

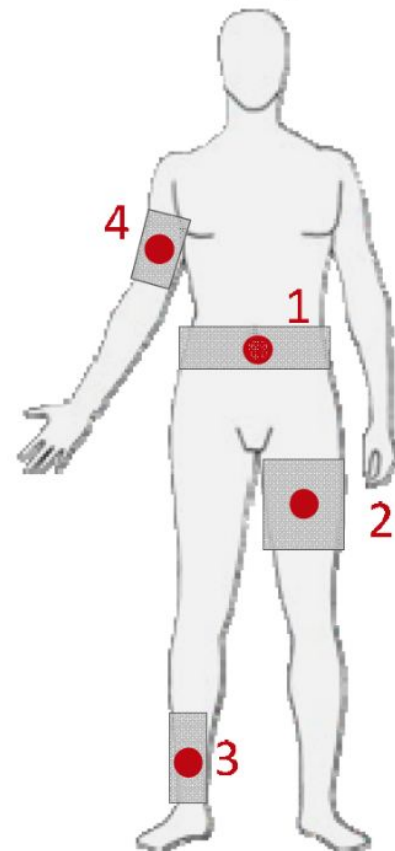
Modelli Probabilistici per le Decisioni

Amrani Hamza	807386
Brumana Mattia	808374
Carta Costantino	808417

a.a 2018/2019

Obiettivi

- Sviluppo di un modello predittivo che date in input le coordinate di quattro accelerometri posti su un soggetto riconosce l'azione che esso sta compiendo.



Dataset Iniziale (1)

- Dati raccolti durante 8 ore di attività
- 4 soggetti (due uomini e due donne)
- 165.633 campioni

Subject	Genre	Age	Height	Weight	Instances
A	Female	46 y.o.	1.62m	67kg	51,577
B	Female	28 y.o.	1.58m	53kg	49,797
C	Male	31 y.o.	1.71m	83kg	51,098
D	Male	75 y.o.	1.67m	67kg	13,161*

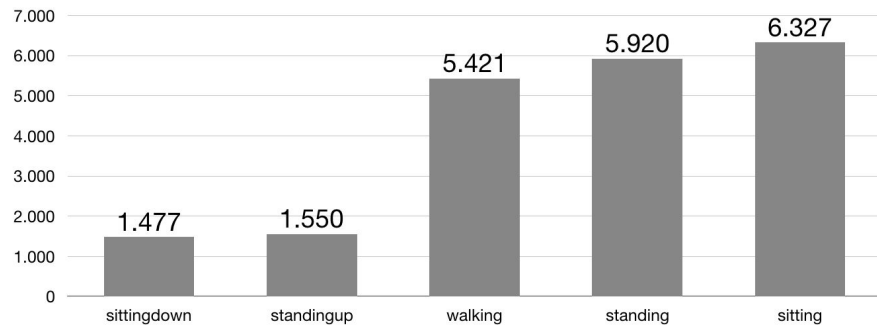
Dataset Iniziale (2)

- Dati anagrafici
- Rilevazione di sensori
- Azione compiuta

	user	gender	age	how_tall_in_meters	...	x4	y4	z4	class
0	debora	Woman	46	1,62	...	-150	-103	-147	sitting
1	debora	Woman	46	1,62	...	-149	-104	-145	sitting
2	debora	Woman	46	1,62	...	-151	-104	-144	sitting
3	debora	Woman	46	1,62	...	-153	-103	-142	sitting
4	debora	Woman	46	1,62	...	-153	-104	-143	sitting

Preprocessing - W. Ugolino *et al.*

- 20 partizioni (4 utenti per 5 classi differenti)
- per ogni accelerometro (una lettura): *roll, pitch, modulo accelerazione*
- per ogni finestra di 1 secondo (8 letture): *varianza di roll e pitch, media e deviazione standard dell'accelerazione*
- Dataset finale: 20.695 record



