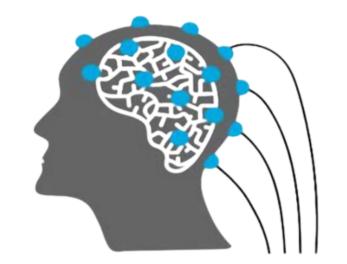
Prof. Michele Nappi Dott.ssa Chiara Pero



Eeg for human recognition

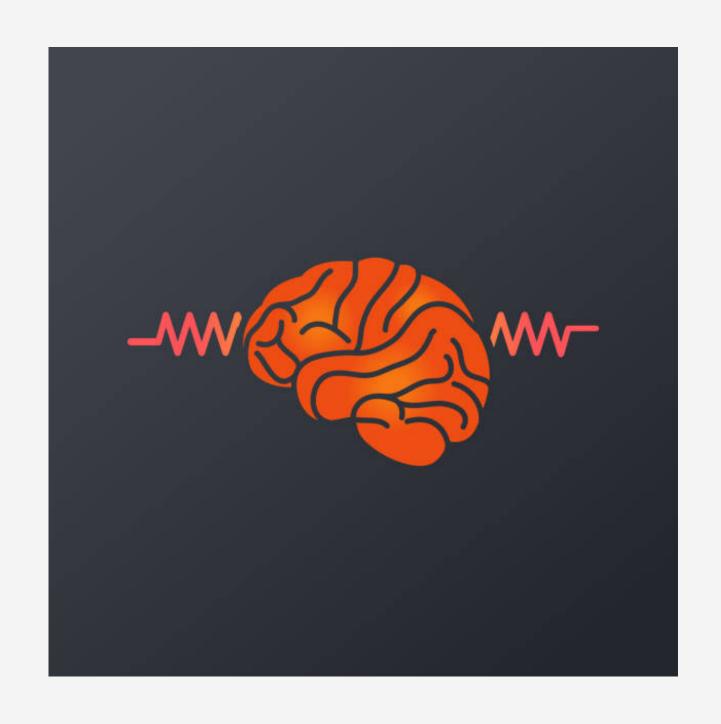
Progetto Fondamenti di Visione Artificiale e Biometria a.a. 2022/23

Mattia d'Argenio Simone Masullo



Programma

- <u>Introduzione</u>
- <u>Descrizione progetto</u>
- Stato dell'arte
- Metodo proposto
- <u>Sperimentazione</u>
- Risulati ottenuti
- Sviluppi futuri





Introduzione

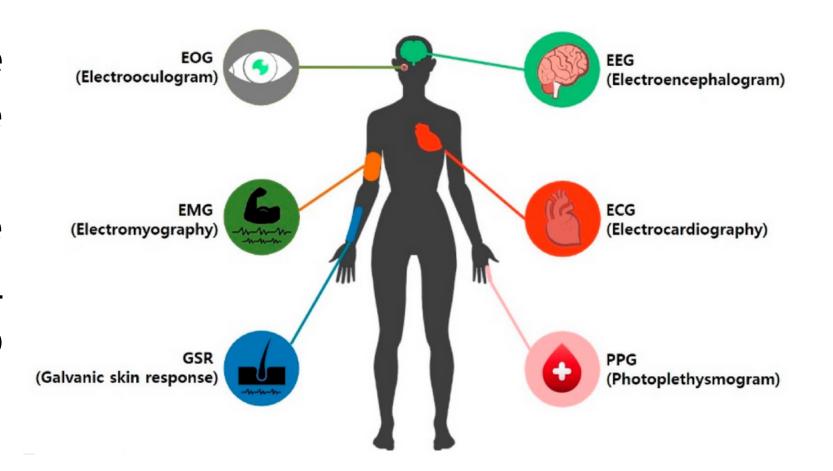
Introduzione alla biometria cognitiva



Introduzione

Cos'è la biometria cognitiva?

Integra le tradizionali modalità biometriche basate su caratteristiche fisiologiche e comportamentali con ulteriori elementi legati "al modo in cui pensiamo, sentiamo e reagiamo". Si basa sui segnali provenienti dal cervello, dal cuore e dal sistema nervoso autonomo.





Campi di applicazione

Sicurezza

Interazione uomo-macchina

Sanità



Introduzione

Task

Identificazione Autenticazione



Emotion recognition



Descrizione del progetto: obiettivo preposto

Descrizione dell'obiettivo preposto



Descrizione progetto

L'identificazione di un soggetto attraverso dati biometrici è un problema comune della computer vision, famoso come **human identification**.

I dati biometrici più utilizzati per istanze di questo problema sono l'impronta digitale, il volto o l'iride, ma esistono molte altre biometrie che offrono diversi vantaggi e svantaggi. Tra queste vi sono i **segnali elettro encefalografici** (Electroencephalogram - EEG), ovvero segnali elettrici provenienti dal cervello.



