

Il ruolo del phase lag index nella classificazione multiclasse di un segnale elettroencefalografico

Università di Verona
Corso di Laurea in Informatica
05/10/2022

Relatore: Prof.ssa Silvia Francesca Storti
Correlatore: Ing. Ilaria Siviero

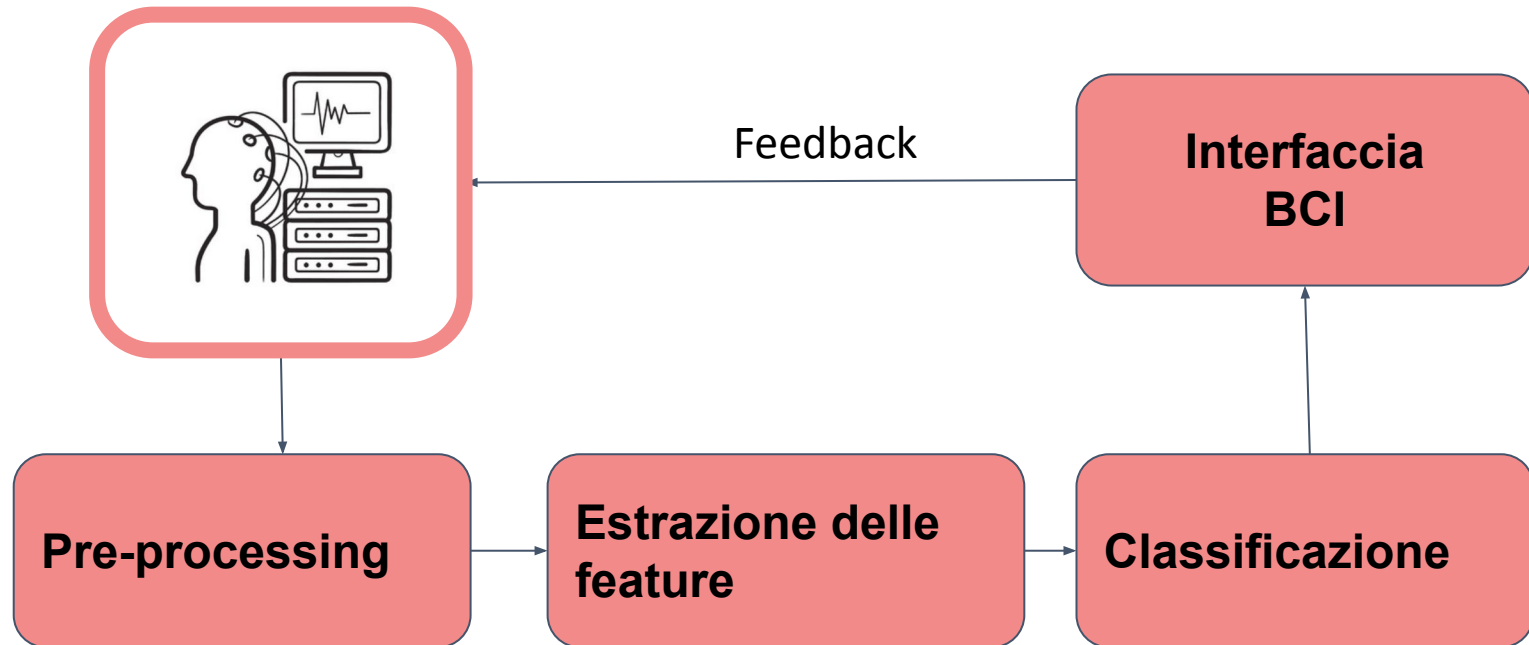
Giulia Doro
VR414426

Introduzione

Un'**interfaccia cervello computer (BCI)** è un sistema di comunicazione diretta tra il cervello e il computer.

Frequentemente utilizzate sono le BCI di tipo **non invasivo**, basate sull'**elettroencefalogramma (EEG)**. L'EEG viene registrato tramite **elettrodi** posti sullo scalpo che registrano l'attività sincrona di gruppi di cellule.

L'**immaginato motorio (MI)** è un processo in cui il soggetto immagina un movimento senza l'esecuzione fisica.



