Intelligenza artificiale e Riabilitazione

Stimare il cambiamento dei sintomi motori nella malattia di Parkinson con dati di vita quotidiana e Fuzzy Mixture of Experts

Lucia Pepa

Dipartimento di Ingegneria dell'Informazione
Università Politecnica delle Marche

In collaborazione con:
Prof. Lucio Ciabattoni
Prof. Luca Spalazzi
Prof.ssa Maria Gabriella Ceravolo
Prof.ssa Marianna Capecci

Background e spazio di ricerca

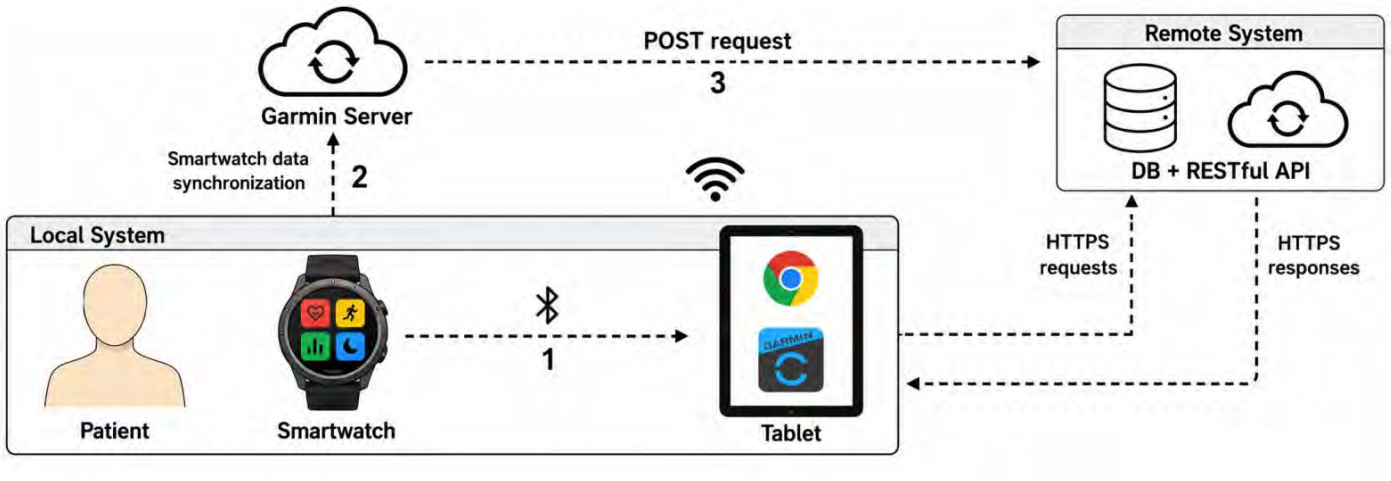
Migliorare la gestione della malattia di Parkinson

Bisogno	Evidenze scientifiche	Limitazioni
Telemonitoraggio: accurata caratterizzazione dello stato funzionale e della progressione di malattia durante la vita quotidiana	Descrivere sintomi e fluttuazioni motorie Smartwatch e smartphone (Fay-Karmon, 2024)	Riconoscimenti puntuali, no traiettorie. Pochi studi in ambito domestico/quotidiano.
Teleriabilitazione: percorsi riabilitativi di intensità adeguata, centrati sulla persona	Accompagnare l'esercizio a casa Supporto mHealth (Ellis, 2019)	Non integrato con telemonitoraggio. Pochi studi in ambito domestico/quotidiano.

RAPIDO – si pone l'obiettivo di:

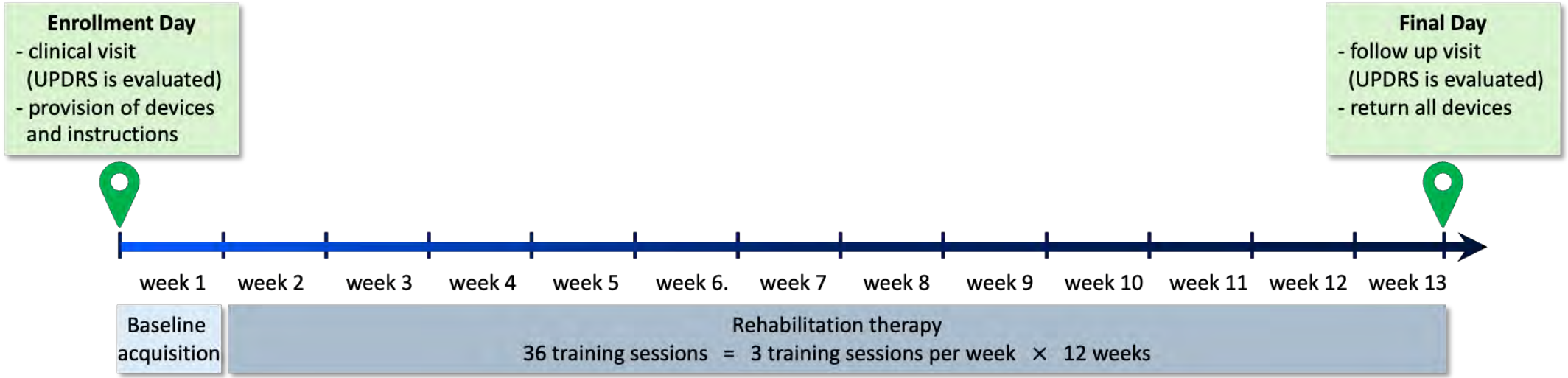
- **proporre un sistema integrato di telemonitoraggio e teleriabilitazione basato su dispositivi di uso comune;**
- **testarlo in ambiente domestico su una popolazione di persone con malattia di Parkinson;**
- **indagare approcci di machine learning per estrarre informazioni utili dai dati di telemonitoraggio**

