EEG Signal - Documentation

# Segnali EEG

I segnali EEG prodotti dal cervello ed acquisiti tramite elettrodi posizionati sulla testa sono vari e vengono classificati in base alla loro frequenza:

* Delta **(δ)**:
  + frequenza 0.5~4Hz;
  + rilevate durante il sonno profondo;
  + l’ampiezza di questo segnale varia tra i 75 e i 200 µV, incrementando all’aumentare della profondità del sonno;
* Theta **(θ)**:
  + frequenza 4~8Hz;
  + rilevate durante fasi di concentrazione silenziosa;
  + l’ampiezza di questo segnale è solitamente <100µV;
* Alpha **(α)**:
  + frequenza 8-14Hz;
  + rilevate durante fasi di rilassamento, con occhi chiusi;
  + l’ampiezza di questo segnale varia tra i 10 ed i 20 µV;
* Beta **(β)**:
  + frequenza 14~30Hz:
  + rilevate durante fasi di all’erta o di ragionamento;
  + l’ampiezza di questo segnale è <10µV;
* Gamma **(γ)**:
  + frequenza >30Hz;
  + rilevate durante fasi di elaborazione delle informazioni. Possono essere stimolate tramite attività visiva (vedi VEP).
  + l’ampiezza di questo segnale è <2µV;

[Informazioni ottenute da: <https://ietresearch.onlinelibrary.wiley.com/doi/epdf/10.1049/iet-bmt.2014.0040>]

# Preprocessing

Durante l’acquisizione di un segnale EEG possono esserci diversi fattori che portano alla contaminazione dei dati ottenuti. È dunque necessario effettuare un’attività di preprocessing adeguata al fine di eliminare i dati non necessari. Ad esempio è possibile che durante l’acquisizione, alcuni *elettrodi* non siano posizionati in modo corretto, oppure che captino segnali non necessariamente provenienti dal cervello. Tali dati non desiderati rappresentano il rumore. Il preprocessing finalizzato all’eliminazione di rumore può riguardare quattro categorie:

1. **rimozione dei canali sporchi (bad channels):** durante un’acquisizione di segnale EEG si ottengono diversi valori, provenienti da diverse fonti, detti canali. Ogni canale viaggia su diverse frequenze. Come anticipato, accade talvolta che i segnali abbiano valori palesemente errati (troppo alti o troppo bassi). Ciò è dovuto al malfunzionamento di elettrodi o al cattivo posizionamento degli stessi. Per evitare di lavorare su dati errati, è necessario rimuovere questi canali;
2. **interpolazione:** per interpolazione si intende la predizione di un valore basata sui valori che si hanno a disposizione. Una volta identificato e rimosso un canale con dati errati, vorremmo avere comunque a disposizione dei dati corretti per eseguire successivamente delle analisi. È possibile predirre i valori corretti di un canale errato a partire dai dati degli altri canali (corretti). In tal modo è anche possibile generare nuove *features*, ovvero proprietà, che potrebbero non venire rilevate durante l’acquisizione dei dati o non avere abbastanza dati per essere utilizzate;
3. **filtering:** un canale potrebbe non essere totalmente errato ma presentare solo alcuni valori che si discostano palesemente dall’andamento naturale del segnale. In tal caso è possibile applicare filtri *passa-alto*, *passa-basso* e *passa-banda* per ignorare gli intervalli indesiderati. Un’ulteriore tecnica di filtering utilizzata durante il preprocessing dei segnali EEG è il cosiddetto *notch filter* che, piuttosto che agire sui valori assunti dal segnale, agisce su una frequenza (o intervalli di frequenze). Un esempio pratico di utilizzo del notch filter, nel preprocessing di segnali EEG, è dovuto al’acquisizione indesiderata di segnali elettrici degli elettrodi stessi. I segnali elettrici solitamente viaggiano su due frequenze: 50Hz o 60Hz. Tramite il notch filtering vengono dunque eliminate queste frequenze poiché indesiderate.
4. **artifact rejection:** consiste nell’eliminare i dati acquisiti che non riguardano segnali provenienti dal cervello - ad esempio il battito delle palpebre, che porta valori con ampiezza tra i 100 e i 200 µV, per un intervallo di 250ms.