Stato dell'arte

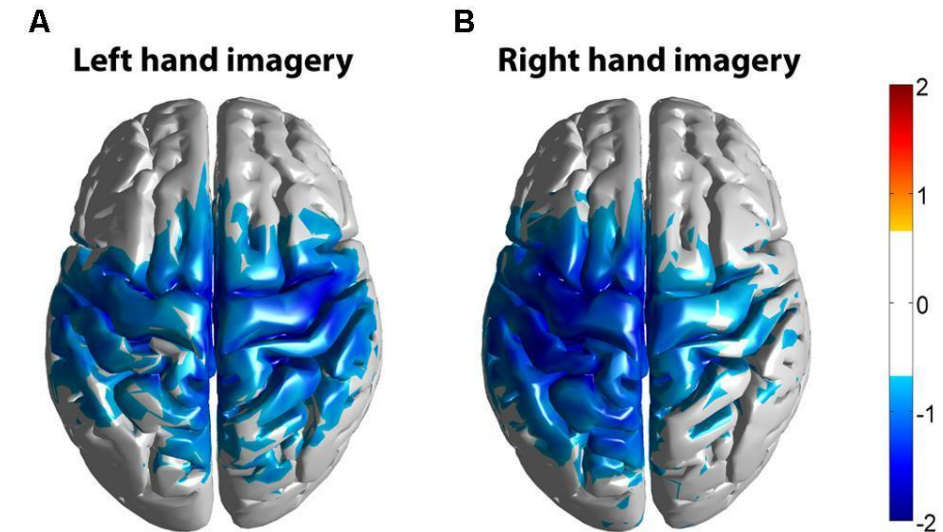
Numerose ricerche utilizzano vari metodi di analisi per l'estrazione delle feature come **Common Spatial Pattern (CSP)** e **Power Spectral Density (PSD)** per ottenere i migliori risultati nel classificare movimenti immaginati.

Si possono così implementare **BCI** basate su segnali EEG sempre più efficienti per aiutare le persone che non possono comunicare con il mondo esterno attraverso comuni canali di comunicazione.

Lo scopo della ricerca è quello di **classificare** tra diversi tipi di task di tipo immaginato motorio in soggetti sani tramite misure di connettività come **Phase Lag Index (PLI)**.

Si utilizza:

- PLI
- Classificatore multiclasse e binario



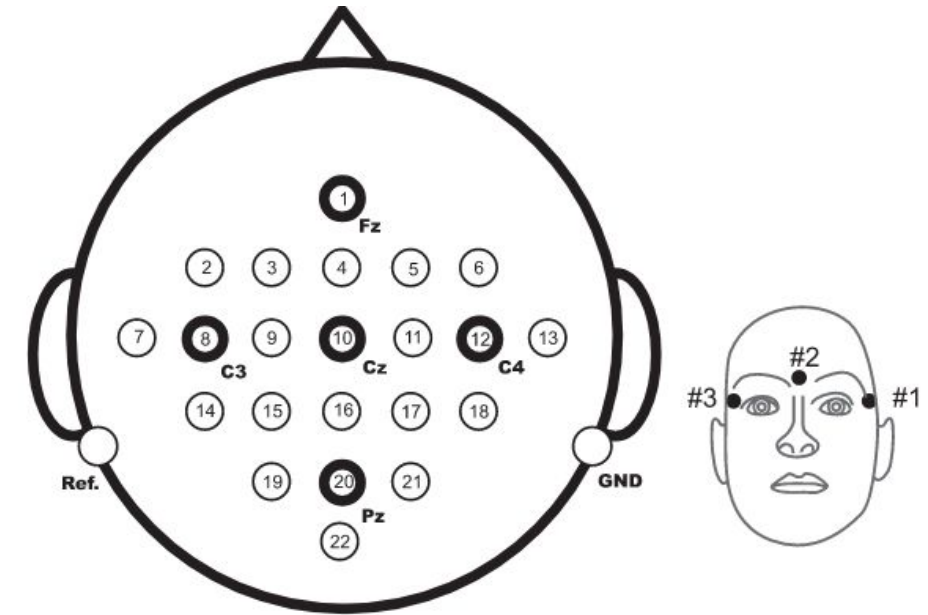
Atlan et. Al, *National Defence University*, 2019

