifiquen en un d'ells en funció de les seves característiques. Segons el teorema de Bayes, LDA estima la probabilitat de que una observació, donat un determinat valor dels predictors, pertanyi a cada una de las classes de la variable qualitativa,  $P(Y=k|X=x)P(Y=k|X=x)$ . Finalment s'assigna l'observació a la classe k per a la que la probabilitat predita és més alta.

Per obtenir aquesta classificació s'ha utilitzat:

```
TYPE.TYPE = 'LDA';
```



## 10. Projecció a futur

En futures versions es podria realitzar un preprocessat més exhaustiu. Incloent la millora del senyal d'entrada aplicant una reducció d'artefactes EOG.

Per altra banda, per una millor classificació també es podria realitzar una extracció de característiques més extensa, incloent així alguna característica temporal que no s'ha inclòs en l'anàlisi d'aquest treball. A més es podrien afegir més característiques espacials. D'aquesta manera el senyal quedarà més caracteritzat i es podrà classificar millor. Es per aquest motiu que el processat en freqüència i el filtratge adaptatiu funcionen millorant el senyal.

# 11. Conclusió

S'ha pogut verificar que els senyals EEG són difícils de classificar, degut a la gran quantitat d'interferències que poden ocórrer durant la gravació dels senyals i que s'ha de realitzar una classificació molt focalitzada amb l'usuari en concret. El fet d'utilitzar paràmetres d'altres usuaris no es de molta utilitat, s'ha de basar l'anàlisi en el propi senyal. Això dificulta encara més el processat i la classificació.

# Annex 1. Lliurables del projecte

Llista de fitxers lliurats i la seva descripció.

- PAC1: Definició del projecte.
- PAC2: Document explicatiu de la part del programa realitzat fins al punt d'entrega. En aquest cas, es defineixen les dues primeres etapes del programa: preprocessat i extracció de característiques.
- PAC3: Memòria final. En aquest lliurable ja s'inclouen totes les parts del classificador BCI explicades per tant, s'inclou el preprocessat, l'extracció de característiques i el classificador.
- PAC4: Presentació del projecte. Document de presentació on s'explica de manera gràfica el resultat del projecte i les accions realitzades.

## Annex 2. Codi font (extractes)

```
% Step 1: Preparar les dades =====%
fprintf(1, '--> 1: Preparar les dades.\n', filename);
fprintf(1, '\n1a: Carregar el fitxer %s.\n', filename);
[s, HDR]=sload(filename);

% Step 1b: Extracció de trigger i categorització =====%
fprintf(1, '\n1b: Extracció de trigger i classlabels.\n');
%----- extracció per event table
ix = find((HDR.EVENT.TYP>hex2dec('300')) & (HDR.EVENT.TYP<hex2dec('30d'))); %
0x0300..0x03ff
[i, j, HDR.Classlabel] = unique(HDR.EVENT.TYP(ix));
HDR.TRIG = HDR.EVENT.POS(HDR.EVENT.TYP==hex2dec('300'));
%HDR.TRIG = HDR.EVENT.POS(ix);
t0 = HDR.EVENT.POS(ix);
t0 = (t0(1) - HDR.TRIG(1))/HDR.SampleRate; % temps entre inici del trial i el
final.

W = spatialfilter(s, 'CSP', eegchan);

% Step 2: Extracció de característiques =====%
eegchan = HDR.InChanSelect;
fprintf(1, 'Step 3: feature extraction.\n');
p = 9;
MODE.MOP = [0, p, 0]; % order of AAR model
MODE.UC = 0.0085; % update coefficient of AAR model

% 2a: Time-dependent parameters - motivated by log(Hjorth)
fprintf(1, '\n2a: TDP (log(Hjorth)).\n');
f1 = tdp(s(:, eegchan), p, MODE.UC);

% 2b: Adaptive Autoregressive Parameters
fprintf(1, '\n2b: Adaptive Autoregressive parameters (AAR).\n');
f2 = [];
for ch = 1:length(eegchan),
    X = tvaar(s(:, eegchan(ch)), MODE.MOP, MODE.UC);
    X = tvaar(s(:, eegchan(ch)), X); % AAR estimation
    f2 = [f2, X.AAR, log(X.PE)];
end;

fprintf(1, '\n2b: Differential Adaptive Autoregressive parameters (AAR).\n');
f4 = [];
ds = diff(s);
for ch = 1:length(eegchan),
    X = tvaar(ds(:, eegchan(ch)), MODE.MOP, MODE.UC);
    X = tvaar(ds(:, eegchan(ch)), X); % AAR estimation
    f4 = [f4, X.AAR, log(X.PE)];
end;

% 3c: bandpower
fprintf(1, '\n3c: bandpower.\n');
bands = [8, 13; 12, 33]; % definició de les bandes de freqüència
win = 1; % longitud en segons de la finestra de suavitzat (smoothing window)
s1 = s;
s1(isnan(s1))=0;
f3 = bandpower(s1, HDR.SampleRate, bands, win);
```

```

% Step 4: Classificació ===== %
fprintf(1, '--> 4: Classificació.\n');
TrialLen = 4; % segons
SegmentLen = HDR.NRec; % mostres
NoS = ceil(9*HDR.SampleRate/SegmentLen);
MODE.T = reshape((1:NoS*SegmentLen), SegmentLen, NoS)';
MODE.WIN = MODE.T(:,1) > 3*HDR.SampleRate+1; % cue @ t=3s.
MODE.t0 = t0;
MODE.t = [min(MODE.T(:)):max(MODE.T(:))]/HDR.SampleRate;
MODE.Segments = MODE.T;
MODE.Fs = HDR.SampleRate;

K = 1;
NG = ceil([1:length(HDR.Classlabel)]'/K);

%TYPE.TYPE = 'LDA'; % Classificador
TYPE.TYPE = 'SVM';
% Altres classificadors: MDA, LD2, LD3, LD4, RDA, GDBC, SVM, RBF, NBC, aNBC,
LDA/GSVD, MDA/GSVD, LDA/sparse, MDA/sparse

% Trobar el millor segment utilitzant cross-validation
CC1 = findclassifier(f1, HDR.TRIG, [HDR.Classlabel, NG], MODE, [], TYPE);
CC2 = findclassifier(f2, HDR.TRIG, [HDR.Classlabel, NG], MODE, [], TYPE);
CC3 = findclassifier(f3, HDR.TRIG, [HDR.Classlabel, NG], MODE, [], TYPE);
CC4 = findclassifier(f4, HDR.TRIG, [HDR.Classlabel, NG], MODE, [], TYPE);

d1 = [ones(size(f1,1),1), f1]*CC1.weights;
d2 = [ones(size(f2,1),1), f2]*CC2.weights;
d3 = [ones(size(f3,1),1), f3]*CC3.weights;
d4 = [ones(size(f4,1),1), f4]*CC4.weights;

```

## Annex 9. Bibliografia

<https://blog.sincrolab.es/2017/03/20/que-es-y-que-podemos-medir-con-el-electroencefalograma-eeg/>

<https://lamenteesmaravillosa.com/tipos-de-ondas-cerebrales/>

<http://www.ub.edu/pa1/node/130>

<http://biosig.sf.net>

[www.bci2000.org](http://www.bci2000.org)

<http://arxiv.org/abs/cs.DB/0608052>