RAPIDO: il percorso a domicilio



Garmin Vivosmart 4 e tablet Samsung A7 con accesso a K-REHAB

- 95 persone con MdP arruolate
- 59 M, 36 F
- età media: 67 $\pm$ 8
- H&Y: 2.1 $\pm$ 0.9
- BMI: 26 $\pm$ 5 kg/m^2
- 85 persone contribuiscono all'analisi



Obiettivo di ricerca

- **Fattibilità**
- **stimare cambiamento dei sintomi motori (UPDRS-III) dai dati smartwatch**
  - **regressione → outcome continuo**
  - **classificazione → outcome categorizzato**

Outcome continuo: ΔUPDRS-III = punteggio al follow-up − punteggio al baseline

Outcome categorizzato	Soglia clinica (MCID: Mininally Clinical Important Difference)	Pazienti
Migliorato	$\Delta \leq -3{,}25$	21
Stabile	$-3{,}25 < \Delta < +4{,}63$	44
Peggiorato	$\Delta \geq +4{,}63$	20

Raccolta dati

39 variabili ricevute o ricavate dallo smartwatch con frequenza giornaliera:
- 3 variabili di attività motoria
- 14 variabili di frequenza cardiaca
- 10 variabili di metabolismo
- 12 variabili di stress

Lo stress e il metabolismo sono indici stimati dal dispositivo (da attività motoria e variabilità della frequenza cardiaca)
Le variabili possono essere correlate.

Pre-processamento e controllo qualità:
- Esclusione dei giorni con uso <14 h
- media settimanale
- interpolazione del primo ordine su un massimo di due settimane consecutive mancanti



Week #	steps	distanceInMeters	...
1	5436	7645	...
2	8644	10064	...
...	...	...	...

Ogni partecipante: matrice `weeks x 39`
`weeks`: settimane di monitoraggio (circa 13)