Feng et Al., *SMC*, 2020
Boernama et Al., *AIMS*, 2021
Kim et Al., *MBEC*, 2018

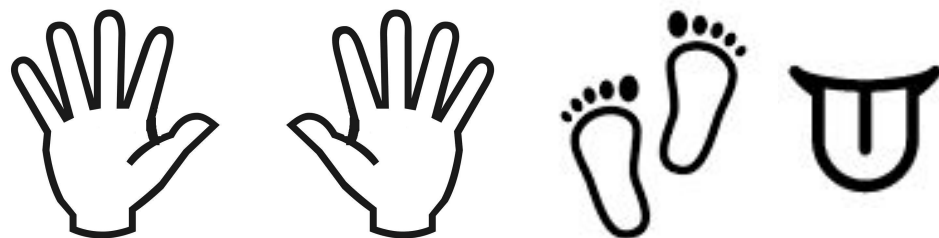
Materiali e Metodi: data set

Si utilizza il data set 2a della “**BCI Competition IV 2008**”.
Le caratteristiche principali sono:

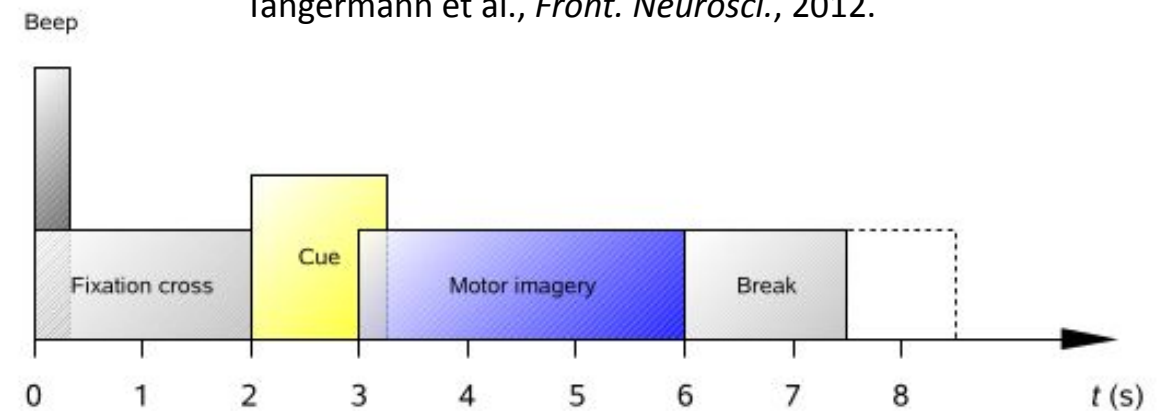
- 9 soggetti
- 4 MI: mano sinistra, mano destra, entrambi i piedi e lingua
- 22 canali EEG e 3 EOG
- frequenza di campionamento: 250 Hz
- 72 trial per ogni classe



Tangemann et al., *Front. Neurosci.*, 2012.



MI TASK



Tangemann et al., *Front. Neurosci.*, 2012.

Materiali e Metodi: pre-elaborazione

Rimozione della **baseline**

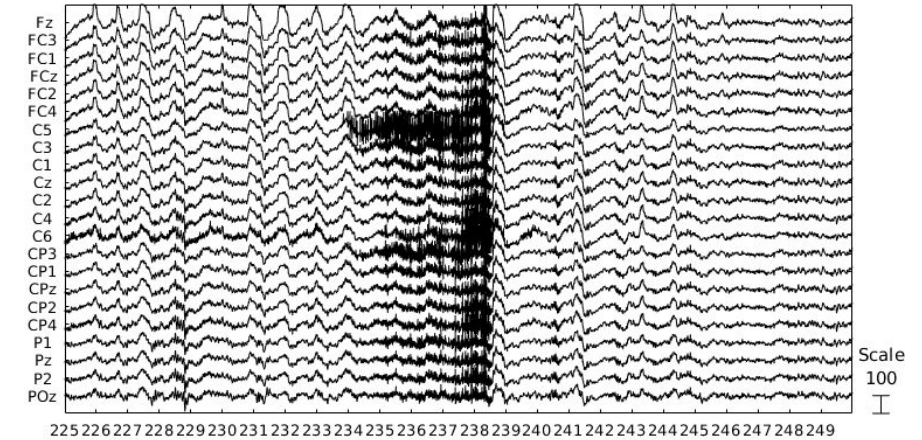
Filtro di Butterworth:

- Filtro di tipo Infinite Impulse Response (IIR)
- Fase nulla
- Assenza di ripple
- Applicato come passa alto e passa basso

