



Analisi di serie temporali per il monitoraggio di bambini con emiplegia

Davide Marchi
Giordano Scerra



UNIVERSITÀ
DI PISA



AInCP

Introduzione

*Utilizzo dell'**intelligenza artificiale** per creare una diagnosi e una riabilitazione personalizzata per i bambini con **emiplegia***



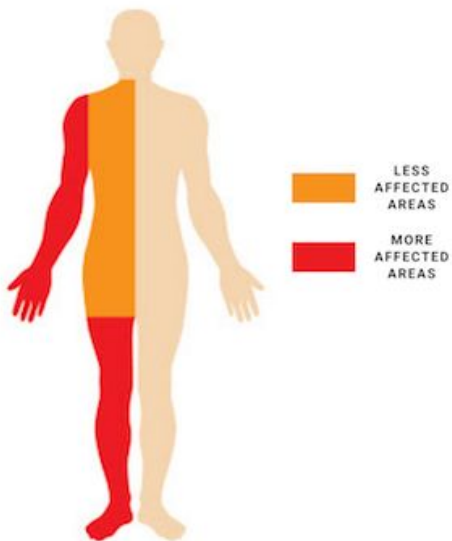


Emiplegia e PCI

Introduzione

HEMIPLEGIA

AFFECTS ONLY ONE SIDE OF THE BODY



I soggetti dello studio sono bambini con Paralisi Cerebrale Infantile (PCI)

Nello specifico **PCI a tipo emiplegia**:

- Colpisce un solo emisfero cerebrale risultando in una parziale immobilità di un lato del corpo



I nostri obiettivi

Introduzione

Realizzazione di uno strumento di monitoraggio a partire da dati registrati (Activity Counts) da sensori indossabili non invasivi chiamati **attigrafi**

Datetime	x_D	y_D	z_D	x_ND	y_ND	z_ND
2017-05-06 18:31:04	15	12	6	8	0	0
2017-05-06 18:31:05	0	6	26	0	6	4
2017-05-06 18:31:06	0	1	14	0	10	4

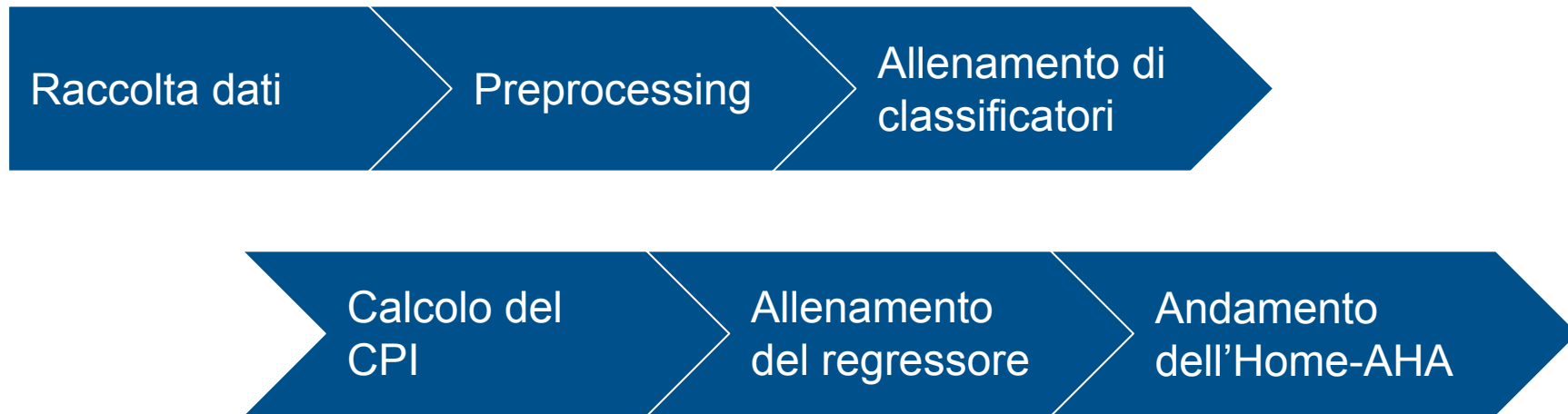




Struttura della tesi

Introduzione

La tesi è strutturata secondo il seguente schema





Raccolta dati

Raccolta dati e preprocessing

Utilizzo di attigrafi durante:

Sessioni AHA	Sessioni WEEK
<ul style="list-style-type: none">● 20 minuti● Semistrutturate● Personale clinico assegna Assisting Hand Assessment	<ul style="list-style-type: none">● 6 giorni● Non strutturate● Non supervisionate

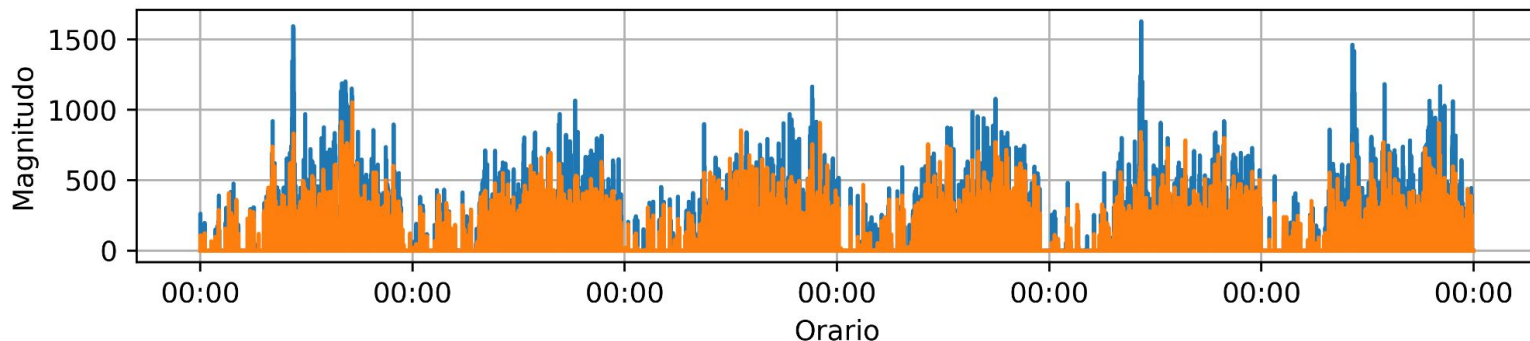


Preprocessing

Raccolta dati e preprocessing

Calcolando la **magnitudo** del vettore spostamento a partire dagli AC otteniamo:

- 60 serie temporali a **due variabili** delle sessioni AHA
- 60 serie temporali a **due variabili** delle sessioni WEEK



