



Federated
HAR

HAR

Federated
Learning

Dataset

Modellazione

Risultati

Approccio
centralizzato
Approccio federato

Conclusioni

Federated Human Activity Recognition

Dario Comanducci

Master in Data Science and Statistical Learning
MD2SL 2023/24

Università degli Studi di Firenze
Scuola IMT Alti Studi Lucca

30 Aprile 2025

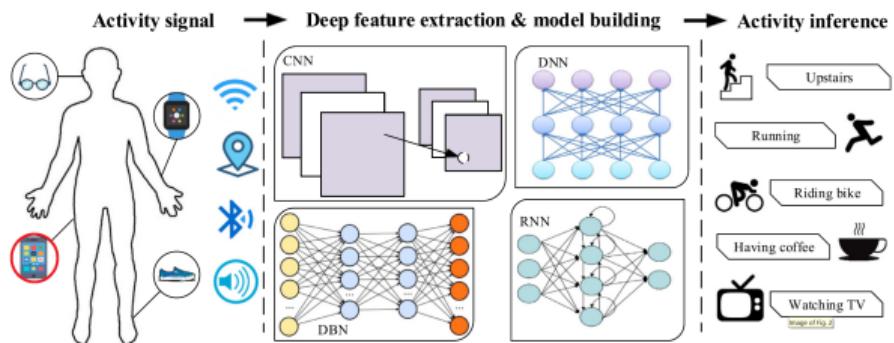
Human Activity Recognition



- ▶ La *Human Activity Recognition* (**HAR**) è un ramo di ricerca volto a riconoscere le attività umane utilizzando segnali di varia natura
- ▶ **Smartphones** come *sensori*: accelerometri e giroscopi
- ▶ Tecniche di Deep Learning

Problematiche

- ▶ problemi di **privacy**
- ▶ problemi di data **storage** centralizzati
- ▶ annotazioni spesso mancanti



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

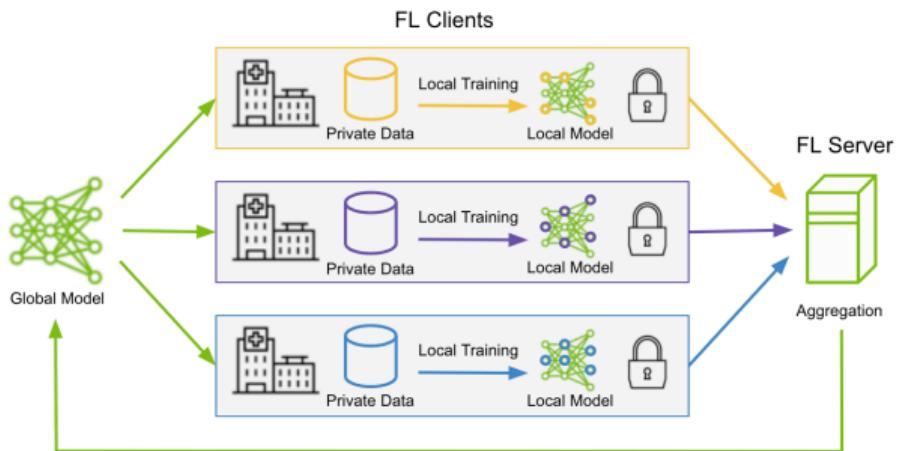
Approccio centralizzato
Approccio federato

Conclusioni



Federated Learning

- ▶ framework di addestramento *non* centralizzato
- ▶ utilizzo di dispositivi decentralizzati (**clients**) che mantengono i dati sui quali vengono eseguiti addestramenti in locale,
- ▶ gli addestramenti vengono comunicati ad un **server**, senza la necessità di scambiare i dati stessi.



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato
Approccio federato

Conclusioni



Federated Averaging (FedAvg)

Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni

Lato server:

```

1 for  $t = 1 \dots R$  do
2    $m \leftarrow \max(\gamma \cdot K, 1)$ 
3    $\mathcal{S}_t \leftarrow$  (random set of  $m$  clients)
4   for  $k \in \mathcal{S}_t$  in parallel do
5      $\mathbf{w}_{t+1}^{(k)} \leftarrow \text{ClientUpdate}(k, \mathbf{w}_t)$ 
6      $n_k \leftarrow |\mathcal{P}_k|$  ( $\mathcal{P}_k$ : client $_k$  examples)
7   end
8    $n \leftarrow \sum_{k=1}^m n_k$ 
9    $\mathbf{w}_{t+1} \leftarrow \frac{1}{n} \sum_{k=1}^m n_k \mathbf{w}_{t+1}^{(k)}$ 
10 end
```