Re-referenziazione dei dati rispetto all' **average reference**

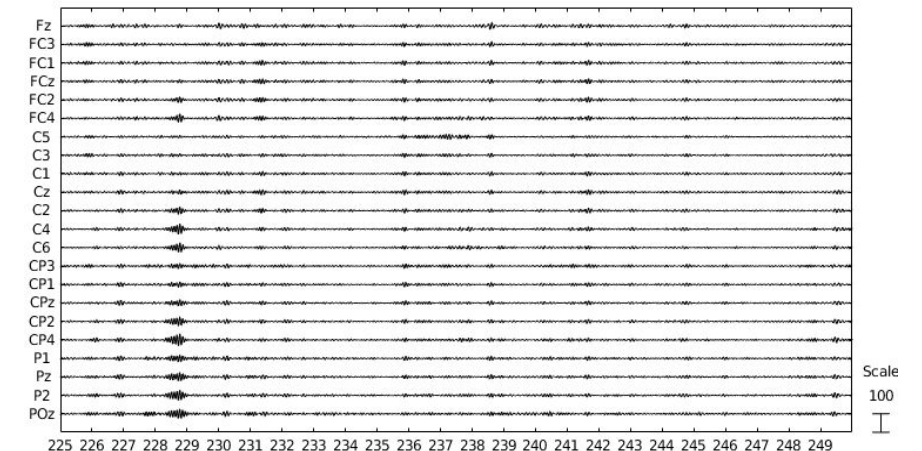
Riorganizzazione degli eventi e rimozione dei **trial artefattati**

EEG GREZZO



Filtraggio nella banda alpha : 8 - 12 Hz
Filtraggio nella banda beta: 13 - 30 Hz

EEG FILTRATO



Materiali e Metodi: estrazione delle feature

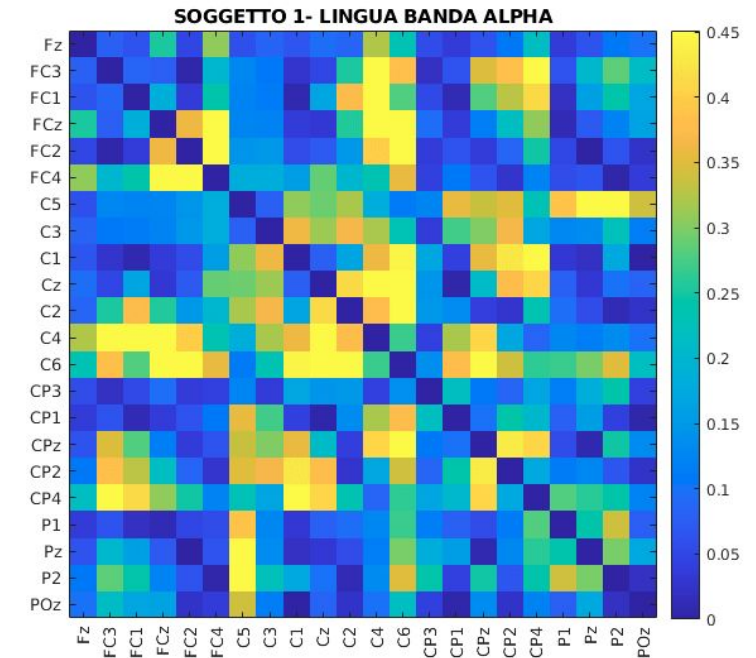
Calcolo del **PLI** per ogni classe e per ogni “trial” privo di artefatti.

Calcolo della **trasformata di Hilbert** per ottenere la fase istantanea dei segnali utile al calcolo del PLI.

Il PLI è un indice che stima la connettività del segnale EEG e misura in particolare la sincronizzazione tra due elettrodi distinti.

$$PLI = \left| \frac{1}{N} \sum_{k=1}^N \text{sign}(\sin(\phi_l(k) - \phi_m(k))) \right| \in [0, 1]$$