Descrizione progetto

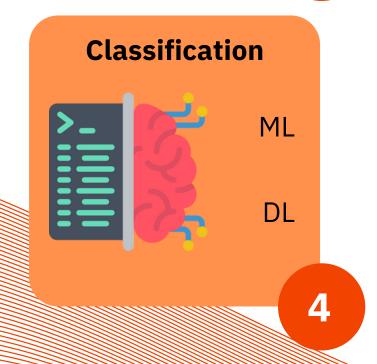
Biosignal acquisition

Signal elicitation protocol

Filtering
De-noising
Segmentation
Artifact removal

2

Features extraction Feature selection/ extraction



L'idea del progetto è stata quella di risolvere un'istanza del comune problema di human identification basandosi sui segnali EEG relativi ad ogni singolo soggetto. Per farlo vengono esaminate diverse tecniche di preprocessing note ed adoperati algoritmi di machine learning e deep learning.



Biometric EEG Dataset

Descrizione del dataset utilizzato



Biometric EEG Dataset

Il dataset comprende le risposte EEG di 21 soggetti a 12 stimoli diversi, suddivisi in 3 sessioni cronologicamente distinte. Sono stati inoltre considerati stimoli mirati a suscitare diversi stati affettivi, al fine di facilitare future ricerche sull'influenza delle emozioni per task biometrici basati su EEG.



Stato dell'arte

Descrizione dello stato dell'arte e dei lavori correlati



Stato dell'arte

BED: A New Data Set for EEG-Based Biometrics, 2021

Representation Learning and Pattern Recognition in Cognitive Biometrics: A Survey, 2022

The PREP pipeline: standardized preprocessing for largescale EEG analysis, 2015

State-of-the-art methods and future perspectives for personal recognition based on electroencephalogram signals, 2015



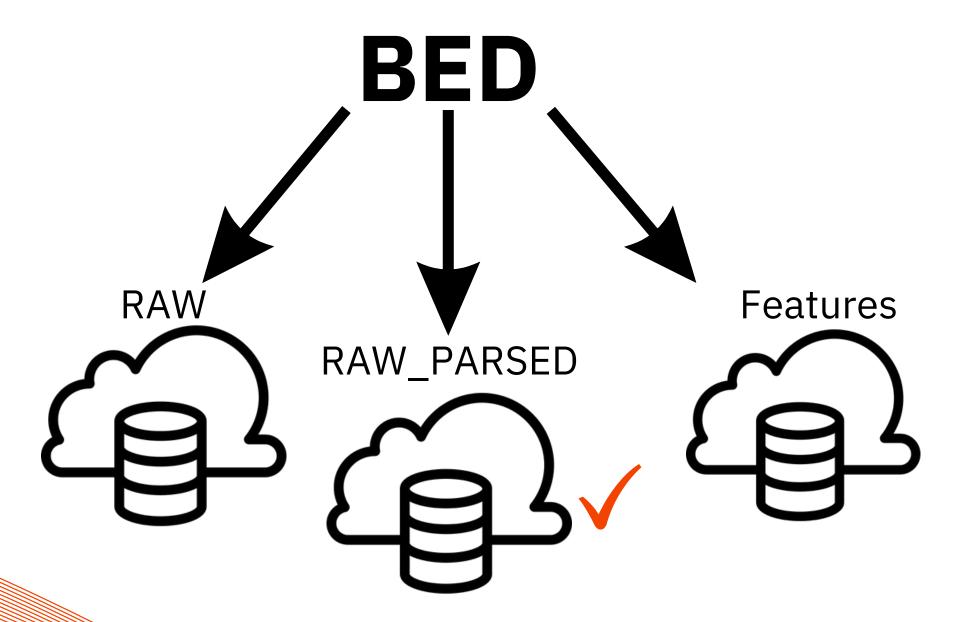
Stato dell'arte

A novel deep learning approach for classification of eeg motor imagery signals, 2016	90%
Human identification from brain eeg signals using advanced machine learning method eeg-based biometrics, 2016	94.4%
Human identification with electroencephalogram (eeg) signal processing, 2012	95.1%
Support vector machine approach for human identification based on eeg signals, 2020	99.1%



Approfondimento del metodo e sulle tecniche utilizzate per lo svolgimento del progetto







Tecniche di pre-processing utilizzate

Filtro passa-banda

- 1. eliminare la frequenza del segnale elettrico proveniente dai dispositivi stessi (50Hz);
- 2. eliminare le frequenze poco significative