Il cuore della data science: dai dati all'informazione

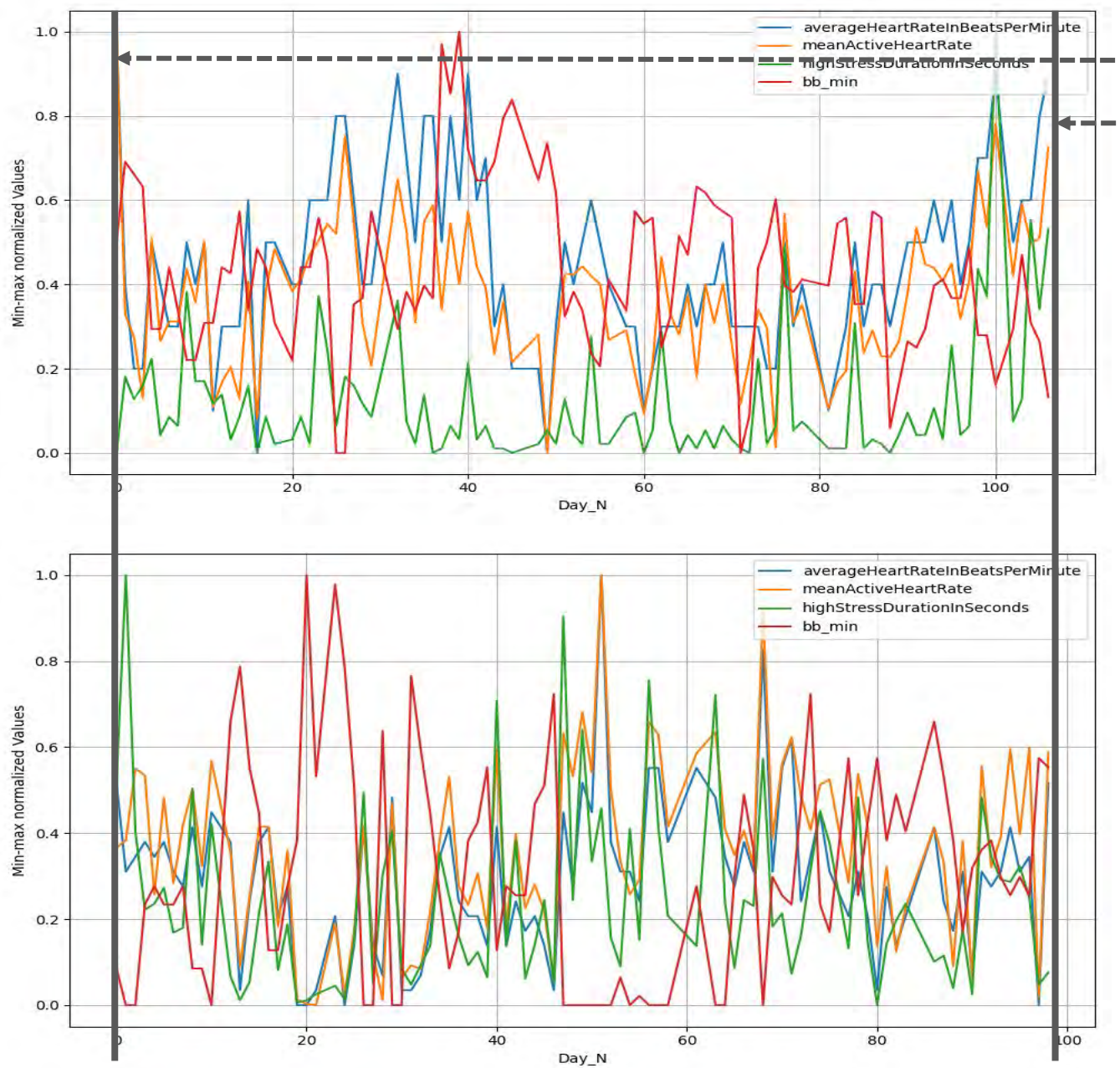
user1

+5 punti

...