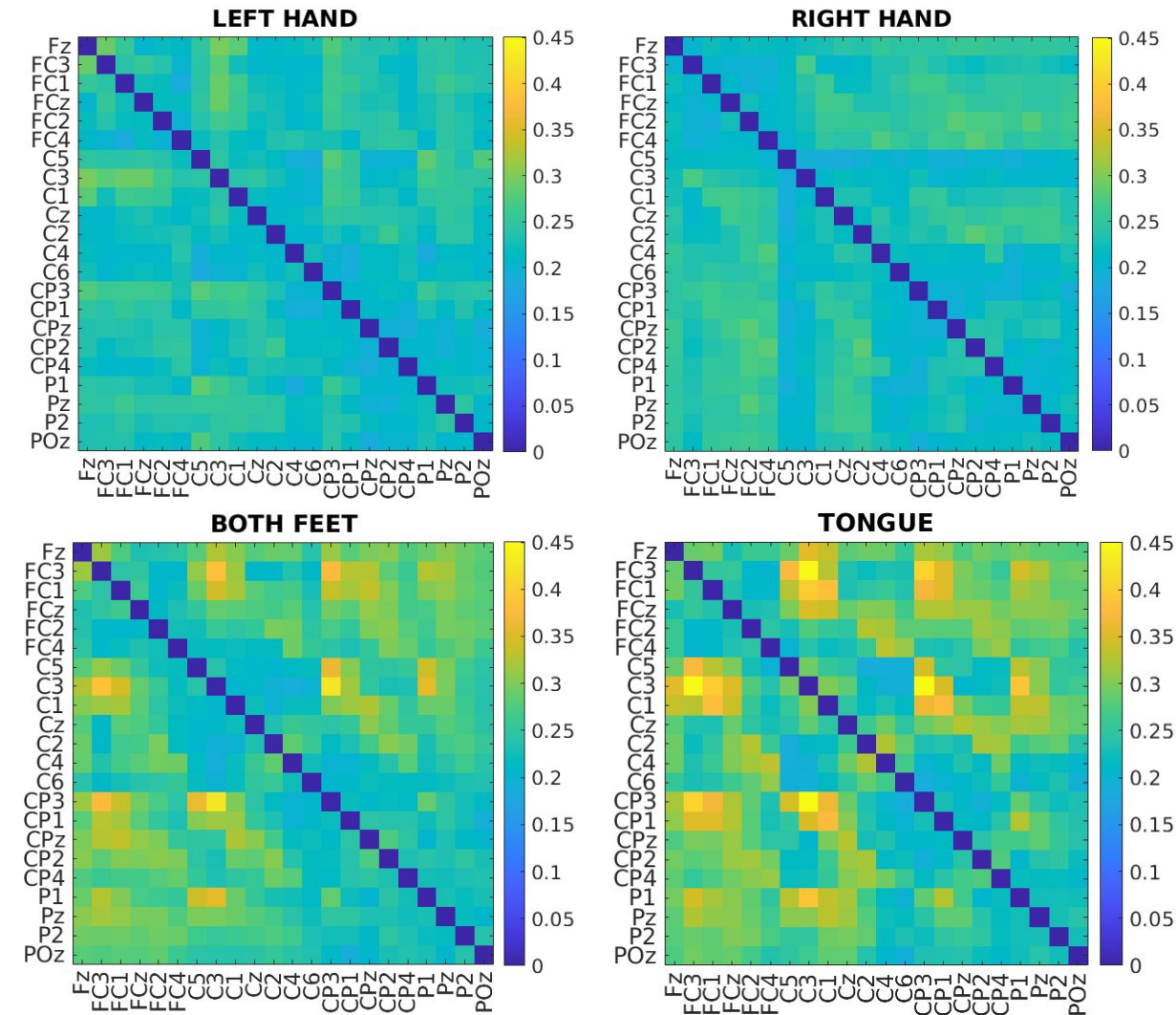
Viene misurato sia per la banda alpha sia per la banda beta