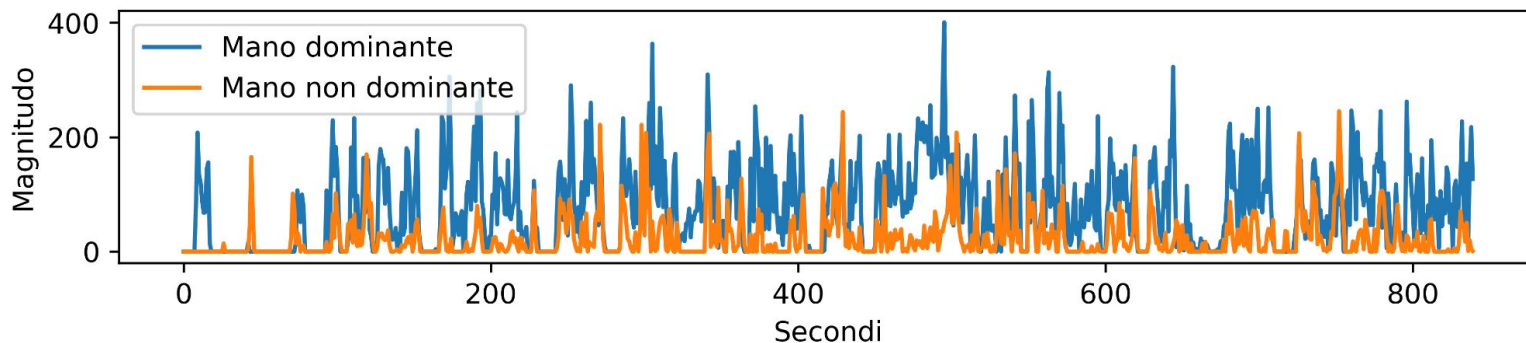


Manipolazione delle serie temporali

Dataset e preprocessing

Tre approcci per trattare le serie temporali:

1. Concatenazione
2. Differenza
3. Asimmetry Index



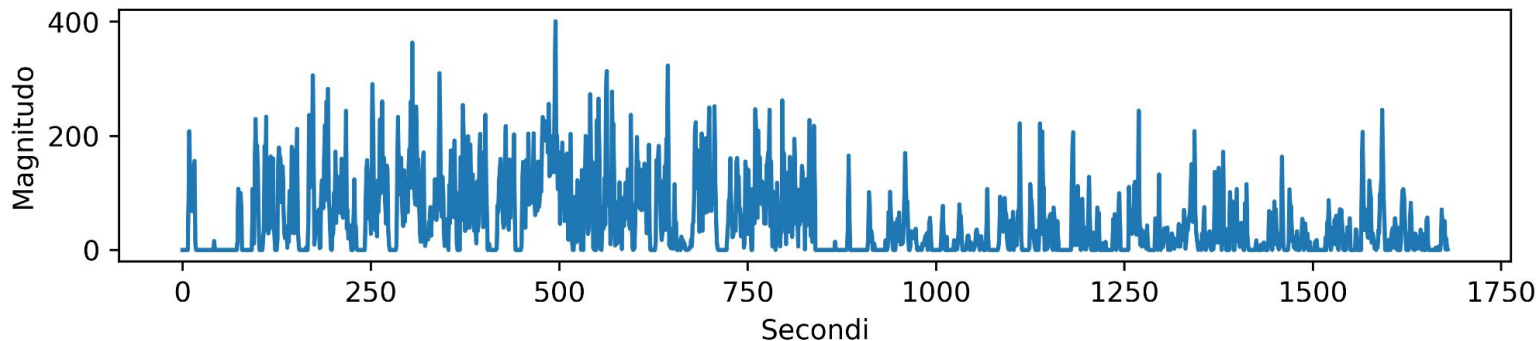


Manipolazione delle serie temporali

Dataset e preprocessing

Tre approcci per trattare le serie temporali:

1. **Concatenazione**
2. Differenza
3. Asimmetry Index



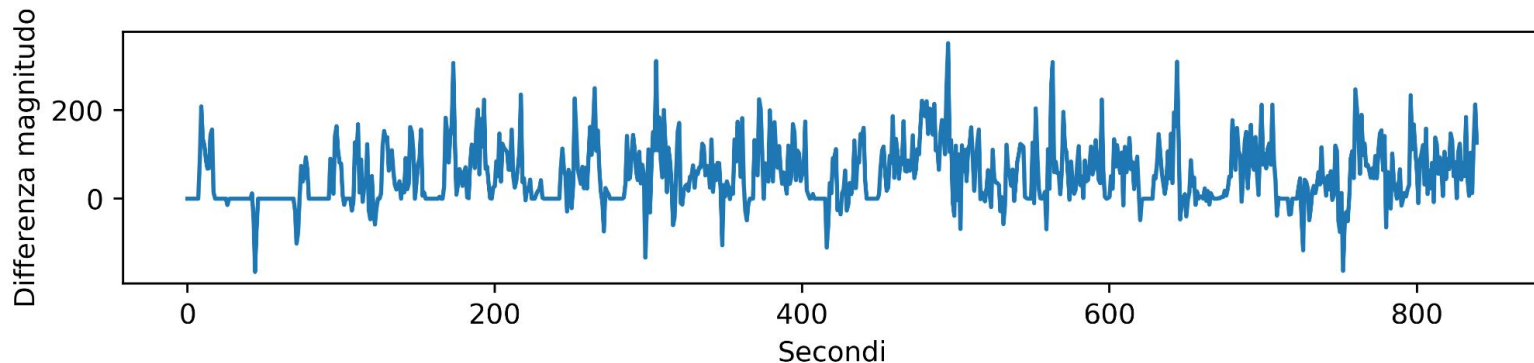


Manipolazione delle serie temporali

Dataset e preprocessing

Tre approcci per trattare le serie temporali:

1. Concatenazione
2. **Differenza**
3. Asimmetry Index



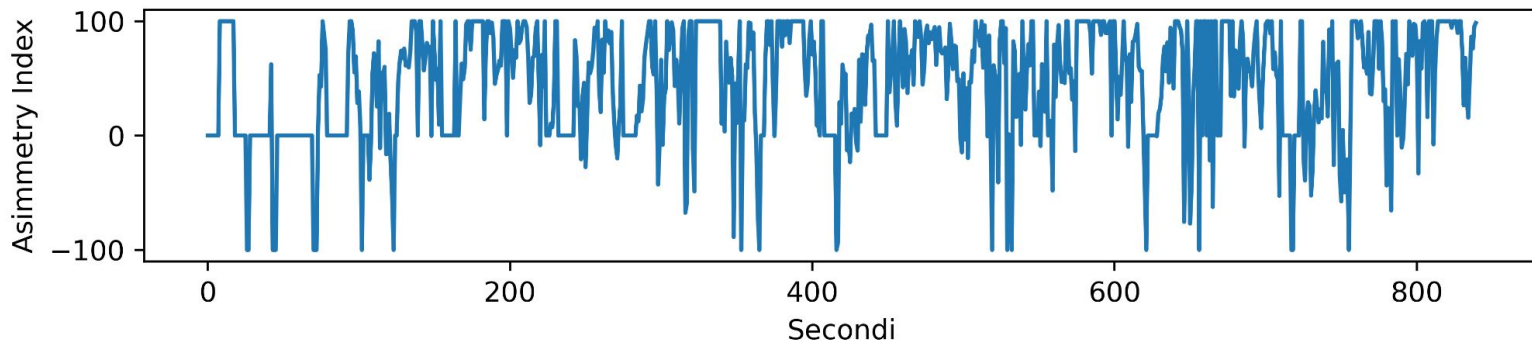


Manipolazione delle serie temporali

Dataset e preprocessing

Tre approcci per trattare le serie temporali:

1. Concatenazione
2. Differenza
3. **Asimmetry Index**



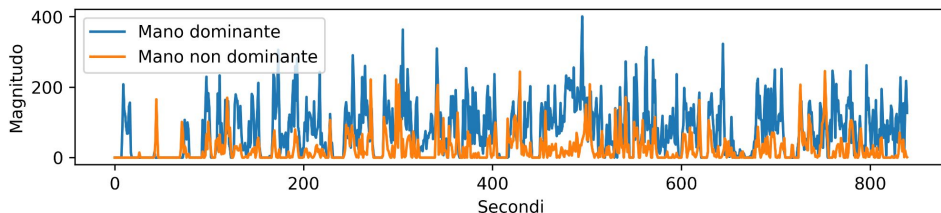


Suddivisione dei sample

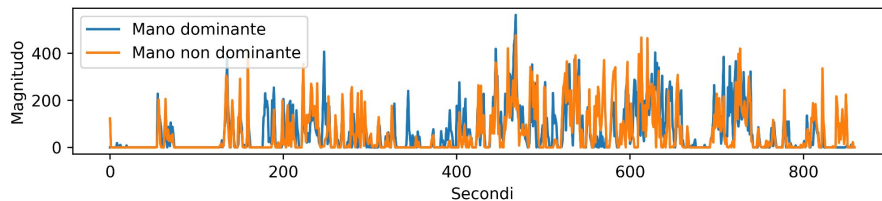
Analisi

Approccio iniziale:

Lasciare che algoritmi non supervisionati provino a raggruppare le serie temporali in modo significativo



?



?

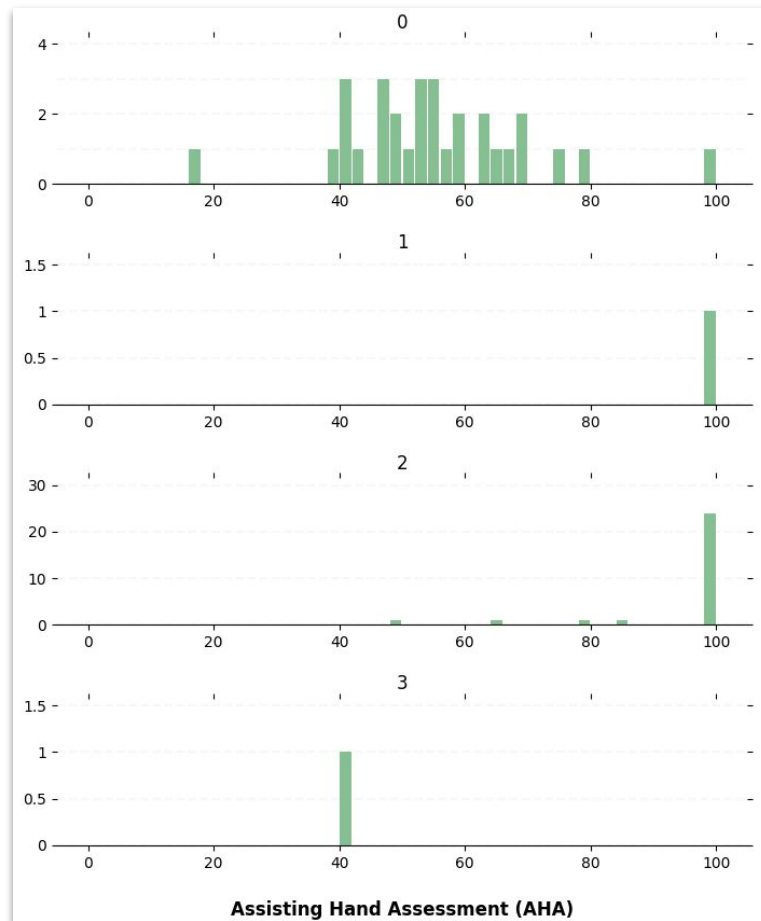


Cluster multipli

Analisi

Tentativo di suddividere le serie temporali
delle sessioni AHA in **più cluster**

Esempio di clustering →



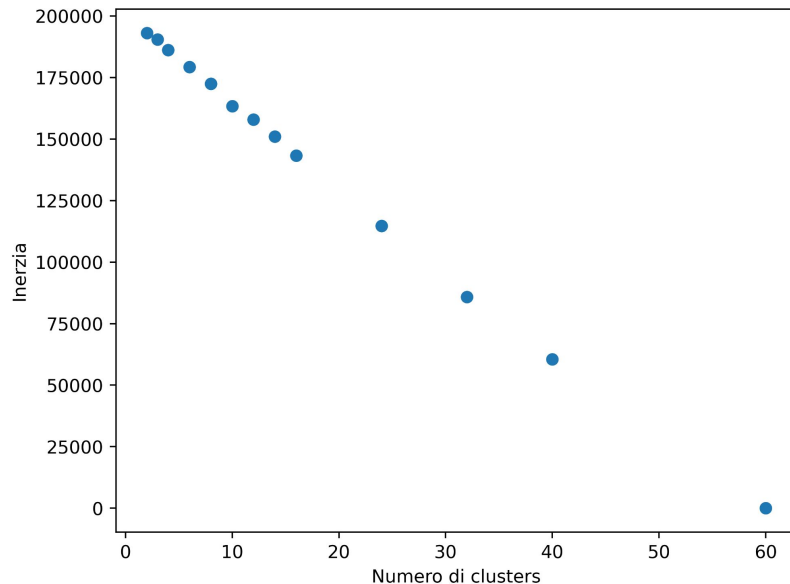


Cluster multipli

Analisi

Tentativo di suddividere le serie temporali delle sessioni AHA in **più cluster**

Elbow plot dell'inerzia →





Classificazione binaria

Analisi

Modelli **supervised** o **unsupervised** per decidere se un sample appartenga ad un soggetto con emiplegia o ad un soggetto del gruppo di controllo

Esploriamo:

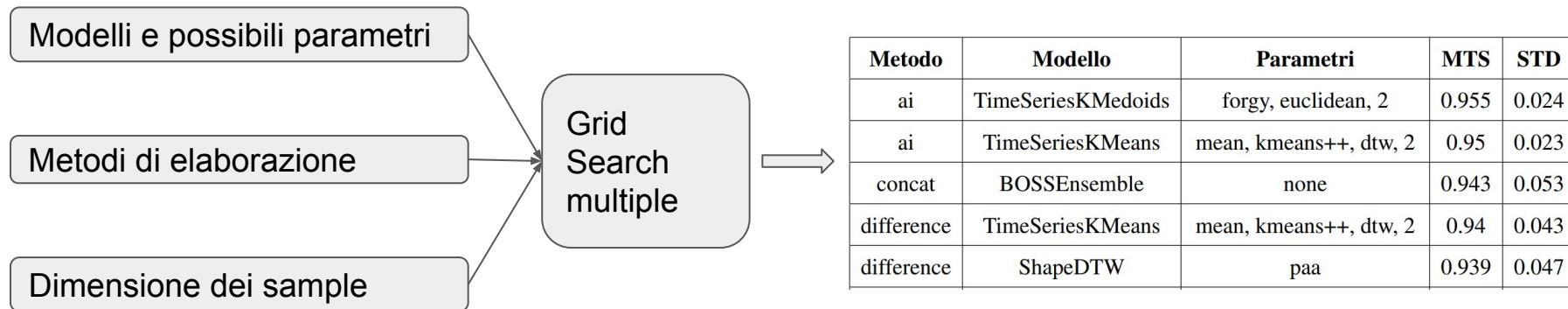
- Dimensioni dei sample
 - 300, 600, 900 secondi
- Tipi di elaborazioni delle serie temporali
 - Concatenazione, Differenza, Asimmetry Index
- Algoritmi con diversi parametri
 - KMeans, KMedoids, ShapeDTW, BOSSEnsemble



Model selection

Analisi

Grid search multiple per poter confrontare i vari approcci



MTS = Mean Test Score calcolato con f1 score



Risultati della model selection

Analisi

La grid search ha
prodotto i seguenti
risultati per sample da
900 secondi

Metodo	Modello	Parametri	MTS	STD
difference	TimeSeriesKMeans	mean, kmeans++, dtw, 2	1.0	0.0
ai	TimeSeriesKMeans	mean, kmeans++, euclidean, 2	1.0	0.0
difference	TimeSeriesKMedoids	forgy, dtw, 2	0.983	0.033
difference	ShapeDTW	paa	0.983	0.034
ai	TimeSeriesKMedoids	random, euclidean, 2	0.983	0.034
ai	ShapeDTW	paa	0.983	0.034
ai	BOSSEnsemble	none	0.95	0.041
concat	BOSSEnsemble	none	0.948	0.07
difference	BOSSEnsemble	chi2	0.934	0.062
concat	TimeSeriesKMedoids	forgy, dtw, 2	0.77	0.136



Risultati della model selection

Analisi

La grid search ha
prodotto i seguenti
risultati per sample da
600 secondi

Metodo	Modello	Parametri	MTS	STD
ai	TimeSeriesKMeans	mean, kmeans++, euclidean, 2	1.0	0.0
ai	TimeSeriesKMedoids	forgy, euclidean, 2	0.988	0.024
ai	BOSSEnsemble	none	0.976	0.029
difference	TimeSeriesKMeans	mean, kmeans++, euclidean, 2	0.976	0.03
difference	TimeSeriesKMedoids	forgy, dtw, 2	0.976	0.03
ai	ShapeDTW	paa	0.952	0.044
difference	ShapeDTW	raw	0.951	0.045
difference	BOSSEnsemble	none	0.947	0.08
concat	BOSSEnsemble	chi2	0.947	0.08
concat	TimeSeriesKMeans	mean, kmeans++, euclidean, 2	0.896	0.208



Risultati della model selection

Analisi

La grid search ha
prodotto i seguenti
risultati per sample da
300 secondi

Metodo	Modello	Parametri	MTS	STD
ai	TimeSeriesKMedoids	forgy, euclidean, 2	0.955	0.024
ai	TimeSeriesKMeans	mean, kmeans++, dtw, 2	0.95	0.023
concat	BOSSEnsemble	none	0.943	0.053
difference	TimeSeriesKMeans	mean, kmeans++, dtw, 2	0.94	0.043
difference	ShapeDTW	paa	0.939	0.047
ai	BOSSEnsemble	none	0.938	0.043
difference	TimeSeriesKMedoids	forgy, dtw, 2	0.933	0.037
ai	ShapeDTW	paa	0.933	0.026
difference	BOSSEnsemble	none	0.928	0.045
concat	ShapeDTW	paa	0.809	0.045



Risultati della model selection

Analisi

La grid search ha
prodotto i seguenti
risultati per sample da
300 secondi

Metodo	Modello	Parametri	MTS	STD
ai	TimeSeriesKMedoids	forgy, euclidean, 2	0.955	0.024
ai	TimeSeriesKMeans	mean, kmeans++, dtw, 2	0.95	0.023
concat	BOSSEnsemble	none	0.943	0.053
difference	TimeSeriesKMeans	mean, kmeans++, dtw, 2	0.94	0.043
difference	ShapeDTW	paa	0.939	0.047
ai	BOSSEnsemble	none	0.938	0.043
difference	TimeSeriesKMedoids	forgy, dtw, 2	0.933	0.037
ai	ShapeDTW	paa	0.933	0.026
difference	BOSSEnsemble	none	0.928	0.045
concat	ShapeDTW	paa	0.809	0.045

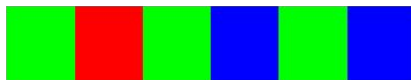


Continuous Performance Indicator

Analisi

1. Suddivisione della settimana in sample da 300 secondi
2. Classificazione del sample tra:
 - a. **Sample** non valido
 - b. **Sample** di un soggetto con emiplegia (**SE**)
 - c. **Sample** di un soggetto del gruppo di controllo (**SC**)
3. Calcolo del CPI