Normalizzazione e Discretizzazione

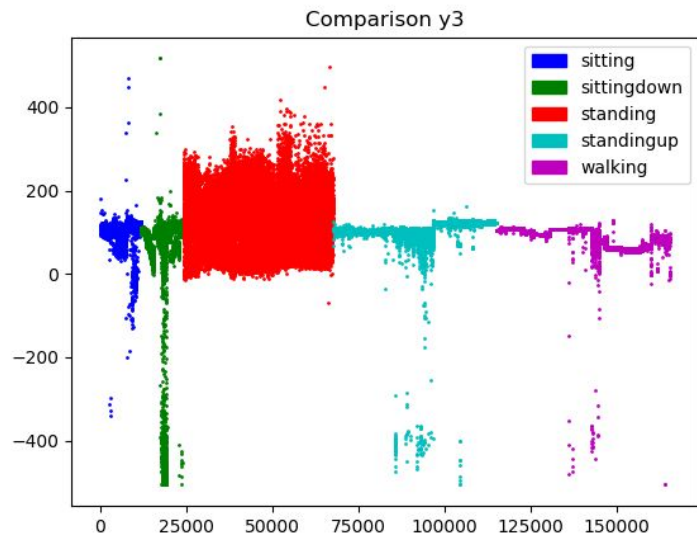
- Operazioni svolte con l'obiettivo di rendere più efficiente la distribuzione dei dati
- Normalizzazione: dati ridimensionati in intervallo $[0,1]$
- Discretizzazione: dati distribuiti nell'intervallo di interi $[0,10]$

Preprocessing - la “nostra” versione

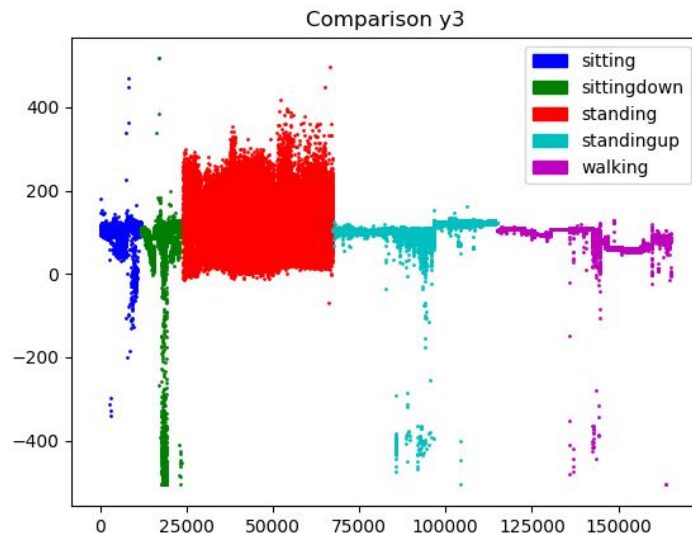
- Analisi più accurata dei valori iniziali contenuti nel dataset
- Rilevazione (e conseguente eliminazione) di alcuni record che non rispettano la distribuzione
- Dataset finale: 20.665 record

Pulizia dataset

- Originale

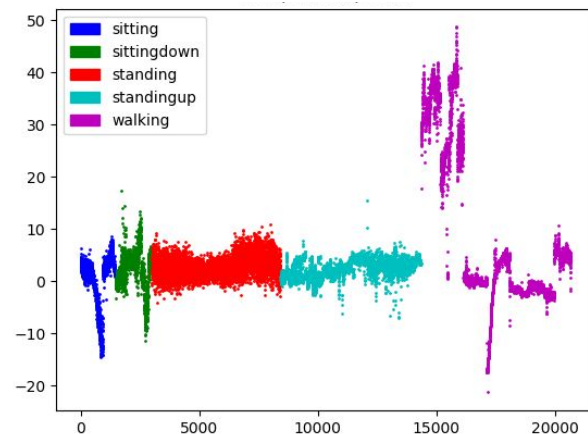
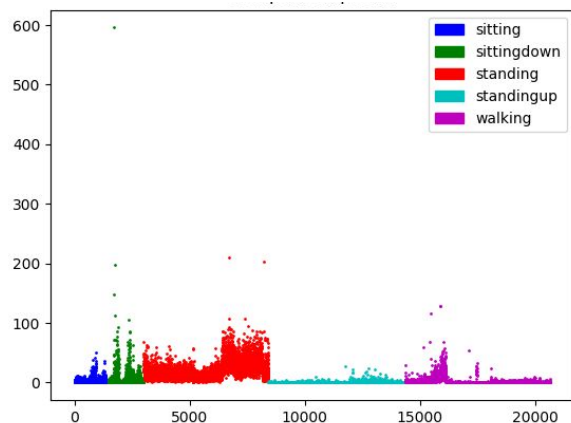


- Dopo la pulizia



Calcolo della media

Nella nostra versione sostituiamo il calcolo della varianza con il calcolo della media, in quanto quest'ultima denota una distinzione più marcata tra le classi.