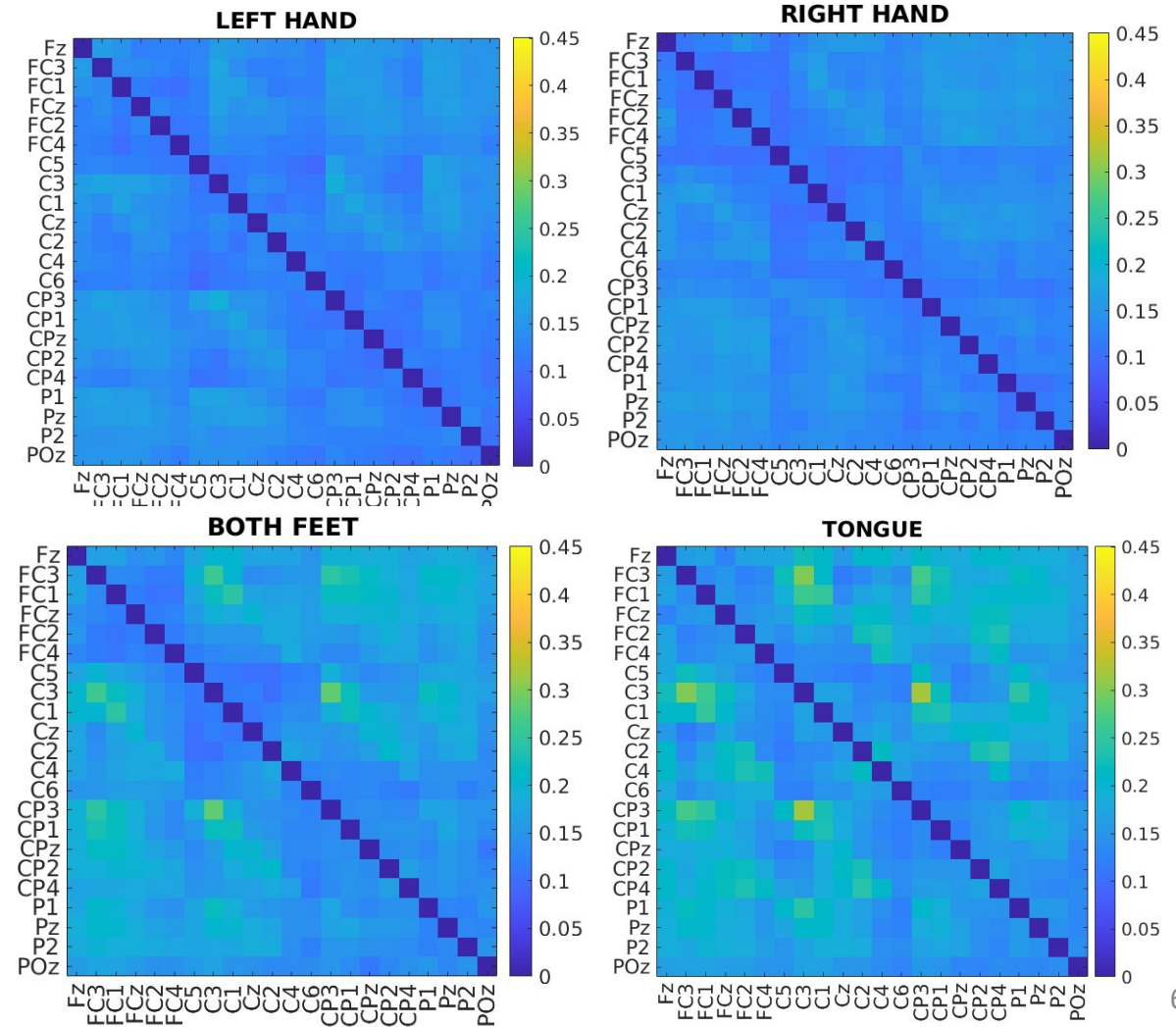
Esempio di matrice di connettività di dimensione 22x22

Materiali e Metodi: estrazione delle feature

PLI medio nelle 4 classi banda alpha Soggetto 1



PLI medio nelle 4 classi banda beta Soggetto 1



Materiali e Metodi: classificazione

Classificazione **bianaria** e **multiclasse**

Partizionamento per entrambe le bande:

- 80% training
- 20% test

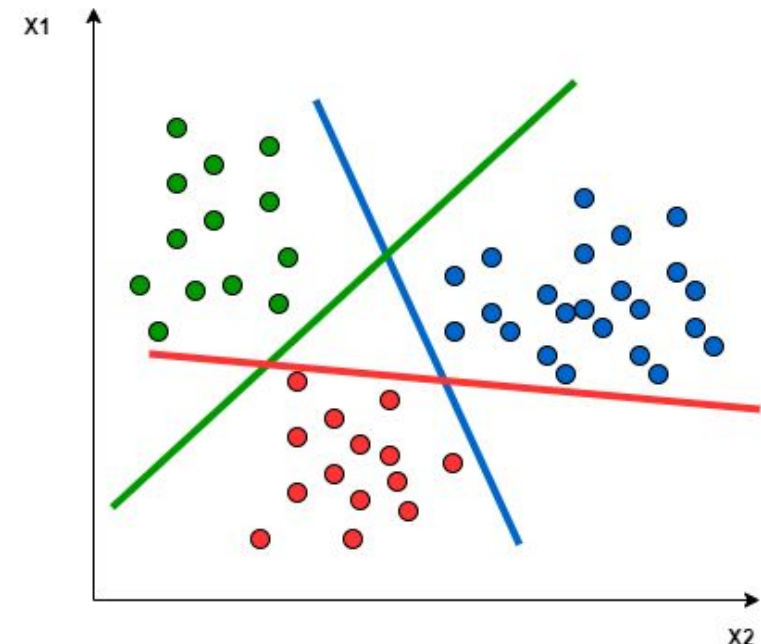
Si utilizza il metodo di classificazione “**Support Vector Machines**” (SVM) :