Modello



$$CPI = \frac{\#SC}{\#SC + \#SE} \cdot 100 = \frac{3}{3+2} \cdot 100 = 60$$

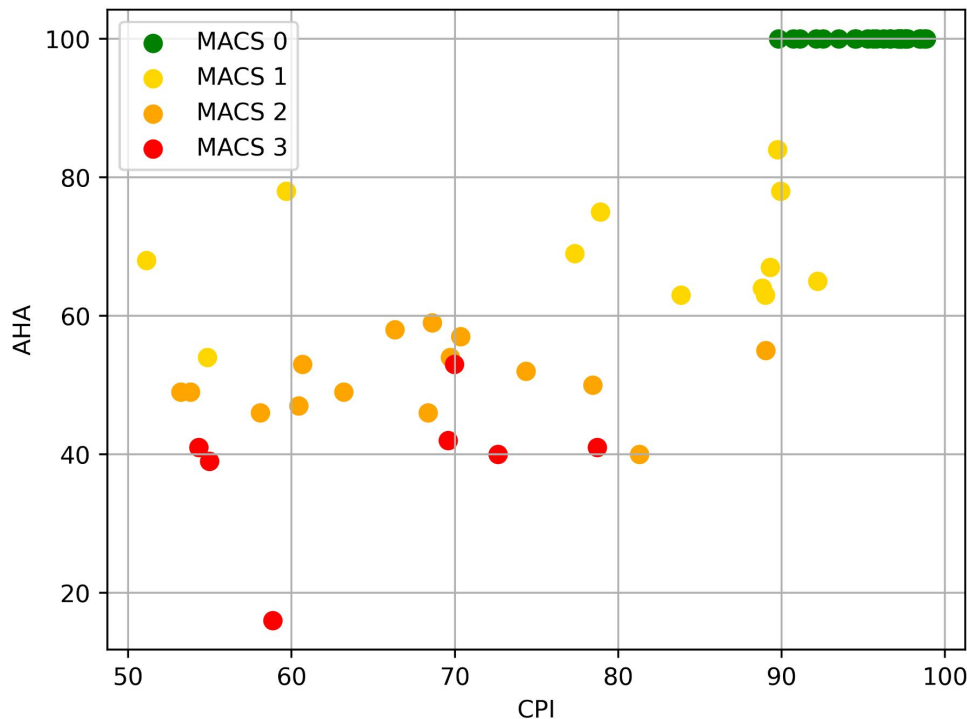


Valutazione del CPI

Analisi

La significatività del CPI è stata misurata calcolando la **correlazione di Pearson** (ρ) tra CPI e AHA

$$\rho_{CPI,AHA} = 0.826$$





Utilizzo del CPI

Dashboard

Necessità di:

- Visualizzare l'andamento delle prestazioni
- Avere una misura dell'affidabilità dell'indicatore
- Renderlo comprensibile al personale clinico

Soluzione:

- Realizzazione di un **dashboard** →

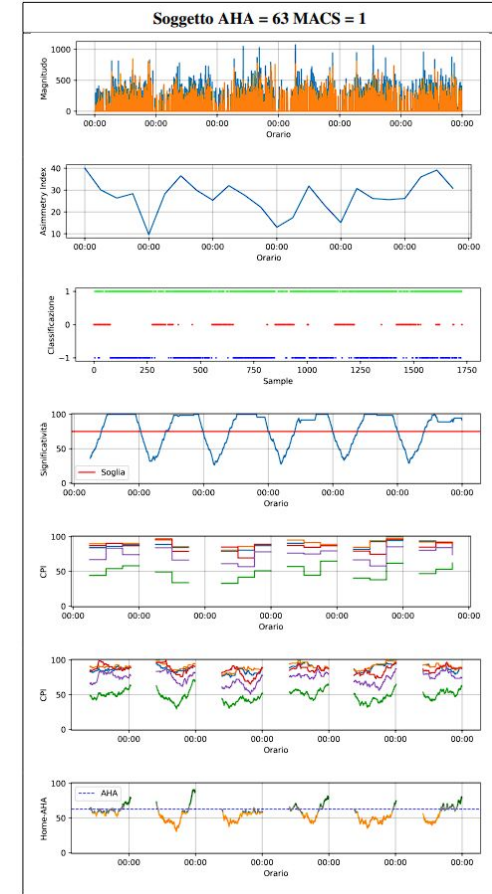
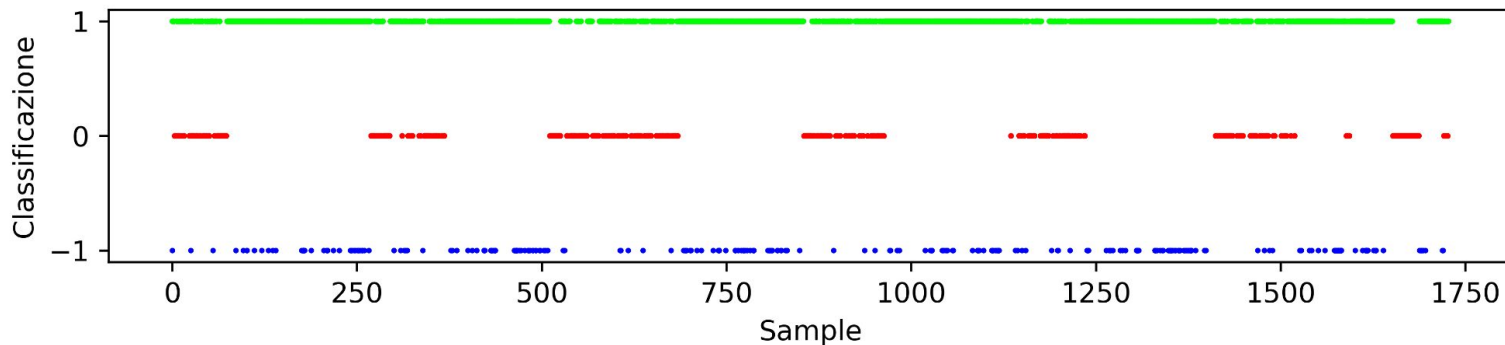
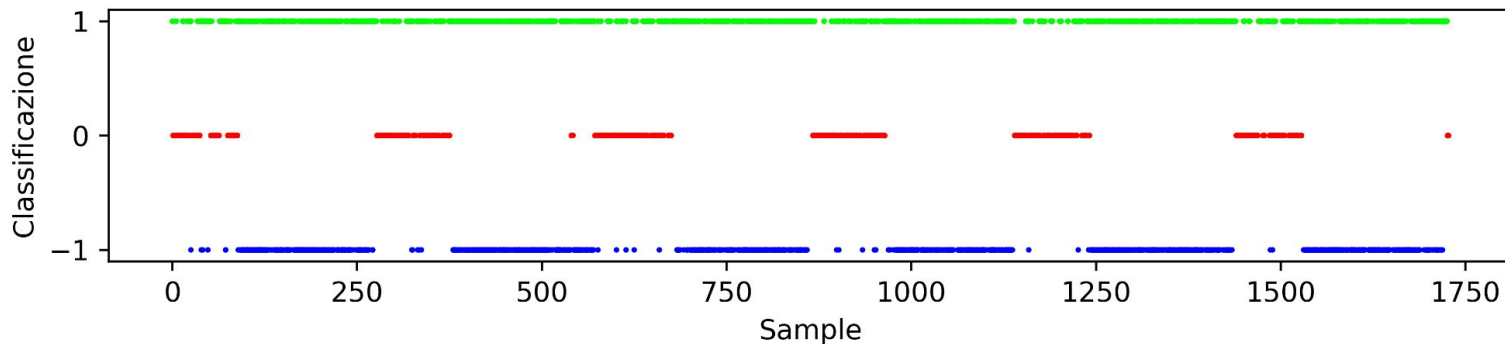




Grafico delle predizioni

Dashboard







Finestre e significatività

Dashboard

Usiamo **finestre da 6 ore** (72 sample):