userN

-10 punti



t0

t3

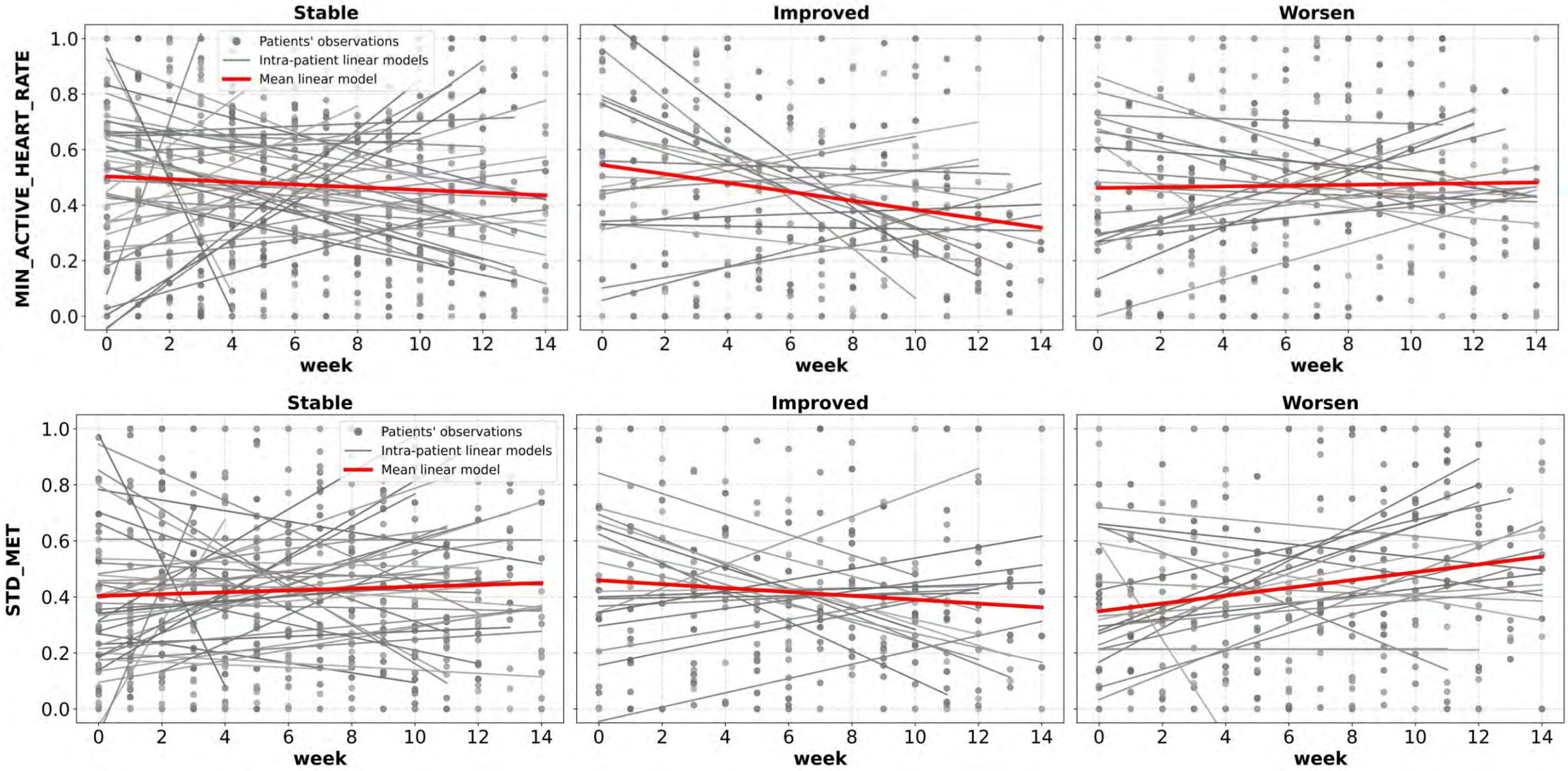


UPDRS score



OBIETTIVO: ci sono relazioni tra le ***traiettorie*** registrate a casa nei 3 mesi e il ***cambiamento clinico*** riscontrato dai medici?

Analisi esplorativa: regressione lineare



Analisi esplorativa: correlazione e analisi statistica

Matrice di coefficienti angolari: $A = [a_{pj}] \in \mathbb{R}^{P \times 39}$
P: numero dei partecipanti, j = 1…39 sono le 39 variabili estratte dallo smartwatch

Matrice di correlazione di Spearman su A

RIDONDANZA: alcune variabili sono fortemente correlate

A priori feature selection: $\rho < 0.9$

Rimangono 34 features

Kruskal-Wallis e post-hoc Dunn's

tra ogni colonna $a_{\cdot j}$ e outcome clinico categorico

pochi confronti risultano statisticamente significativi: transitori e fluttuazioni della MdP possono influenzare il regressore

occorre calcolare lo slope in modo da massimizzare il contenuto informativo e il rapporto segnale/rumore

Creazione del dataset

$$\left\{\boldsymbol{X}^{(p)}\right\}_{p=1}^{P}, \qquad \boldsymbol{X}^{(p)} \in \mathbb{R}^{weeks \times 34}$$

Per ognuna delle 34 variabili, si calcola lo slope sulle seguenti finestre temporali:
- prima / ultima settimana
- prima / penultima settimana
- prima / terzultima settimana

Ogni partecipante può contribure con un massimo di 3 osservazioni al dataset finale:

$$\{\mathbf{X}_slope^{(p)}\}_{p=1}^{P}, \quad \mathbf{X}_slope^{(p)} = [x_slope_{kj}^{(p)}] \in \mathbb{R}^{3x34}$$

- a tutte le osservazioni dello stesso paziente si assegna lo stesso target (categorico o continuo)
- lo slope dataset finale è composto di 242 osservazioni: 11 partecipanti hanno dati solo in 2 delle ultime 3 settimane, e uno soltanto in una settimana

Fuzzy Mixture of Experts

La relazione tra trend fisiologici ed esito può dipendere dal profilo clinico.

Informazione	Uso durante l'inferenza
Profilo clinico	Determina i pesi degli esperti
Trend dello smartwatch	Ogni esperto stima un ΔUPDRS-III
Somma pesata	Produce un'unica previsione continua

DATI CLINICI A T0:
Età
Durata di malattia
BMI
MoCA
Hoehn e Yahr

Fuzzy C-means:
ogni paziente p può appartenere a più profili clinici k.