γ : frazione dei K client per round

Lato client: ClientUpdate(k, \mathbf{w})

```

1  $\mathcal{B} \leftarrow$  split local data  $\mathcal{P}_k$  into batches
  of size  $B$ 
2 for  $i = 1 \dots E$  do
3   for batch  $b \in \mathcal{B}$  do
4      $\mathbf{w} \leftarrow \mathbf{w} - \eta \nabla L_k(\mathbf{w}; b)$ 
5   end
6 end
7 return  $\mathbf{w}$  to server
```

η : learning rate



Problematiche

- ▶ Svariati dataset da enti diversi
- ▶ Differenti tipologie di azioni
- ▶ Differenti sensori impiegati
- ▶ Differenti unità di misura

Su oltre 40 dataset valutati in Napoli et al. (2024), **6 con categorie condivise** e segnali acquisiti tramite **accelerometri e giroscopi degli smartphones**:

- ▶ sit/stand
- ▶ walk/run
- ▶ stairs up/down

I dataset:

- ▶ Ku-HAR (KH)
- ▶ MotionsSense (MS)
- ▶ RealWorld-T/W (RW-T/W)
- ▶ UCI (UCI) **no run**
- ▶ WISDM (WDM) **no stairs**

Standardizzazione

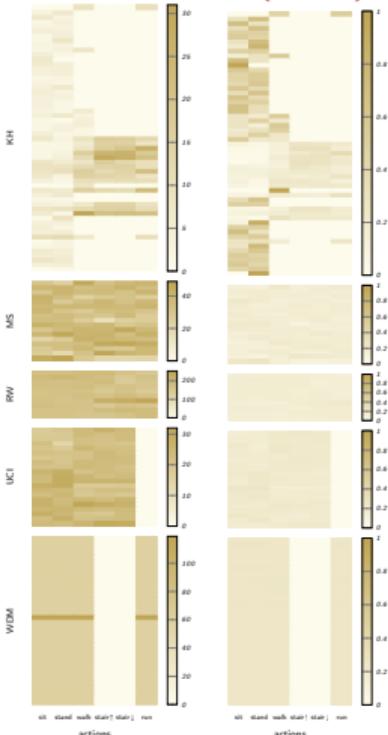
- ▶ stesso campionamento (20 Hz)
- ▶ stesse unità di misura (m/s^2 e rad/s)
- ▶ chunk di 3 s per azione



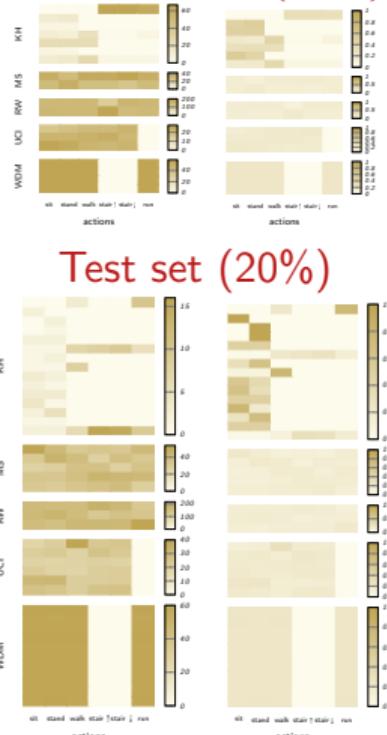
II Benchmark DAGHAR

Training/Validation/test set

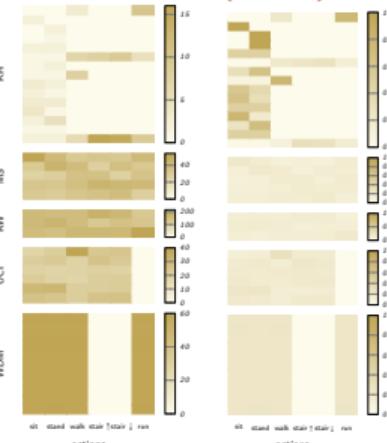
Training set (70%)



Validation set (10%)



Test set (20%)



utenti

dataset	train.	valid.	test	tot.
KH	57	7	15	79
MS	17	2	5	24
RW	10	2	3	15
UCI	21	3	6	30
WDM	36	4	11	51

esempi

dataset	train.	valid.	test	tot.
KH	1386	144	426	1956
MS	3558	420	1062	5040
RW	8400	1764	2628	12792
UCI	2420	340	690	3450
WDM	8736	944	2596	12276

Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

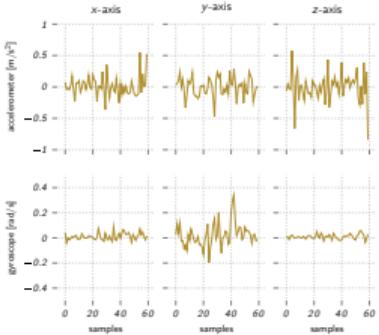
Approccio federato

Conclusioni

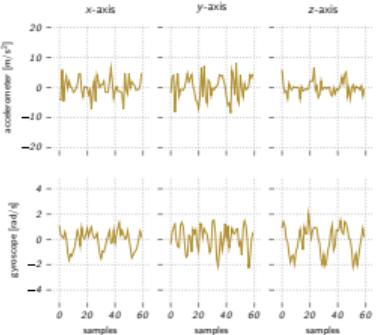
II Benchmark DAGHAR

Esempi dei dati raw

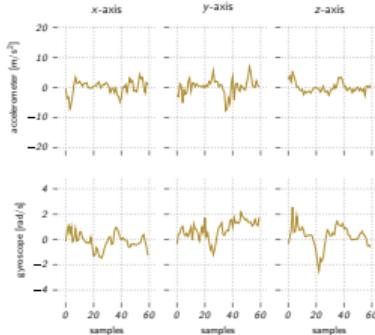
sit



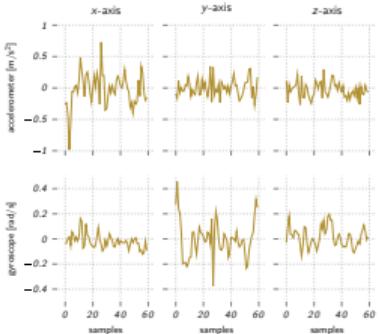
walk



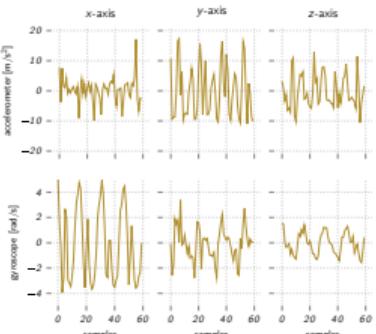
stair↑



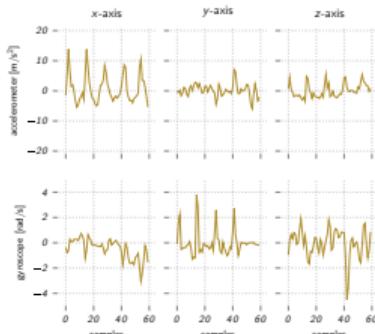
stand



run



stair↓



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Scenario supervisionato

Architettura del classificatore (label *one-hot encoding*)
dati → RNN/LSTM → embedding → MLP → softmax

Mitigazione overfitting
dropout + layer normalization

Loss: Cross-Entropy

$$L(\hat{\mathbf{y}}_n, \mathbf{y}_n)_{n=1}^N = \frac{\sum_{n=1}^N \ell(\hat{\mathbf{y}}_n, \mathbf{y}_n)}{N}$$
$$\ell(\hat{\mathbf{y}}_n, \mathbf{y}_n) = - \sum_{c=1}^C y_{nc} \log \text{softmax}(\hat{y}_{nc})$$

Possibili configurazioni

- ▶ RNN/LSTM:
mono/bi-direzionali,
1 o 2 strati interni
(64 o 128 neuroni)
- ▶ embedding:
ultimo output o
ultimo stato interno
del modello ricorrente
- ▶ MLP:
1 strato interno
(64 o 128 neuroni)
- ▶ dropout:
0, 0.1 ... 0.9

Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Prove sperimentali

Federated HAR

Approccio centralizzato

- ▶ selezione del learning rate
- ▶ selezione della migliore architettura
- ▶ analisi delle matrice di confusione

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato
Approccio federato

Approccio federato

- ▶ impiego della miglior architettura centralizzata per server e client
- ▶ test con fino a 20 client *uniformi* (DAGHAR completo)
- ▶ test con client reali (solo RW-T)
- ▶ analisi delle metriche aggregate e per client sul modello migliore

Conclusioni



Selezione del learning rate

Esperimenti centralizzati

Algoritmo di selezione

- ▶ Addestramento in una **singola** epoca
- ▶ ad ogni batch
 - ▶ incremento **graduale** del learning rate
 - ▶ calcolo della loss raggiunta
- ▶ Il punto a **massimo decremento** della loss indica il learning rate più promettente

Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni

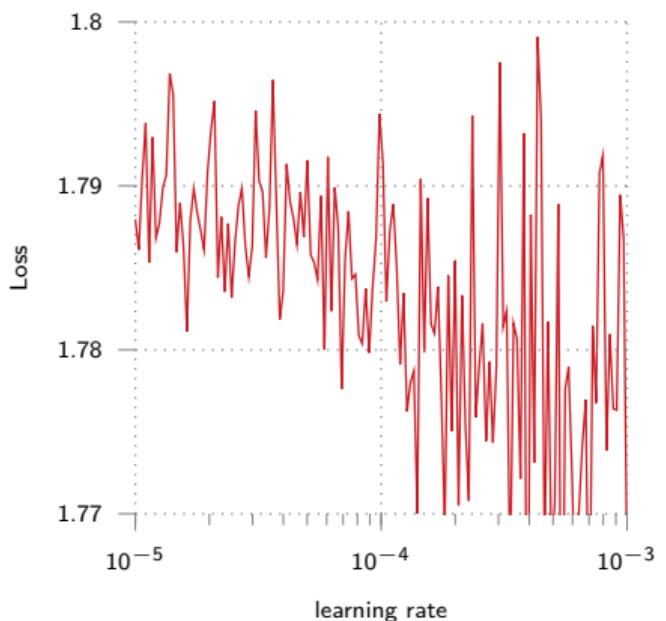


Selezione del learning rate

Esperimenti centralizzati

Algoritmo di selezione

- ▶ Addestramento in una **singola** epoca
- ▶ ad ogni batch
 - ▶ incremento **graduale** del learning rate
 - ▶ calcolo della loss raggiunta
- ▶ Il punto a **massimo decremento** della loss indica il learning rate più promettente



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Selezione del learning rate

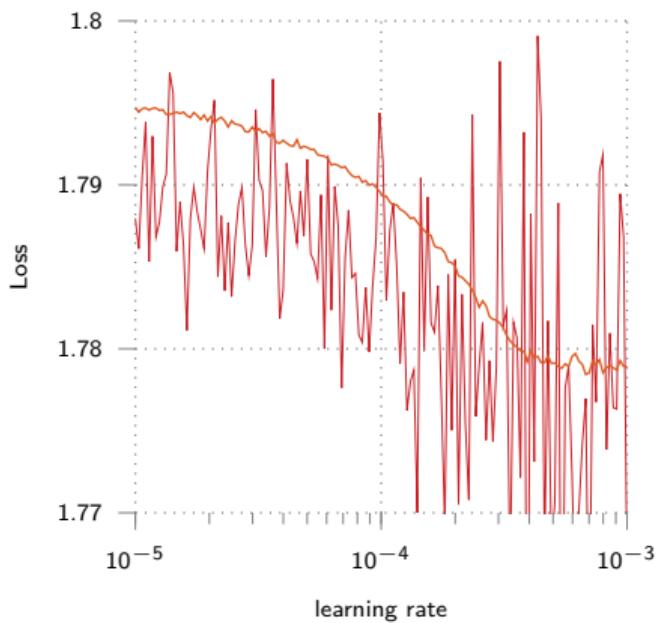
Esperimenti centralizzati

Algoritmo di selezione

- ▶ Addestramento in una **singola** epoca
- ▶ ad ogni batch
 - ▶ incremento **graduale** del learning rate
 - ▶ calcolo della loss raggiunta
- ▶ Il punto a **massimo decremento** della loss indica il learning rate più promettente

Versione *robusta*

- ▶ ripetere il range test N volte
- ▶ mediare le loss ottenute
 - ▶ smoothing (Savitzky-Golay, 3° ordine)
 - ▶ derivata alle differenze centrali (3° ordine)
 - ▶ max. drec. \equiv minimo assoluto della derivata