Confronto tra varianza (sinistra) e media (destra) dei valori di pitch1

HILL-CLIMB

Algoritmo:

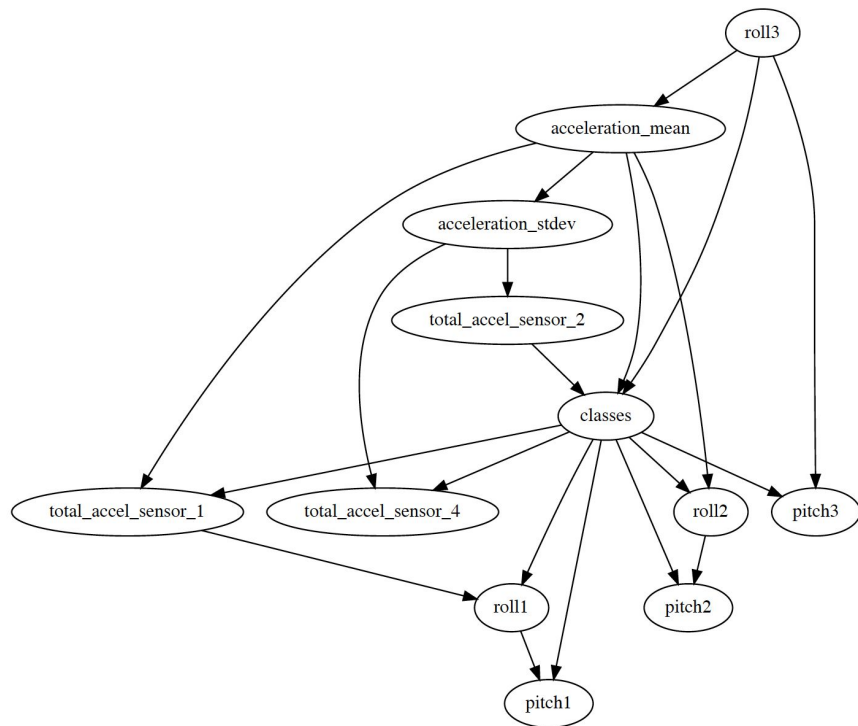
1. Inizia con una soluzione vuota o casuale (detta *best solution*);
2. Si fa una copia della soluzione, dopodiché la si muta leggermente;
3. Valutazione della nuova soluzione: se è migliore della *best solution*, quest'ultima viene sostituita con la nuova soluzione;
4. Si ritorna al punto 2. e si ripete il procedimento.

GREEDY

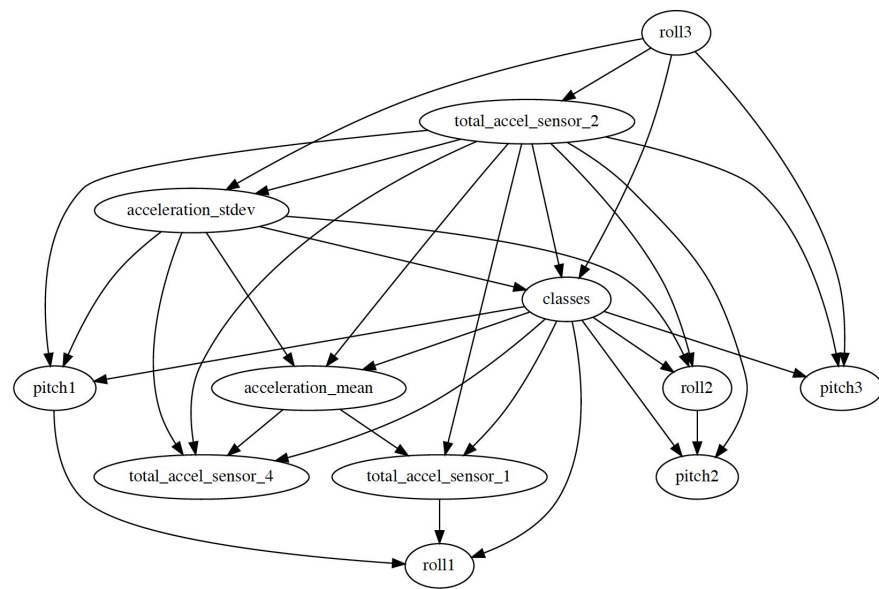
L'algoritmo greedy usato nella Pomegranate costruisce l'ordinamento topologico in maniera *greedy*, ovvero sceglie la variabile successiva migliore da aggiungere all'ordinamento già esistente ad ogni step. Trovata la struttura migliore, i parametri vengono stimati utilizzando la *Maximum Likelihood Estimation*.