[Informazioni ottenute da: <http://learn.neurotechedu.com/preprocessing/>]

## Tecniche utilizzate in BED dataset

Nel paper presentato a corredo della creazione del dataset BED viene illustrato il processo di preparazione dei dati acquisiti, che ha permesso di trasformare i dati originali (raw) in dati analizzati (parsed). Tale processo è comunemente conosciuto come **prep pipeline** (catena di preparazione). Le operazioni sui dati sono state effettuate utilizzando la libreria EEGLAB per MathLab.

La **fase di preprocessing** è consistita in quattro operazioni:

1. eliminazione del rumore tramite *filtering*;
2. valutazione e stima dei dati dei segnali al fine di ottenere valori più veritieri;
3. rilevamento ed *interpolazione* dei bad channels.
4. *Feature extraction* per ottenere i dati parsed a partire dai dati raw. Le tecniche utilizzate a questo scopo sono state: **Mel-frequency ceptral coefficients (MFCC)**, **Auto-regression reflection coefficients (ARRC)**, **Spectral feature extraction (SPEC)** approfondite di seguito.

### MFCC

L’algoritmo di **Mel Frequency Cepstral Coefficients** consente di estrarre features dai canali a disposizione. Solitamente questa tecnica viene utilizzata per segnali audio, ma data la similitudine della distribuzione dei valori tra i segnali audio e quelli provenienti dal cervello sono stati svolti diversi studi a supporto dell’applicazione di questo algoritmo ai segnali EEG.

Nel caso di BED, la tecnica MFCC è stata applicata con una leggera variante. Innanzitutto i dati raw sono stati trasferiti nel dominio delle frequenze temporali attraverso l’applicazione della **trasformata di Fourier**. Allo spettro delle frequenze ottenuto è stata applicata la **Mel frequency filterbank** ovvero un insieme di filtri che elimina le frequenze non percepibili dall’uomo e le approssima a frequenze percepibili (Immaginato nel contesto dei segnali audio è più chiaro). I filtri operano su intervalli di frequenze, l’output dell’applicazione di un filtro su un intervallo di frequenze rappresenta uno dei coefficienti MFCC.

In BED dataset la filterbank contiene 18 filtri, ma gli output (coefficienti) considerati sono solo 12. I 6 coefficienti scartati sono i coefficients, che spesso non contengono informazioni utili ai fini di machine learning, come suggerito in “*Steady-State Visual Evoked Potentials for EEG-Based Biometric Identification”.* Tale operazione viene applicata ad ognuno dei 14 canali, per un totale di 168 features estratte.

### ARRC

La tecnica dell’**Auto Regression modeling (AR modeling)** è una tecnica parametrica utilizzata per la predizione del valore di un segnale a partire dalla somma lineare pesata dei precedenti valori. Formalmente, un AR model di ordine *p* prevede il valore del segnale *s(n)* di lunghezza *N* come somma dei precedenti *p* segnali:

dove rappresenta l’i-esimo coefficiente dell’AR model. Per stabilire il valore dei coefficienti viene utilizzato l’insieme di **equazioni di Yule-Walker**, così definito:

In BED dataset è stato utilizzato un modello di ordine *p = 12* per 14 canali, per un totale di 168 features estratte.