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Selezione del learning rate

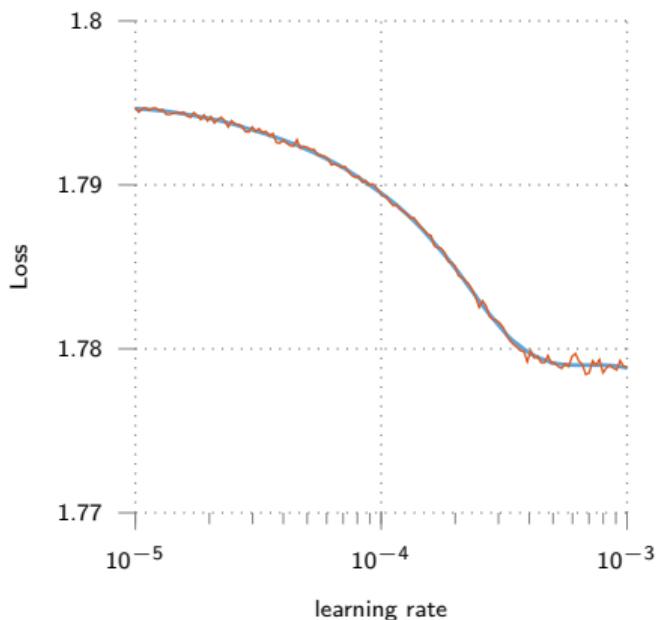
Esperimenti centralizzati

Algoritmo di selezione

- ▶ Addestramento in una **singola** epoca
- ▶ ad ogni batch
 - ▶ incremento **graduale** del learning rate
 - ▶ calcolo della loss raggiunta
- ▶ Il punto a **massimo decremento** della loss indica il learning rate più promettente

Versione *robusta*

- ▶ ripetere il range test N volte
- ▶ mediare le loss ottenute
- ▶ smoothing (Savitzky-Golay, 3° ordine)
- ▶ derivata alle differenze centrali (3° ordine)
- ▶ max. drec. \equiv minimo assoluto della derivata



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Selezione del learning rate

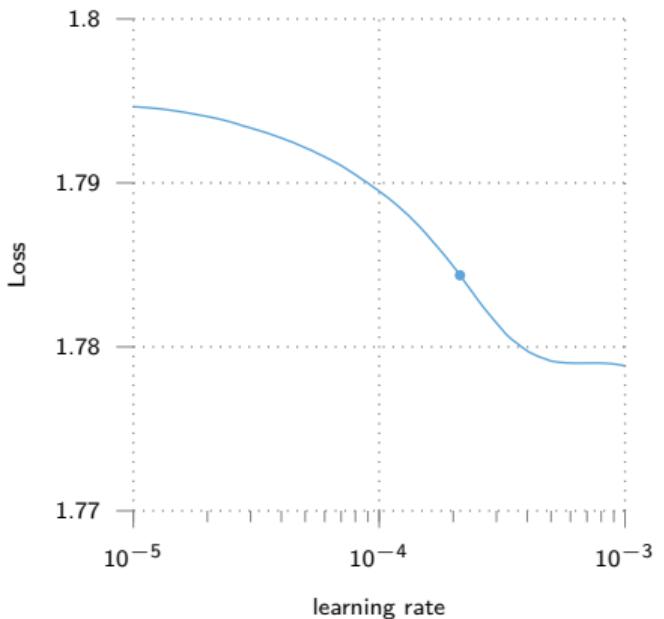
Esperimenti centralizzati

Algoritmo di selezione

- ▶ Addestramento in una **singola** epoca
- ▶ ad ogni batch
 - ▶ incremento **graduale** del learning rate
 - ▶ calcolo della loss raggiunta
- ▶ Il punto a **massimo decremento** della loss indica il learning rate più promettente

Versione *robusta*

- ▶ ripetere il range test N volte
- ▶ mediare le loss ottenute
- ▶ smoothing (Savitzky-Golay, 3° ordine)
- ▶ derivata alle differenze centrali (3° ordine)
- ▶ **max. drec.** \equiv minimo assoluto della derivata



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Selezione del learning rate

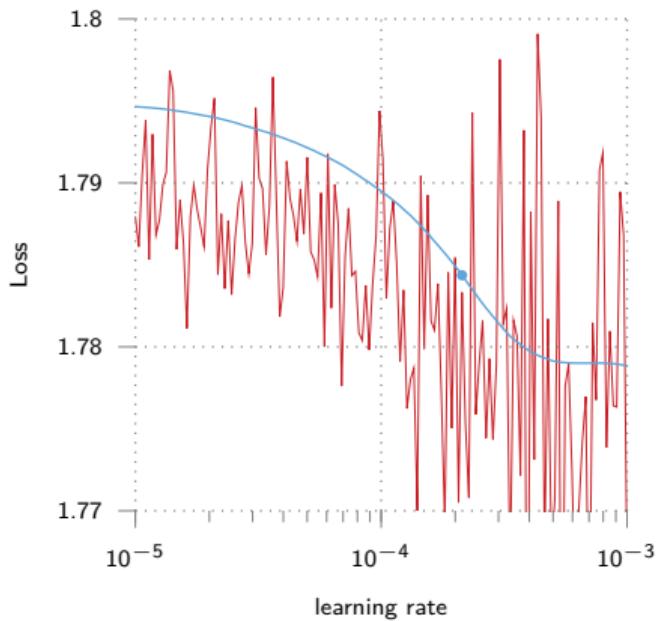
Esperimenti centralizzati

Algoritmo di selezione

- ▶ Addestramento in una **singola** epoca
- ▶ ad ogni batch
 - ▶ incremento **graduale** del learning rate
 - ▶ calcolo della loss raggiunta
- ▶ Il punto a **massimo decremento** della loss indica il learning rate più promettente