PGMPY

- W. Ugolino *et al.*

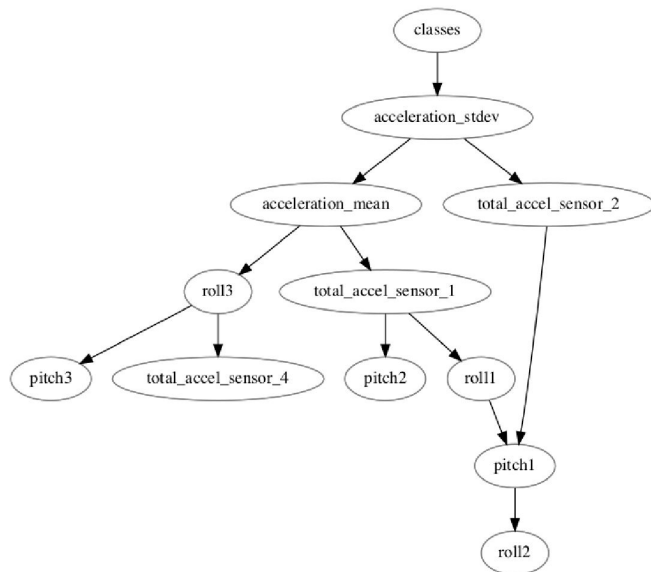


- “Nostra” versione

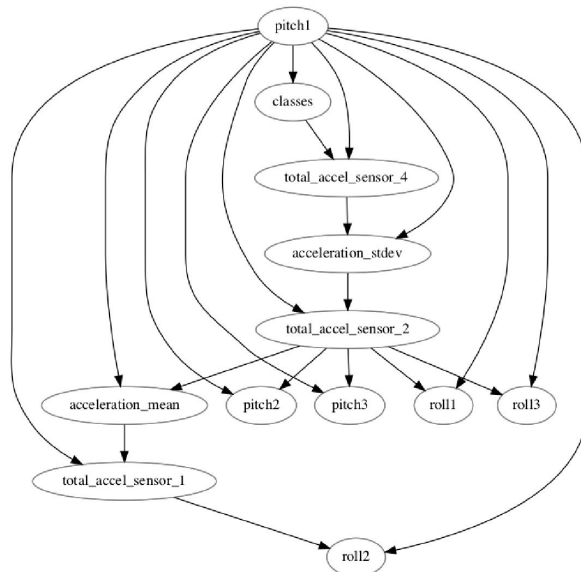


Pomegranate

- W. Ugolino *et al.*



- “Nostra” versione



Risultati Ottenuti

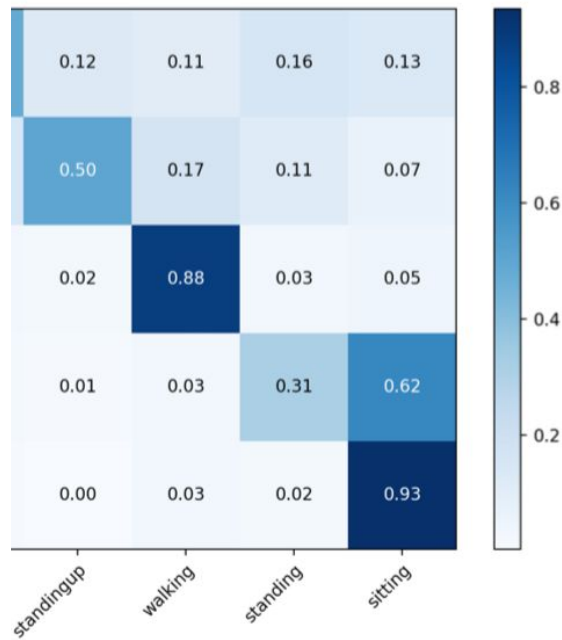
		<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Ugolino Version	<i>PGMPY</i>	0.68%	0.68%	0.62%	0.62%
	<i>POMEGRANATE</i>	0.53%	0.38%	0.41%	0.37%
Our Version	<i>PGMPY</i>	0.89%	0.86%	0.85%	0.82%
	<i>POMEGRANATE</i>	0.94%	0.89%	0.91%	0.90%