- Kernel lineare
- Cross validation del training set: K-fold = 5
- Training con iperparametri ottimizzati

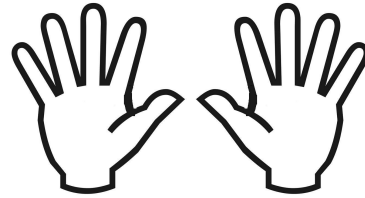
Valutazione dei risultati tramite il calcolo dell’**Accuracy**

Solo per la classificazione **multiclasse** si è eseguito il calcolo del **Fisher Score** per applicare una “**feature selection**”

$$F(i) = \frac{\sum_{j=1}^3 (x_{avg,i}^j - x_{avg,i})^2}{\sum_{j=1}^3 (\frac{1}{n_j - 1}) \sum_{k=1}^{n_j} (x_{k,i}^j - x_{avg,i}^j)^2}$$



Risultati: classificazione binaria



RISULTATI NELLA BANDA ALPHA		
SOGGETTO	ACCURACY VALIDATION	ACCURACY TEST
A01T	51%	70%
A02T	59%	56%
A03T	62%	59%
A04T	52%	52%
A05T	51%	52%
A06T	51%	50%
A07T	60%	58%
A08T	56%	65%
A09T	54%	57%

RISULTATI NELLA BANDA BETA		
SOGGETTO	ACCURACY VALIDATION	ACCURACY TEST
A01T	54%	56%
A02T	61%	41%
A03T	68%	52%
A04T	52%	52%
A05T	67%	44%
A06T	57%	45%
A07T	53%	62%
A08T	60%	54%
A09T	54%	57%

Risultati: classificazione multiclasse

RISULTATI NELLA BANDA ALPHA			
SOGGETTO	FEATURE	ACCURACY VALIDATION	ACCURACY TEST
A01T	50	42%	46%
A02T	231	27%	27%
A03T	231	34%	37%
A04T	100	29%	36%
A05T	175	31%	35%
A06T	160	32%	35%
A07T	70	40%	35%
A08T	175	32%	35%
A09T	100	34%	32%

RISULTATI NELLA BANDA BETA			
SOGGETTO	FEATURE	ACCURACY VALIDATION	ACCURACY TEST
A01T	50	38%	46%
A02T	120	30%	37%
A03T	175	44%	35%
A04T	120	34%	31%
A05T	160	30%	29%
A06T	40	35%	33%
A07T	40	37%	37%
A08T	50	39%	36%
A09T	100	42%	40%

Discussione

- Il classificatore **binario** è più efficiente rispetto a quello multiclasse
- Il classificatore **multiclasse** non ottiene i livelli di accuracy sperati. I risultati non sono necessariamente imputabili al calcolo del PLI, ma bisogna tenere presente la difficoltà nel classificare l'immaginato motorio di task mappati in elettrodi molto vicini tra loro.
- L'**accuracy aumenta** se si sceglie una soglia ad hoc per ogni soggetto e per ogni banda.
- L'utilizzo del **kernel lineare** ha prestazioni migliori rispetto a quello **gaussiano**

Risultati del classificatore multiclasse:

BANDA ALPHA		
METODO	ACCURACY VALIDATION	ACCURACY TEST
SOGLIA MEDIA	37%	29%
FEATURE SELECTION AD HOC	33%	35%

BANDA BETA		
METODO	ACCURACY VALIDATION	ACCURACY TEST
SOGLIA MEDIA	39%	29%
FEATURE SELECTION AD HOC	37%	36%

In futuro si suggerisce di utilizzare il **Fisher Score** anche nella classificazione binaria.