Versione *robusta*

- ▶ ripetere il range test N volte
- ▶ mediare le loss ottenute
- ▶ smoothing (Savitzky-Golay, 3° ordine)
- ▶ derivata alle differenze centrali (3° ordine)
- ▶ **max. drec.** \equiv minimo assoluto della derivata



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Modello migliore

Esperimenti centralizzati

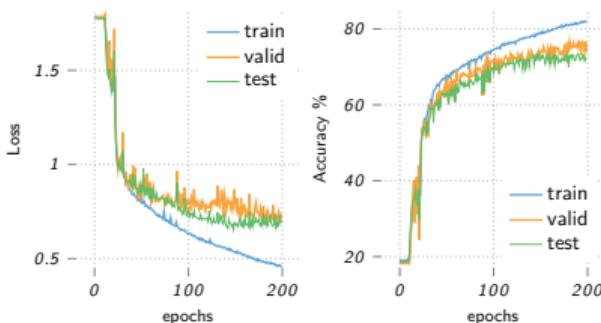
embedding: ultimo hidden state

architettura	l.r. · 10^4	accuracy %			
		train.	valid.	test	epoca
RNN _{64,1} MLP ₆₄	2.131	54.1	54.7	53.5	188
RNN _{64,1} MLP ₁₂₈	1.874	52.8	53.0	52.0	189
RNN _{128,1} MLP ₆₄	1.647	41.8	43.1	42.7	193
RNN _{128,1} MLP ₁₂₈	1.273	44.4	44.6	44.6	193
LSTM _{128,1} MLP ₁₂₈	2.201	73.8	69.3	68.1	168
LSTM _{128,2} MLP ₁₂₈	1.595	76.2	71.0	68.3	150
LSTM _{256,1} MLP ₁₂₈	1.595	77.3	71.6	71.3	159
LSTM _{64,2} MLP ₆₄	2.131	76.2	70.2	66.8	168
LSTM_{64,2}MLP₁₂₈	2.201	79.4	73.3	71.4	168
LSTM _{128,2} MLP ₆₄	1.647	78.4	70.4	69.5	182
LSTM _{128,2} MLP ₁₂₈	1.402	77.6	69.9	67.8	190
LSTM _{256,1} MLP ₁₂₈	1.233	69.7	65.4	64.8	194
LSTM _{256,2} MLP ₁₂₈	0.953	71.1	66.9	65.7	87

Miglior architettura: LSTM bidirezionale con 2 strati nascosti da 64 unità + MLP con 128 neuroni interni

LSTM_{64,2}MLP₁₂₈ + dropout & layer norm.

dropout	accuracy %			
	train.	valid.	test	epoca
0.0	71.4	71.4	66.7	47
0.1	74.9	70.7	70.0	83
0.2	76.0	72.7	71.5	117
0.3	71.9	68.2	67.6	182
0.4	81.9	76.5	72.1	195
0.5	74.1	71.7	70.8	99
0.6	76.4	75.1	71.0	153
0.7	71.8	72.6	68.2	144
0.8	75.5	70.8	70.9	186
0.9	62.2	61.2	60.1	149



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

Approccio federato

Conclusioni



Matrici di confusione

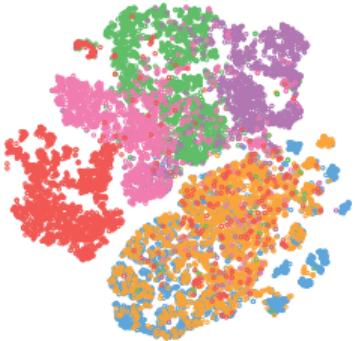
Confronto con t-SNE

Train. set

	sit	stand	walk	stair↑	stair↓	run
sit	0.75	0.29	0	0.01	0	0.02
stand	0.24	0.71	0.01	0.01	0.01	0.07
walk	0	0	0.91	0.04	0.05	0.01
stair↑	0	0	0.01	0.89	0.02	0
stair↓	0	0	0.05	0.05	0.91	0.01
run	0	0	0.01	0	0	0.88