Precision = $TP / (TP + FP)$

Recall = $TP / (TP + FN)$

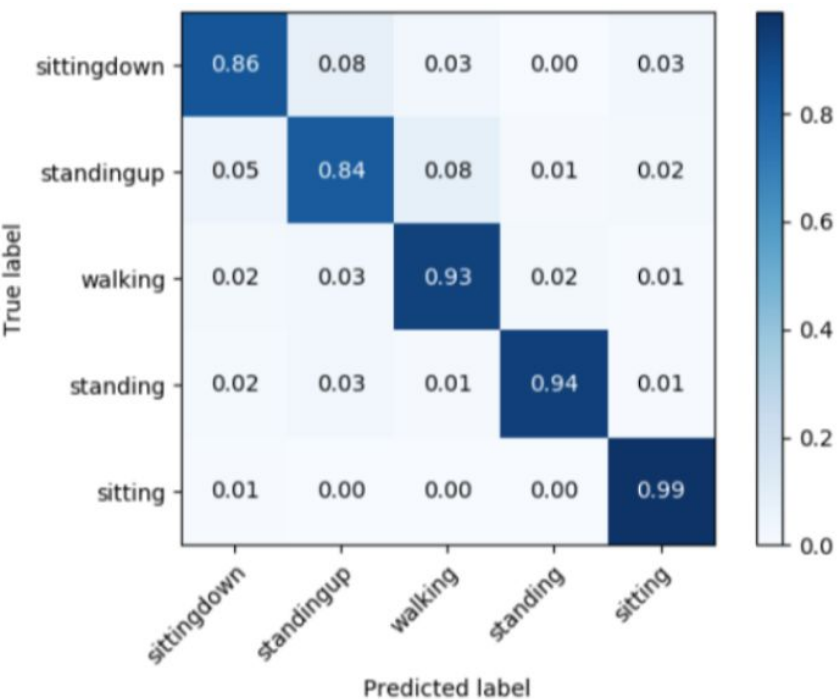
F1 Score = $2 * (Recall * Precision) / (Recall + Precision)$

W. Ugolino *et al.*



	<i>Sitting Down</i>	<i>Standing Up</i>	<i>Walking</i>	<i>Standing</i>	<i>Sitting</i>
<i>F1-Score</i>	0.51	0.58	0.87	0.43	0.71
<i>Precision</i>	0.54	0.67	0.86	0.73	0.58
<i>Recall</i>	0.48	0.50	0.88	0.31	0.93

La "nostra" versione



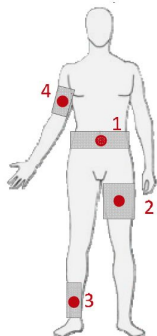
	<i>Sitting Down</i>	<i>Standing Up</i>	<i>Walking</i>	<i>Standing</i>	<i>Sitting</i>
<i>F1-Score</i>	0.83	0.78	0.94	0.96	0.98
<i>Precision</i>	0.80	0.74	0.95	0.98	0.98
<i>Recall</i>	0.86	0.84	0.93	0.94	0.98

Demo

Ambient Assisted Living Demo

This is a demonstration showing the demo of AAL(Ambient Assisted Living Demo).

Scheme of accelerometers positioning



Please select a position:

Random

Load example

	Roll	Pitch	Module acceleration vector
1 Accelerometer	150.391	1.892	179.036
2 Accelerometer	57.056	-17.033	49.425
3 Accelerometer	142.568	-4.26	49.425
4 Accelerometer	-138.606	53.816	167.703

True prediction: sittingdown

Class predicted: sittingdown

Classes	Probabilities
sittingdown	0.84
standingup	0.16
sitting	0.0
standing	0.0
walking	0.0



Conclusioni

- Reti Bayesiane: ottime prestazioni in termini di accuracy
- In particolare:
 - 68% (W. Ugolino *et al.*)
 - 94% (“nostra” versione)
- Sviluppi futuri: ampliare il dataset con rilevazioni effettuate su un numero maggiore di soggetti, con caratteristiche differenti

GRAZIE PER L'ATTENZIONE

A series of horizontal lines in teal and light blue colors, extending across the width of the slide below the text.