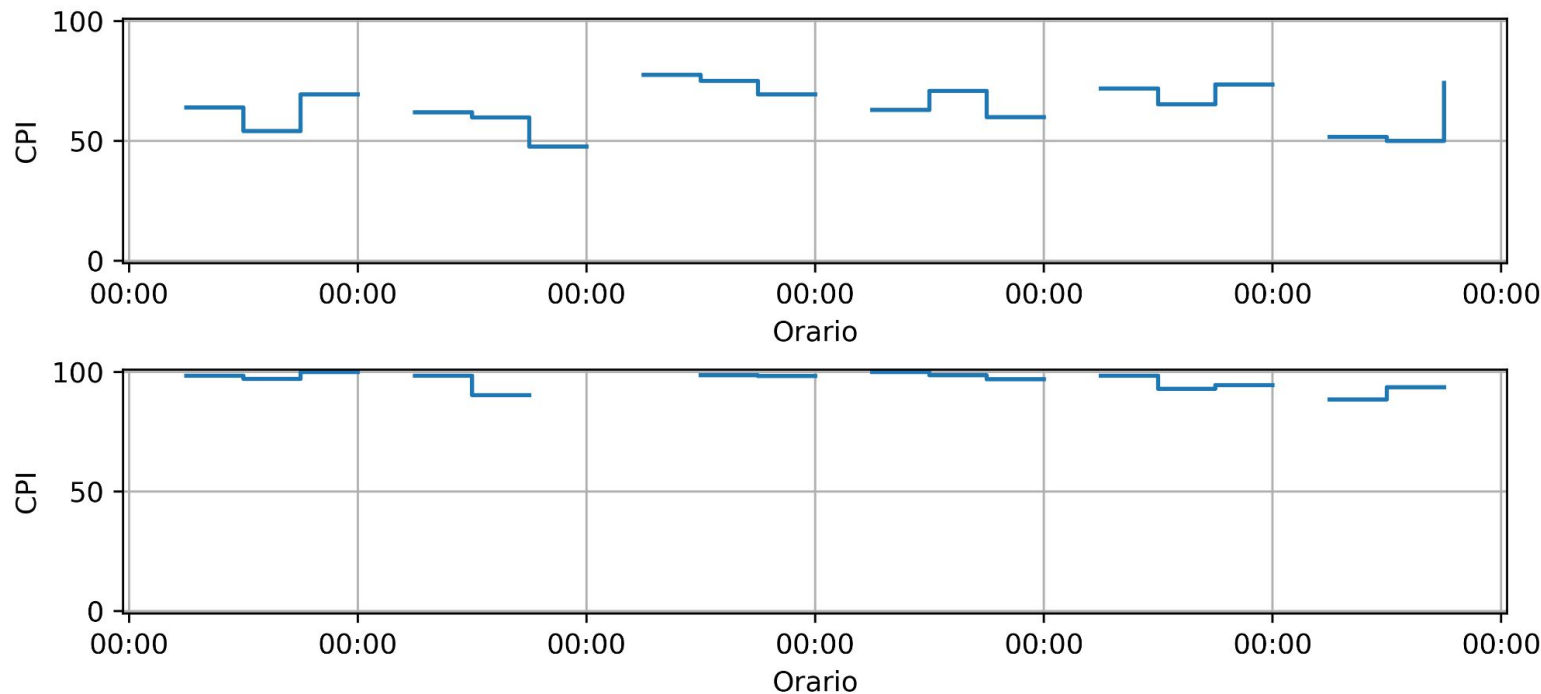
- Disgiunte 
- Scorrevoli 

Per ognuna di loro teniamo conto di una soglia di significatività, che deve essere maggiore del **75%** per calcolare il CPI