Principal Component Analysis --> ridurre la numerosità delle

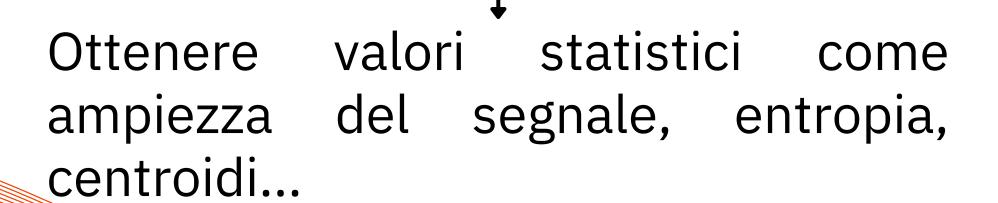
informazioni



Altre tecniche di feature extraction

Power Spectral Density

Wavelet Transform





Scelta del modello

Random Forest

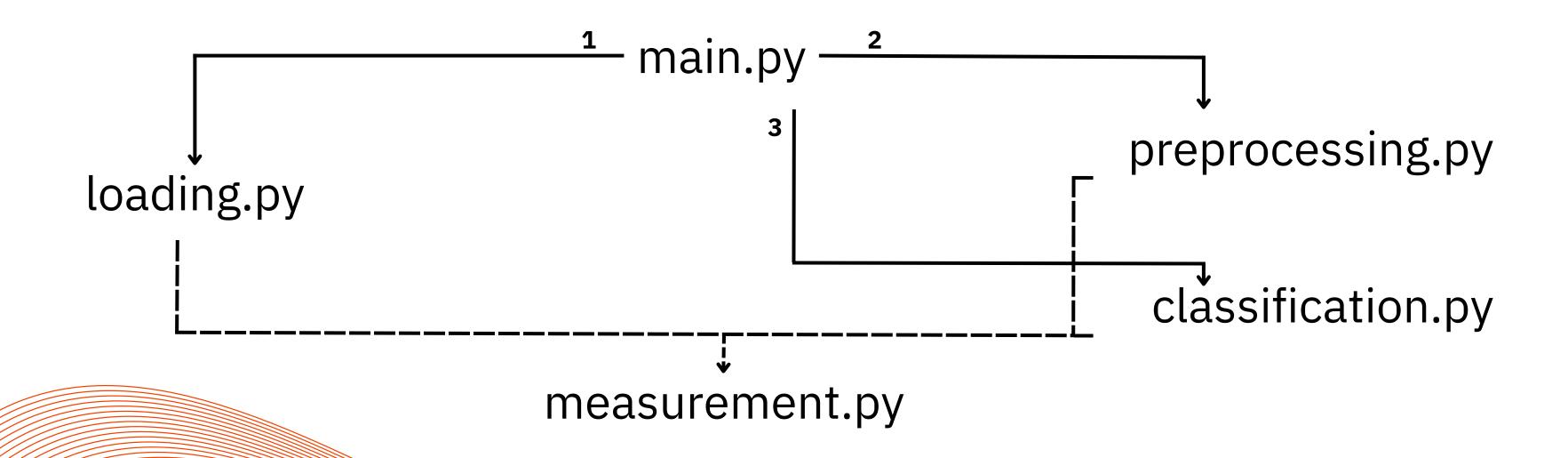
Grid Search

Xgboost



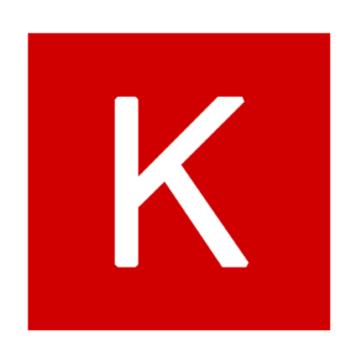
Implementazione delle tecniche di preprocessing e dei modelli utilizzati







Librerie utilizzate



Tensorflow

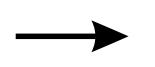
Keras

Scikit-learn





Tecniche di divisione del dataset



Split randomico e split per sessioni

Matrice di confusione



Computazione della matrice di confusione e delle metriche di valutazione del modello



Split randomico

Eseguito secondo il protocollo 70/30: 70% dei dati per il training 30% dei rimanenti per il testing

Split per sessioni

Eseguito utilizzando le prime due sessioni per il training e l'ultima sessione per il testing



Risultati ottenuti

Comparazioni dei risultati ottenuti nei vari modelli



Risultati ottenuti (random split)

	Accuracy	Precision	Recall	FScore
Random forest	99.9%	99.9%	99.9%	0.99
Grid search	99.9%	99.9%	99.9%	0.99
XGBoost	94.9%	94.9%	94.9%	0.97
LSTM	~20,15%	- N/A -	- N/A -	- N/A -



Risultati ottenuti (session split)

	Accuracy	Precision	Recall	FScore
Random forest	99.9%	99.9%	94.1%	0.96
Grid search	98.8%	97.7%	95.1%	0.97
XGBoost	91.8%	80.9%	83.2%	0.60
LSTM	~19,47%	- N/A -	- N/A -	- N/A -



Sviluppi futuri

Descrizione degli sviluppi futuri del progetto.



Sviluppi futuri

L'applicazione risulta performante dal punto L'applicazione risulta performante dal punto di vista dell'accuracy ma non dal punto di vista dell'esecuzione. Al fine di risolvere questo problema si è pensato ad una parallelizzazione del carico di lavoro.