sit stand walk stair↑ stair↓ run
ground truth

2D training embeddings

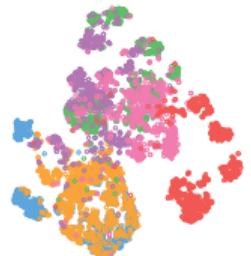


Valid. set

	sit	stand	walk	stair↑	stair↓	run
sit	0.67	0.26	0	0	0	0
stand	0.32	0.74	0.01	0.16	0.02	0
walk	0	0	0.67	0.08	0.08	0.01
stair↑	0	0	0.14	0.62	0.05	0
stair↓	0	0	0.18	0.14	0.82	0.06
run	0	0	0.01	0	0.01	0.92

sit stand walk stair↑ stair↓ run
ground truth

2D validation embeddings

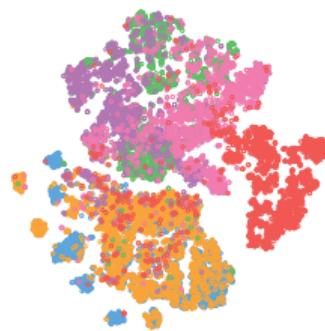


Test set

	sit	stand	walk	stair↑	stair↓	run
sit	0.64	0.4	0	0.01	0.01	0.02
stand	0.34	0.58	0.01	0.02	0.01	0.08
walk	0	0.01	0.77	0.16	0.11	0.02
stair↑	0	0.01	0.07	0.72	0.04	0.01
stair↓	0.01	0	0.16	0.08	0.83	0.04
run	0	0	0	0	0	0.84

sit stand walk stair↑ stair↓ run
ground truth

2D test embeddings



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

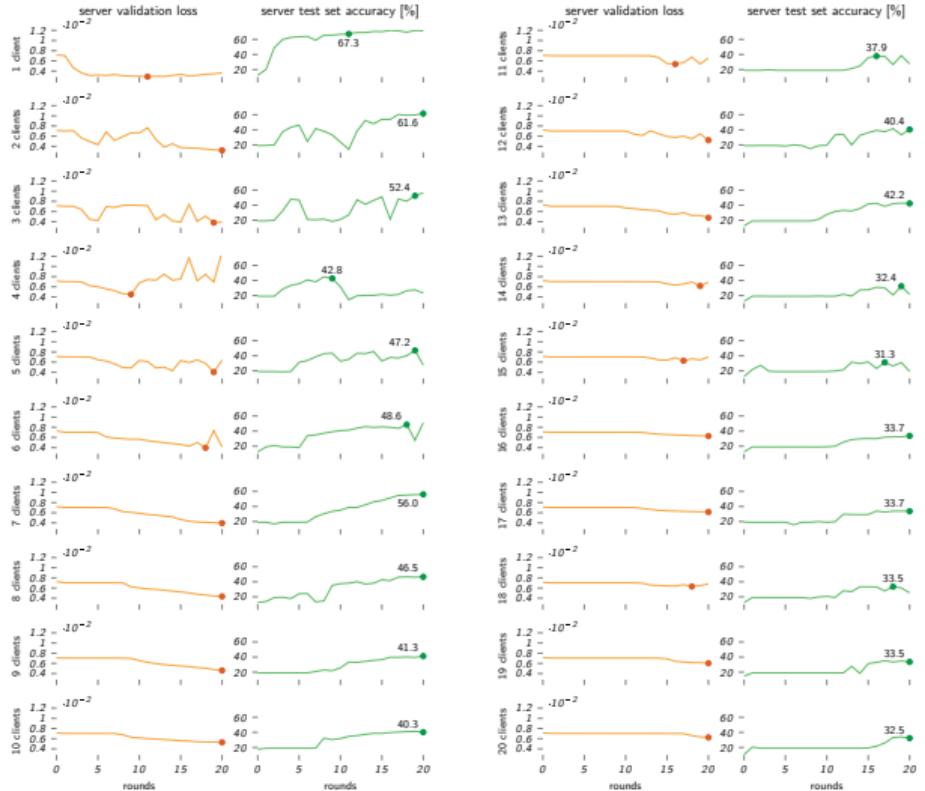
Approccio federato

Conclusioni



Con client uniformi

Esperimenti federati



10 epocha a round

Lato client: 70% training set
30% valid. set

Lato server: validation e test set
del setup centraliz.