Finestre sequenziali disgiunte

Dashboard





Finestra scorrevole

Dashboard

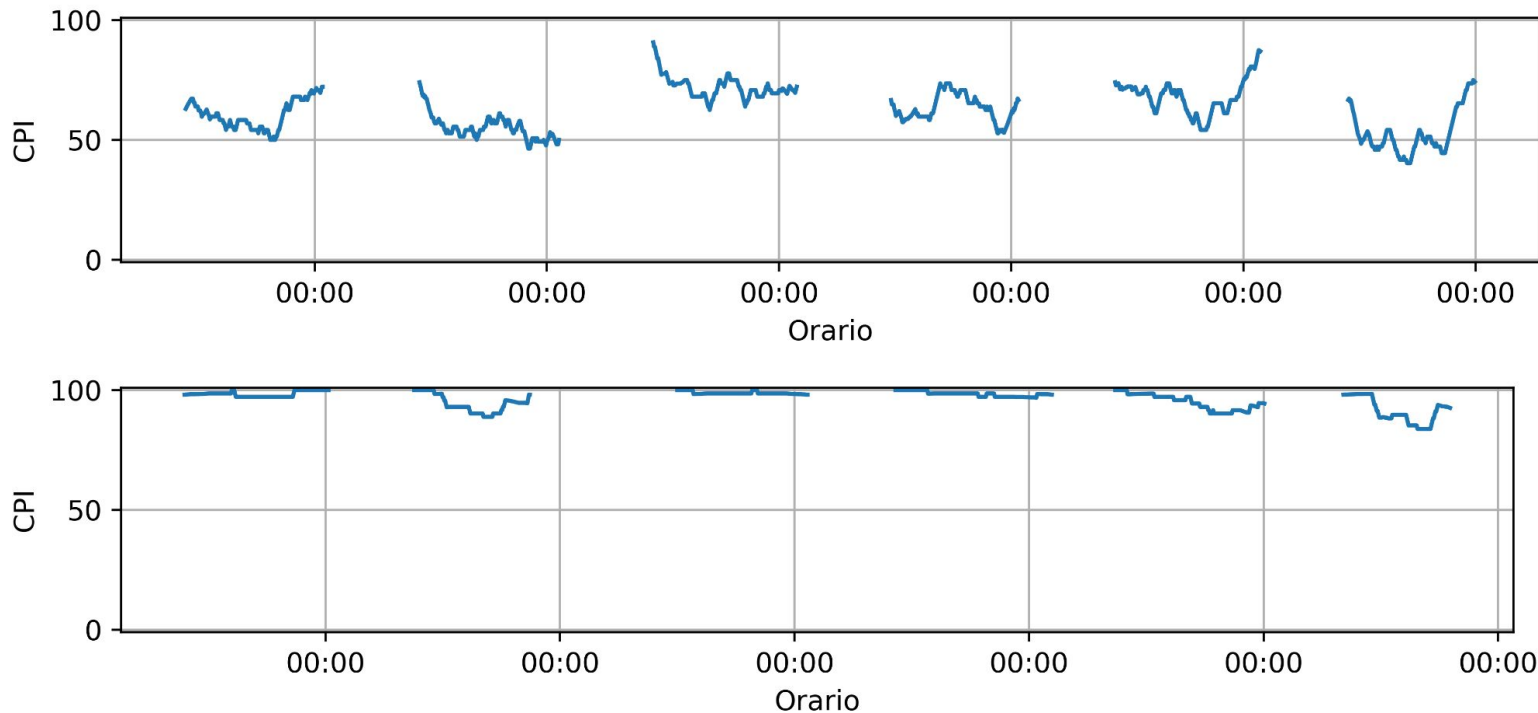
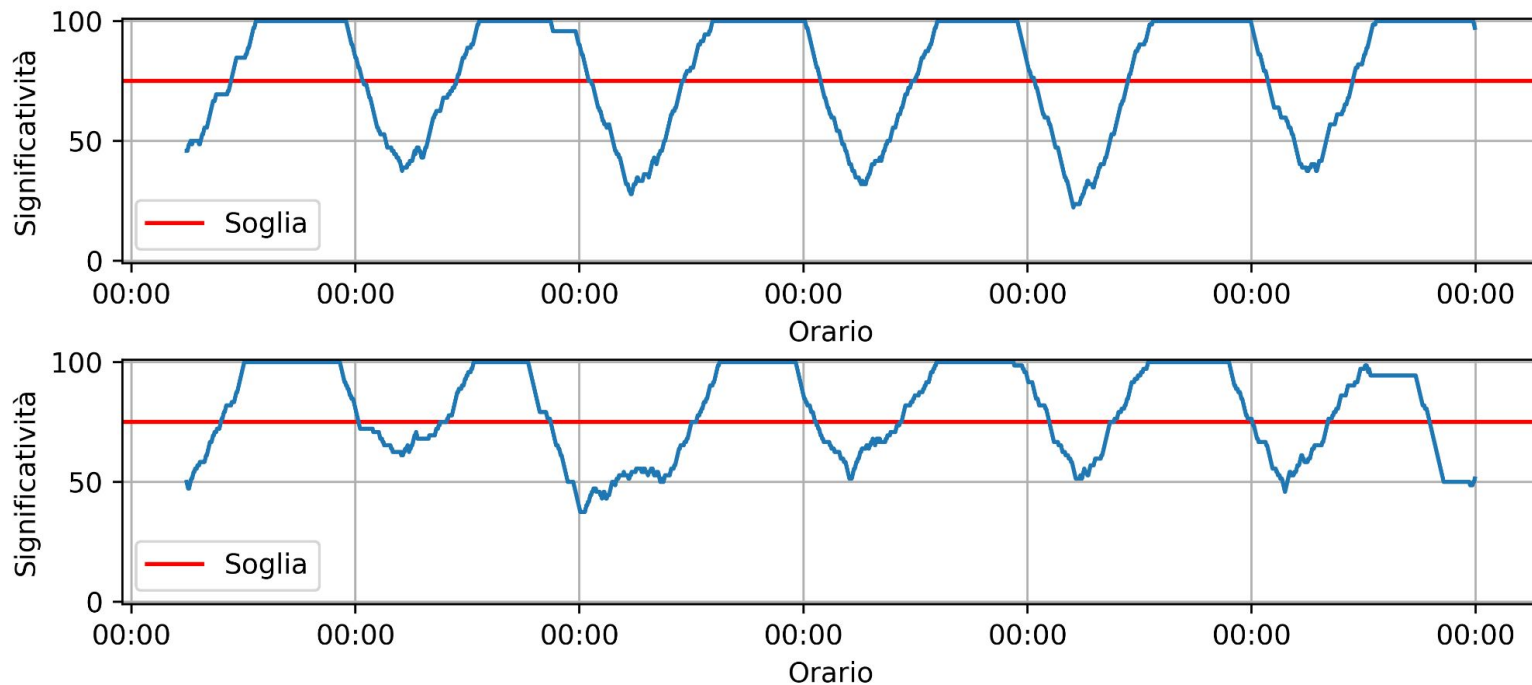