$u_{pk} \in [0,1]$ $\sum_k u_{pk} = 1$

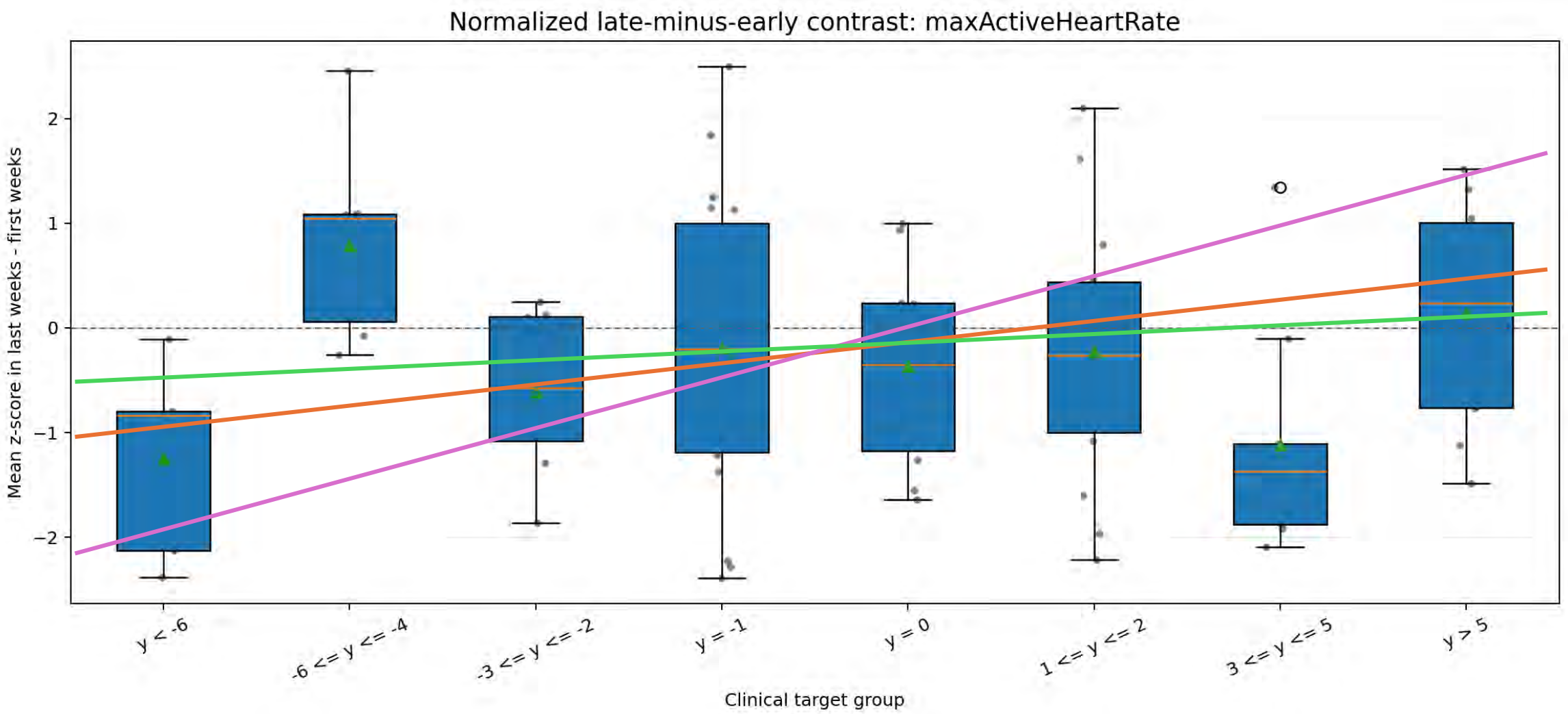
$X_slope \in \mathbb{R}^{242\times34}$

K weighted LASSO experts:
Tutti i pazienti contribuiscono a tutti gli esperti
Le membership u_{pk} pesano il contributo degli esperti per la predizione del paziente p