### SPEC

Dalla rappresentazione dello spettro delle frequenze di un segnale, ottenuto attraverso la trasformata di Fourier, è possibile estrarre alcune misurazioni utilizzate spesso in applicazioni machine learning per effettuare classificazione. Si parla in tali casi di **Spectral Features Extraction**. Anche in BED dataset vengono estratte quattro di queste caratteristiche, ottenute a partire dai segnali theta, alpha, beta e gamma. Le quattro caratteristiche estratte sono:

1. **spectral centroid: (indicatore di posizione)** data la media pesata dello spettro delle frequenze, il centroide è la frequenza rappresentativa del segnale;
2. **spectral bandwidth: (indicatore di variabilità)** rappresenta l’ampiezza dello spettro delle frequenze, che corrisponde alla varianza ponderata dei valori rispetto al centroide. Più bassa è l’ampiezza, più i valori delle frequenze si concentrano intorno al centroide;
3. **spectral crest factor: (indicatore di posizione)** calcola la presenza di picchi nello spettro delle frequenze. Un valore alto indica la presenza di picchi alti;
   * dato
4. **spectral flatness: (indicatore di variabilità)** indica l’omogeneità dei valori. Più il valore è alto, più i valori dello spettro delle frequenze sono ripetuti.

I valori

[Informazioni ottenute da: BED: A new dataset for eeg-based biometrics]

* Fonte: Representation learning and pattern recognition in cognitive biometrics: a survey

(URL: )

* + How to preprocess EEG signals [for human detection]
* Come fare tutto ciò in Python (librerie)

## Conclusioni sul preprocessing

* Criteri di valutazione della bontà del preprocessing
* Scelte

# Addestramento

## Machine Learning

* Breve introduzione
* Applicazione con EEG signals (ed eventuali documentazioni in letteratura)

## Deep Learning

* Breve introduzione
* Applicazione con EEG signals (ed eventuali documentazioni in letteratura)

# 

# Fonti da inserire

<https://www.researchgate.net/publication/369718848_Self-supervised_EEG_emotion_recognition_models_based_on_CNN>

<https://www.mdpi.com/2076-3417/12/5/2527>

<https://github.com/siddhi5386/Emotion-Recognition-from-brain-EEG-signals-/blob/master/Emotion_recognition_using_CNN.ipynb>

<https://www.yimp.it/preprocessing-machine-learning/>

—

Roba nuova

<https://wiki.uiowa.edu/display/hwanglab/EEG+Preprocessing>

<https://www.researchgate.net/figure/Preprocessing-Techniques-for-Scalp-EEG-Signals_fig5_339547751>

<https://www.frontiersin.org/articles/10.3389/fninf.2015.00016/full>

prep pipeline (quella che hanno trovato gli altri ragazzi (pelato + marco+ alessando):[Literature review on **EEG preprocessing**, feature extraction, and classifications techniques](https://www.researchgate.net/profile/Mohamed-Dessouky/publication/338119172_Literature_Review_on_EEG_Preprocessing_Feature_Extraction_and_Classifications_Techniques/links/5e0ccb4c4585159aa4ab4638/Literature-Review-on-EEG-Preprocessing-Feature-Extraction-and-Classifications-Techniques.pdf)

applicazione di prep e altre 3 pipeline:<https://ieeexplore.ieee.org/abstract/document/9047940>

<https://youtu.be/B9ti7boa9jc?list=PLElWgqOQ4bDuuEIGAjpgVf6Cgm4lhWfgo>

Per emozioni :<https://www.frontiersin.org/articles/10.3389/fnins.2020.622759/full>

<https://www.mdpi.com/2227-7080/10/4/79>

<https://www.mdpi.com/2075-4418/12/12/2984>

<https://www.nzbri.org/resources/publications/657/Shoorangiz__2021.pdf>

Human attention(paper che più assomiglia a quello che dobbiamo scrivere noi): [Detection of Human Attention Using EEG Signals | IEEE Conference Publication](https://ieeexplore.ieee.org/abstract/document/8430244)

# A Dynamic Multi-Scale Network for EEG Signal Classification: <https://www.frontiersin.org/articles/10.3389/fnins.2020.578255/full>