Grafico della significatività

Dashboard

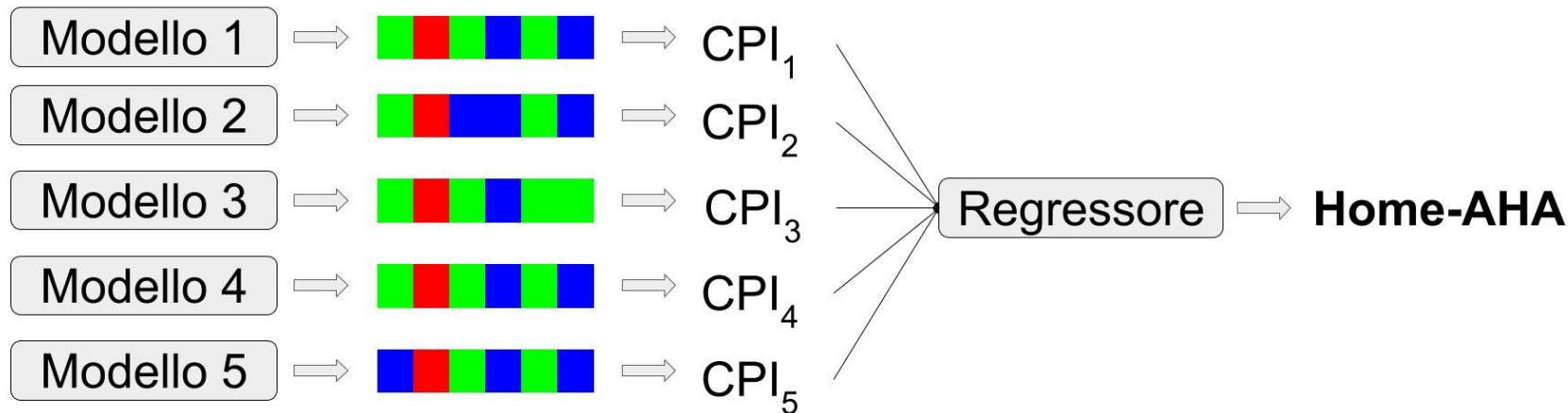




Stimare l'AHA: Home-AHA

Dashboard

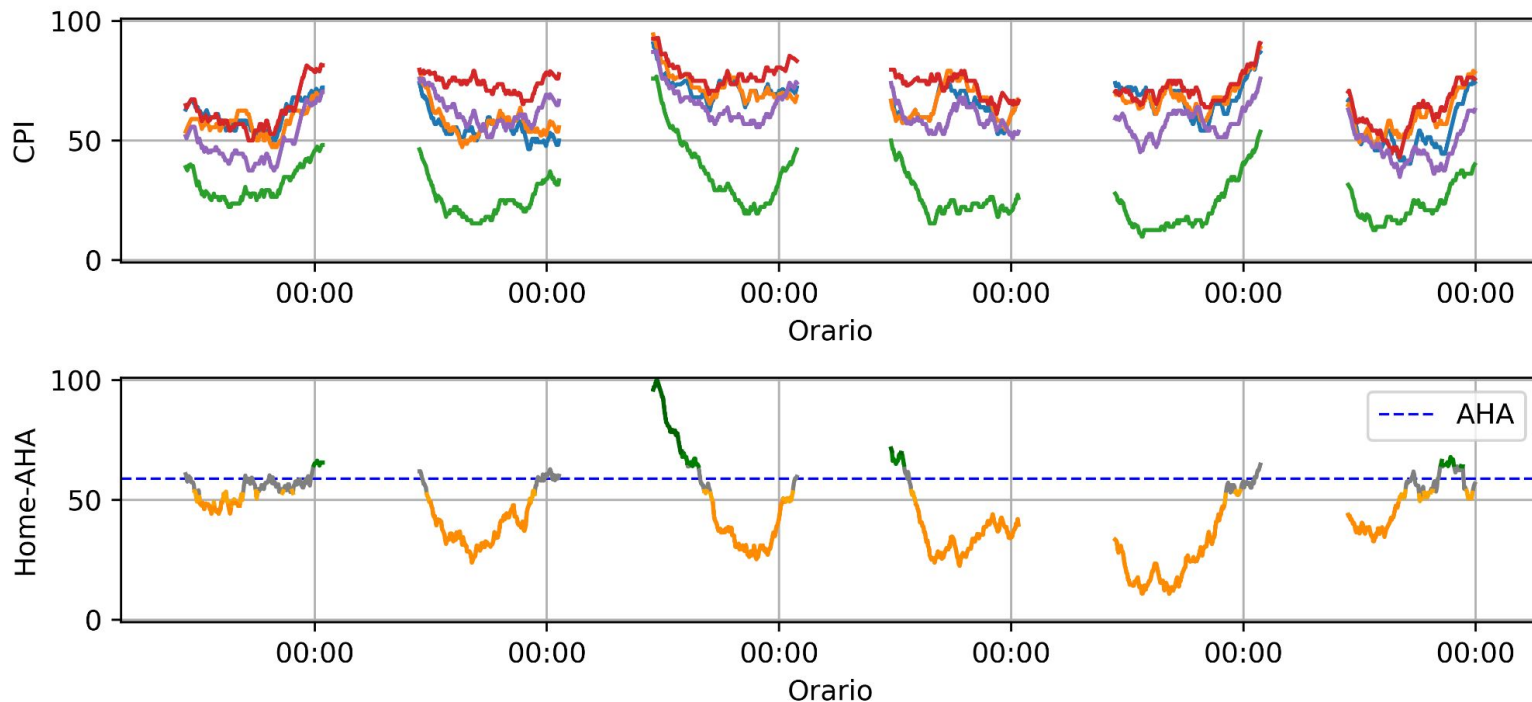
Abbiamo allenato un regressore lineare in grado di stimare l'AHA a partire dal CPI calcolato dai 5 migliori modelli, le cui prestazioni sono state misurate calcolando il **coefficiente di determinazione ($R^2 = 0.747$)**





Andamento dell'Home-AHA

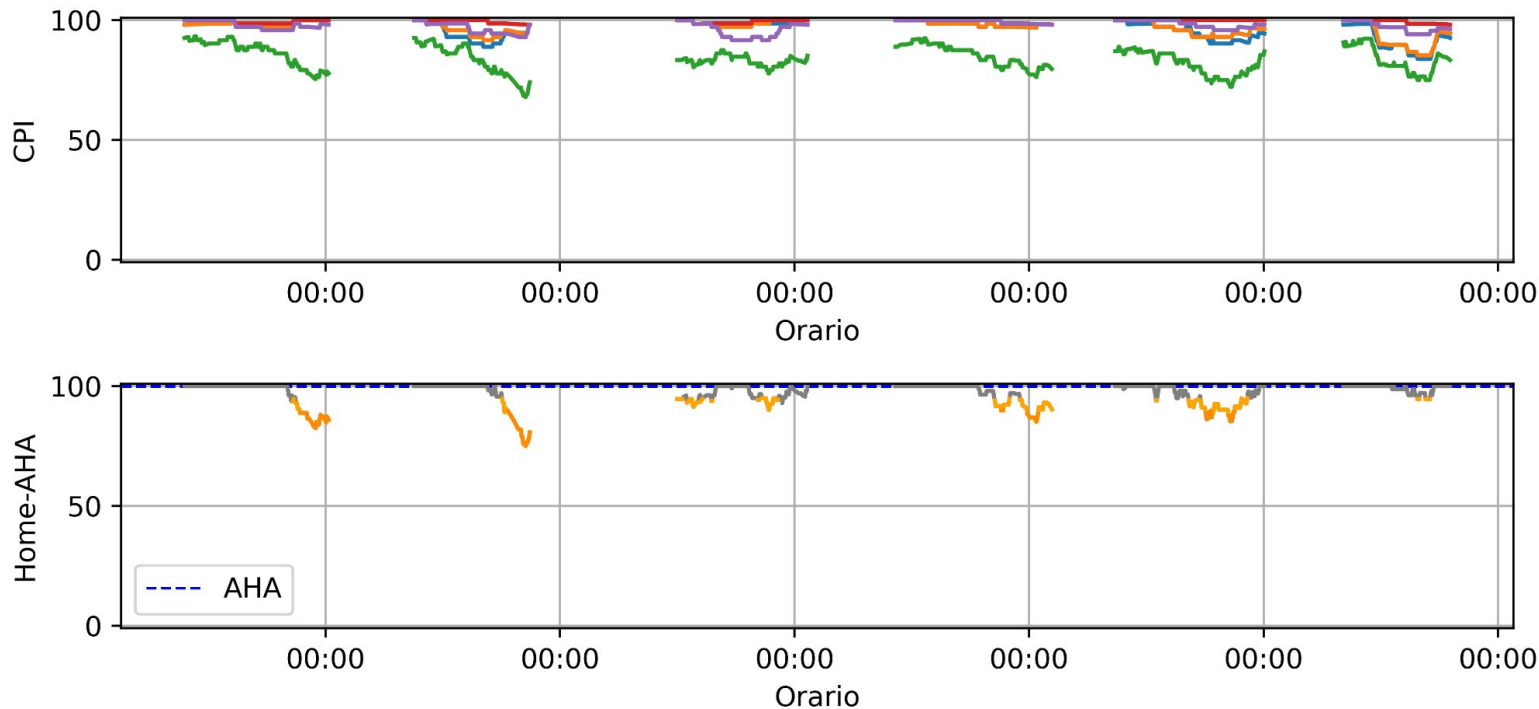
Dashboard → AHA = 59





Andamento dell'Home-AHA

Dashboard → AHA = 100





Risultati ottenuti e contributi

Risultati

- **CPI** come feature
 - Comprovato l'utilizzo del CPI come feature utile per il monitoraggio
- **Grafico dell'Home-AHA**
 - Andamento che stima il punteggio AHA durante i sei giorni del soggetto

Mediante l'implementazione della seguente pipeline





Conclusioni dell'analisi

Conclusioni

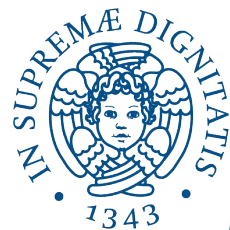
CPI e l'Home-AHA sono frutto dell'addestramento di modelli di ML, e dunque i risultati da loro prodotti sono strettamente dipendenti dal contenuto del dataset di partenza

Possibili **sviluppi futuri** potrebbero dunque essere:

- Aumento della quantità e qualità dei dati
- Esplorazione di ulteriori modelli e parametri
- Gestione e trasferimento dei dati



**Ringraziamo per
l'attenzione**



**UNIVERSITÀ
DI PISA**