- ▶ accuracy decrescente al crescere del numero di client
- ▶ per risultati paragonabili a pochi client, occorrono verosimilmente più round

Con RW-T

Esperimenti federati



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

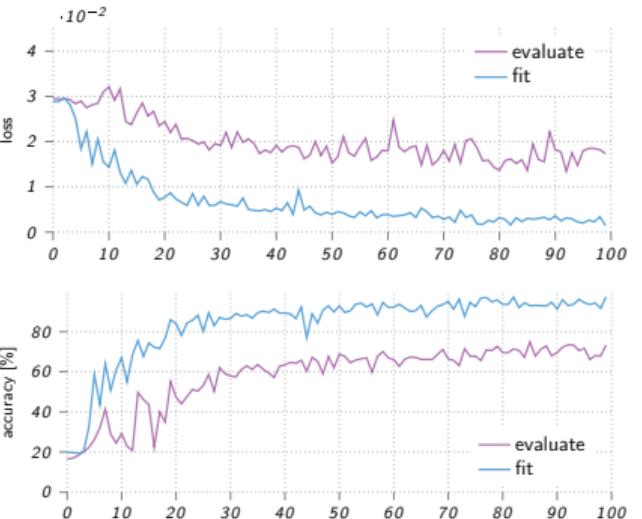
Risultati

Approccio centralizzato

Approccio federato

Conclusioni

epoch	client %	round	accuracy %			
			valid.	test	avg	aggr. eval.
5	40	98	25.2	67.4	50.6	51.4
	50	95	32.4	63.0	50.7	57.7
	60	71	40.1	54.5	48.7	45.9
10	40	52	42.9	65.6	56.6	60.4
	50	81	51.7	64.3	59.3	63.0
	60	49	36.1	57.8	49.1	57.1
20	40	77	65.5	63.4	64.3	62.3
	50	50	50.4	71.6	63.1	68.8
	60	56	52.9	64.3	59.7	66.6



Con RW-T

Esperimenti federati



Federated HAR

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato

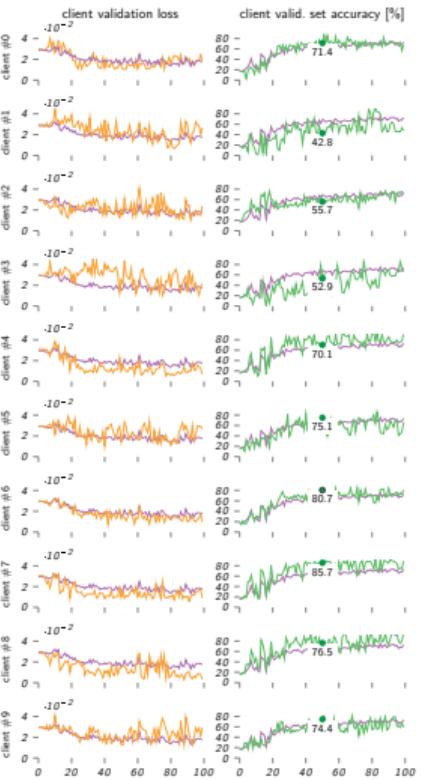
Approccio federato

Conclusioni

epoch	client %	round	accuracy %			
			valid.	test	avg	aggr. eval.
5	40	98	25.2	67.4	50.6	51.4
	50	95	32.4	63.0	50.7	57.7
	60	71	40.1	54.5	48.7	45.9
10	40	52	42.9	65.6	56.6	60.4
	50	81	51.7	64.3	59.3	63.0
	60	49	36.1	57.8	49.1	57.1
20	40	77	65.5	63.4	64.3	62.3
	50	50	50.4	71.6	63.1	68.8
	60	56	52.9	64.3	59.7	66.6

Accuracy dei client al round 50

users									
#0	#1	#2	#3	#4	#5	#6	#7	#8	#9
71.4	42.8	55.7	52.9	70.1	75.1	80.7	85.7	76.5	74.4





Conclusioni

Federated HAR

Riepilogo

Pro

- ▶ base di lavoro flessibile a vari modelli
- ▶ segnali incoraggianti per approccio federato

Contro

- ▶ necessità di prove con più round per approfondire
- ▶ necessità di setup con più utenti

Future work

- ▶ rilassamento dell'approccio supervisionato applicato allo scenario federato
- ▶ impiego di CNN o transformers al posto dello LSTM
- ▶ modelli scalabili in base alla capacità di calcolo dei device

HAR

Federated Learning

Dataset

Modellazione

Risultati

Approccio centralizzato
Approccio federato

Conclusioni