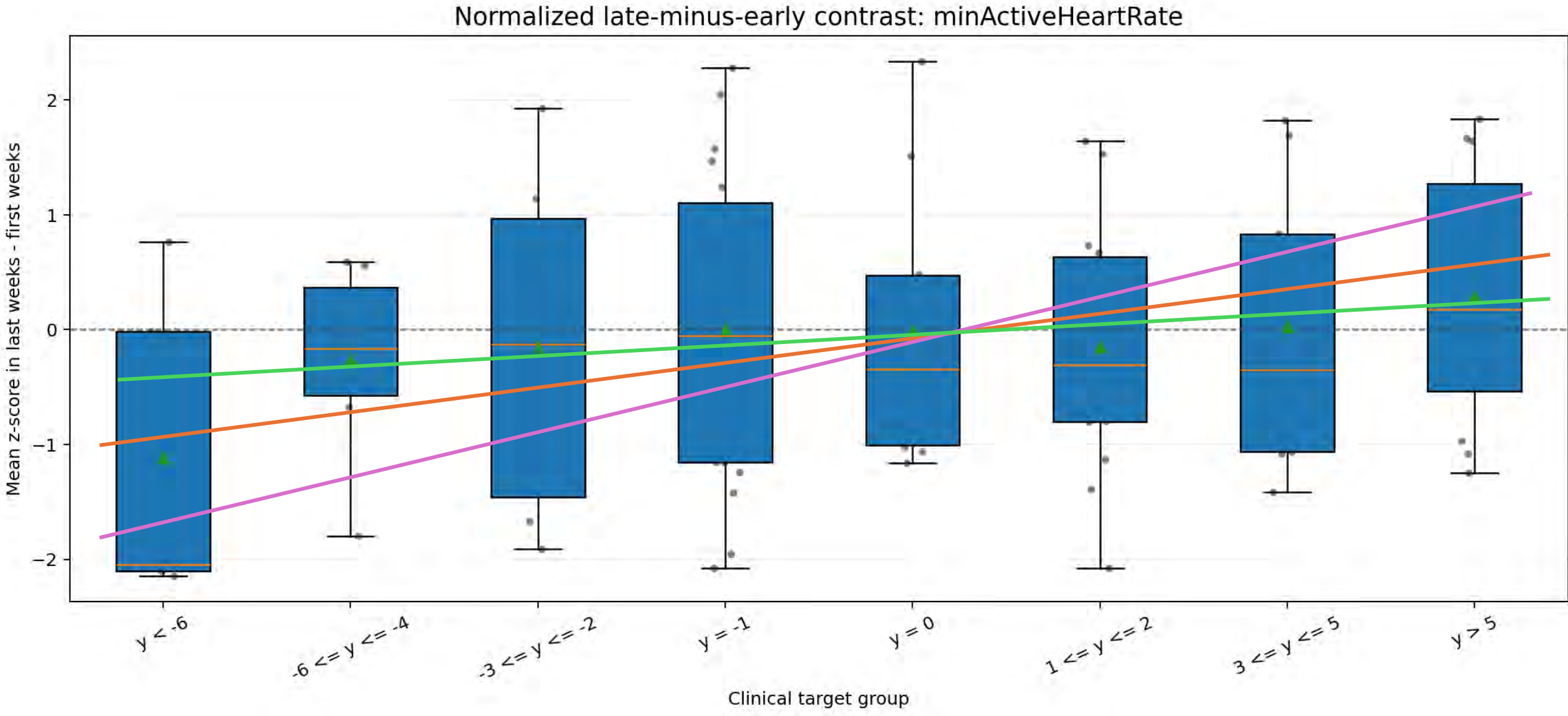
$\hat{y}_i = \sum_k u_{p(i)k}\, \hat{y}_{ik}$

Solo dopo l'aggregazione: migliorato, stabile o peggiorato con le soglie MCID.

Fuzzy Mixture of Experts: intuizione



Fuzzy Mixture of Experts: intuizione



Ogni esperto è una regressione LASSO pesata

$$\hat{y}_{ik} = \beta_{0k} + x_i^T \beta_k$$

Errore quadratico pesato + $\lambda \|\beta_k\|_1$

Elemento	Significato
x_i	Le 39 slope standardizzate
$w_{ik} = u_{p(i)k}$	Più membership, più influenza sull'esperto k
Penalità λ	Riduce i coefficienti e ne azzera alcuni

Target continuo: variazione della MDS-UPDRS Parte III

Leave-Some-Patients-Out Nested Cross-Validation

Tutte le osservazioni dello stesso paziente rimangono nella stessa partizione.

20 ripetizioni con diversi stati iniziali

Outer loop: 17 fold → 5 pazienti esclusi per volta

Inner loop: 16 fold → hyper-parameter optimization

Inner loop training: 15 fold
- scaler
- cluster fitting
- experts training

scaler parameters
centroids
experts coefficients

inner loop validation

$K \in \{2, 3, 4, 5\}$
$m \in \{1{,}5;\ 1{,}7;\ 2{,}0;\ 2{,}2\}$
$\lambda \in \{10{-}4,\ 10{-}3,\ 10{-}2,\ 10{-}1,\ 1\}$

→ Configuration with best MAE

Re-fit on entire
outer loop training
(16 fold)

scaler parameters
centroids
experts coefficients

outer loop test

I confronti con Random Forest e LASSO globale

Baseline	Che cosa confrontiamo	Metriche
Random Forest	Classificazione dell’esito	Accuracy
LASSO globale	Contributo del Fuzzy MoE alla regressione	MAE, RMSE, R^2

Stessi fold, stessi pazienti e stesse osservazioni di valutazione

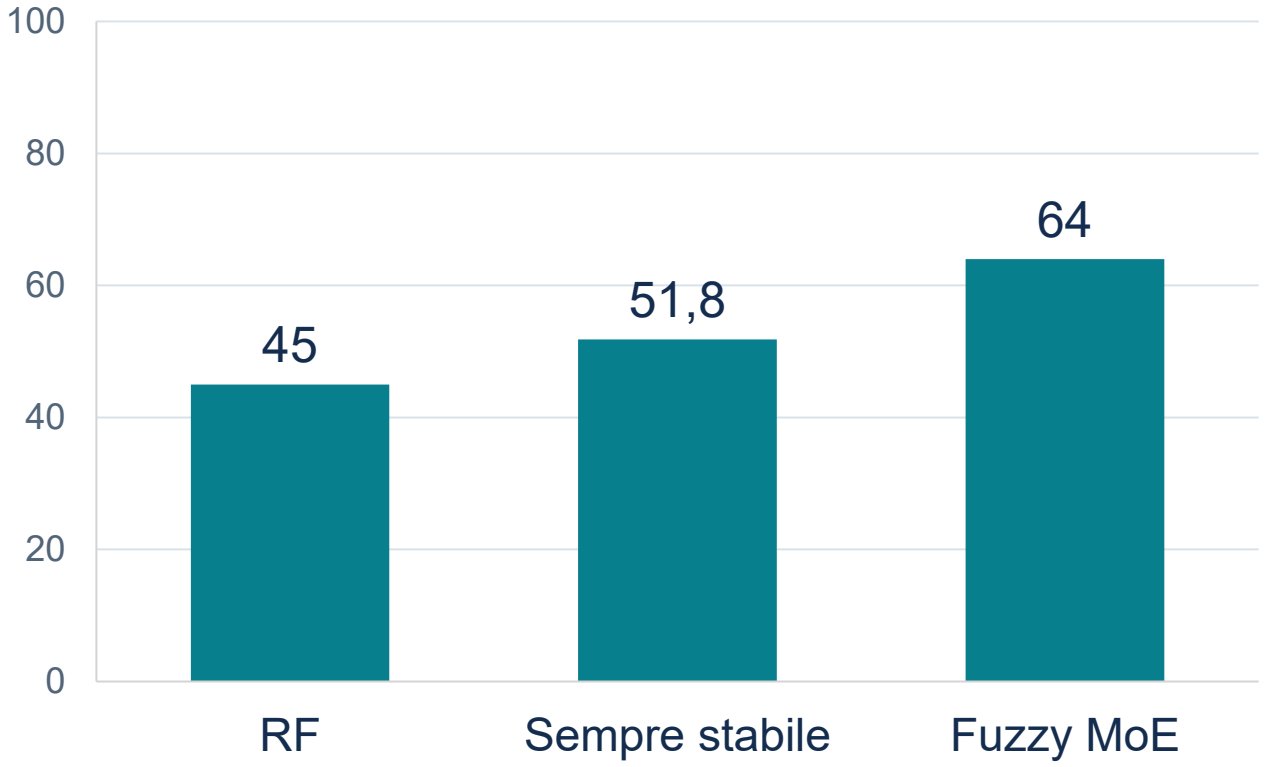
Regressione: errore minore del LASSO globale

Modello	MAE	RMSE	R^2
LASSO	3,5 ± 1,4	4,4 ± 1,7	−1,0 ± 3,8
Fuzzy MoE	3,0 ± 1,1	3,8 ± 1,4	−0,7 ± 3,5

MAE: −0,5 punti, circa −14% rispetto al LASSO

R^2 medio ancora negativo e molto variabile.
MAE < MCID non garantisce affidabilità sul singolo paziente.

Classificazione: Fuzzy MoE 64%, RF 45%



Accuracy
RF 45 ± 8%
Fuzzy MoE 64 ± 4%

Fuzzy MoE
Balanced accuracy 63 ± 8%
Macro-F1 62 ± 6%

+19 punti percentuali rispetto a RF nella stessa validazione

Gli errori cambiano tra le classi

Esito reale nelle righe, previsione nelle colonne

	Migliorato	Stabile	Peggiorato
Migliorato (n=21)	**8.7 (1.8)**	11.2 (1.8)	1.2 (0.9)
Stabile (n=44)	2.9 (1.4)	**29.7 (3.6)**	11.4 (3.2)
Peggiorato (n=20)	0.1 (0.3)	3.9 (1.2)	**16.0 (1.3)**

Il miglioramento viene spesso classificato come stabilità.
Circa un quarto degli stabili viene classificato come peggiorato.

Fuzzy MoE: quattro profili clinici sovrapposti

Refit descrittivo sull’intero dataset: K = 4, m = 1,7, λ = 0,1

Profilo	Età	BMI	Durata	H&Y	MoCA
0	Media	Alto	Media	Medio	Media
1	Media	Basso	Alta	Alto	Bassa
2	Alta	Basso	Bassa	Medio	Media
3	Bassa	Medio	Bassa	Basso	Alta

Membership massima media: 0,59
Solo 28/85 pazienti raggiungono almeno 0,70

IEEE_CEM, Tab. 4 e interpretazione post-hoc, pp. 7–8

Gli esperti: importanza e selezione delle feature

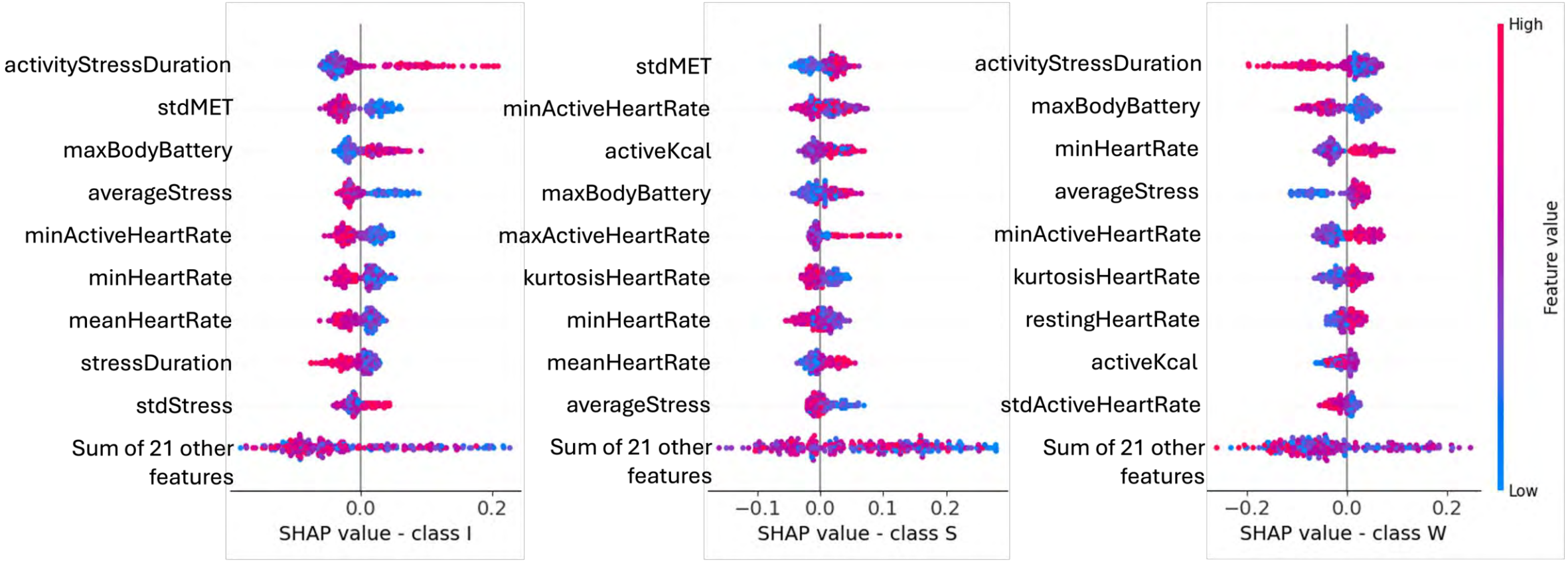
Solo 14 delle 34 features sono usate dagli esperti. Ogni esperto seleziona un massimodi 12 feature.
Associazioni condizionate al modello: gli esperti pesano diversamente una stessa slope feature

Feature	E0	E1	E2	E3
Kurtosis HR	−1,29	−0,96	−0,75	−0,77
HR a riposo	+0,79	+0,62	+0,80	+0,23
Stress medio	+0,96	**+3,01**	+0,71	+1,07
Durata attività	−1,22	0,00	−0,92	−1,25
HR minima	0,00	−0,27	+0,43	+0,21
DS stress	−0,79	−1,06	−0,99	−0,44
HR max in attività	+1,22	+2,18	+0,32	+0,05

Interpretazione clinica dei coefficienti dei regressori LASSO
β positivo → feature tende ad aumentare su pazienti peggiorati
β negativo → feature tende a diminuire su pazienti peggiorati

SHAP: quali trend guidano il Random Forest?

Attività, frequenza cardiaca e stress contribuiscono alle decisioni del modello.



Principali risultati e prossimi passi

I trend quotidiani contengono informazioni sull’esito motorio.

Il Fuzzy MoE ottiene risultati migliori delle baseline e rende leggibili profili e contributi delle feature.

Evidenza preliminare: 85 pazienti, R^2 medio negativo,
previsione vicina al follow-up e assenza di validazione esterna.

Prossimi passi: coorte esterna, verifica prospettica e utilità clinica.
Previsione precoce ed esecuzione edge restano da dimostrare.