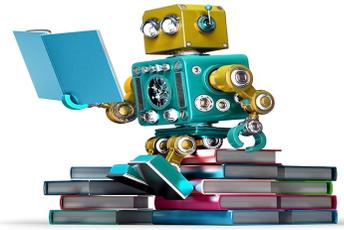




Istituto Nazionale di Fisica Nucleare



SOM neural network e applicazione ai dati Antur

Francesco Conventi^{1,2}, Agostino De Iorio^{1,3}, Biagio Rossi¹, Elvira Rossi^{1,3}

¹ Istituto Nazionale di Fisica Nucleare (INFN)

² Università degli studi di Napoli «Parthenope»

³ Università degli studi di Napoli «Federico II»



UNIVERSITÀ DEGLI STUDI DI NAPOLI FEDERICO II - DIPARTIMENTO DI
FISICA "ETTORE PANCINI"



Università degli Studi di Napoli Parthenope
dipartimento
ingegneria 

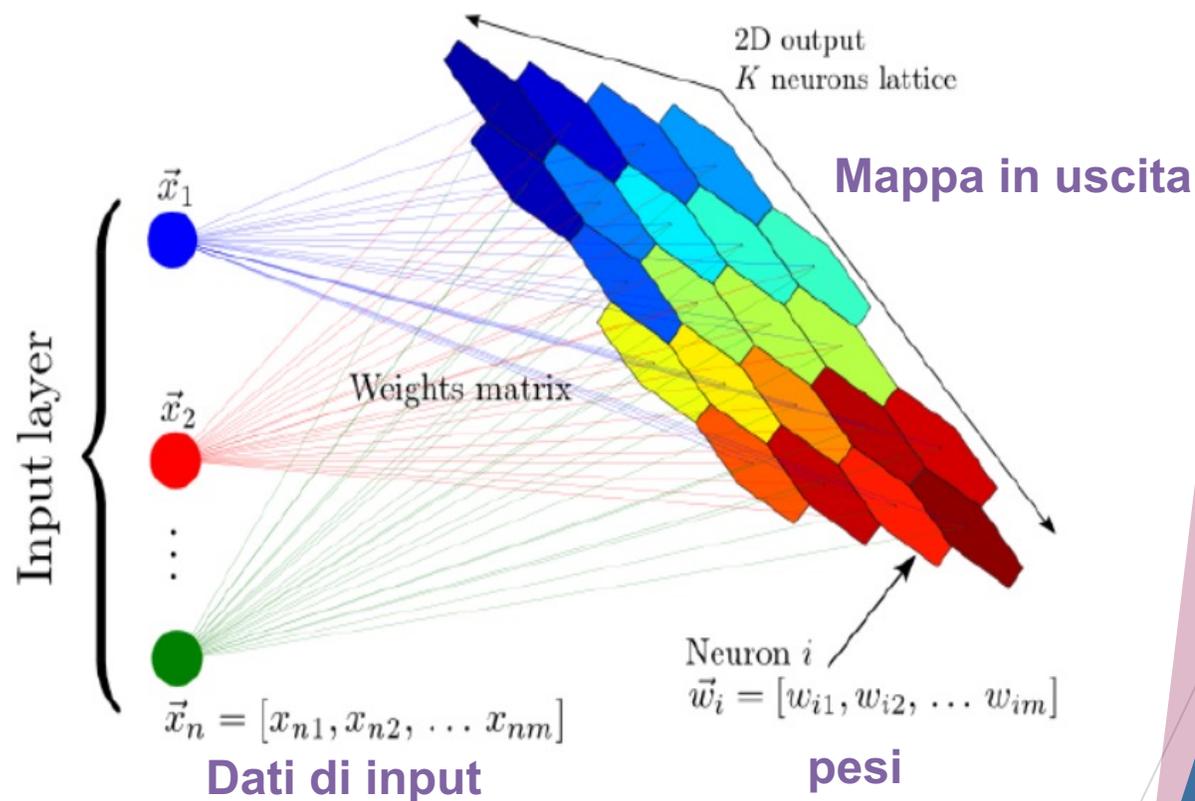
Self-Organizing feature Maps (SOM)

La **Self Organizing Map** o **mappe auto-organizzanti (SOM)** è un particolare tipo di rete neurale caratterizzato da un **apprendimento non supervisionato**.

La rete SOM analizza i dati in ingresso e li classifica ed organizza sulla base di caratteristiche comuni (**cluster**).

L'algoritmo di una SOM accetta in ingresso dati appartenenti ad uno spazio n-dimensionale e li colloca su una mappa bidimensionale (strato di Kohonen) rendendo possibile la visualizzazione e più semplice l'interpretazione dei risultati.

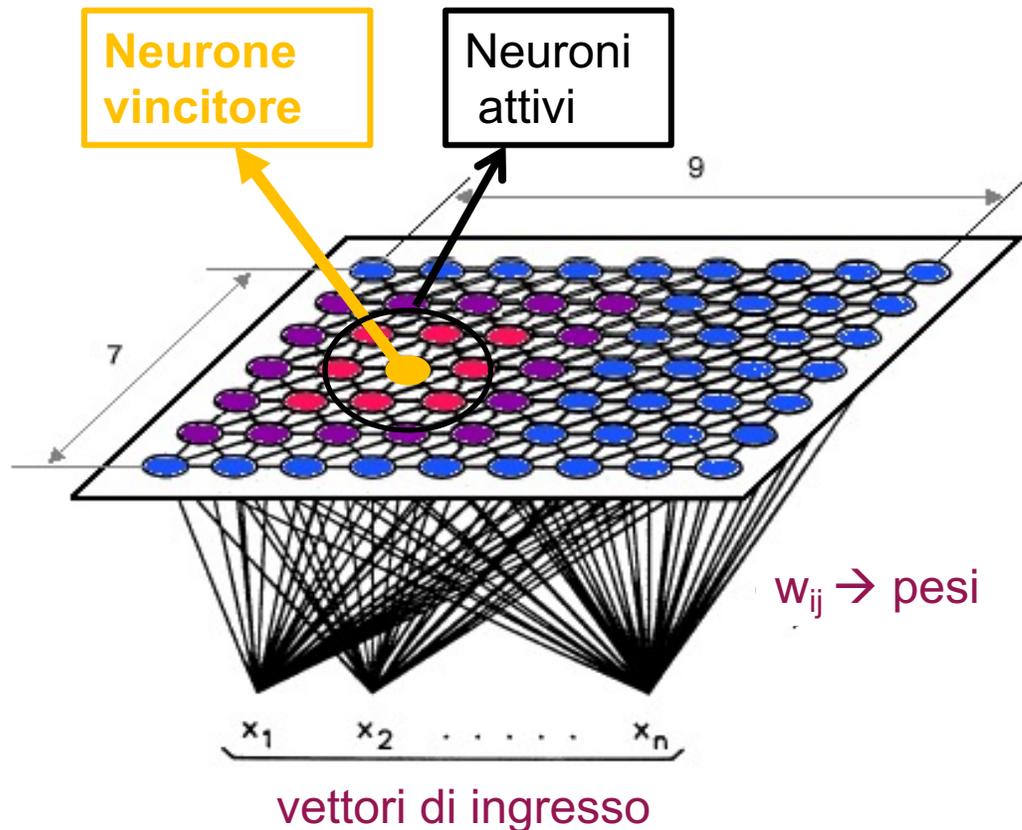
Ciascun dato di ingresso è associato a tutti i neuroni appartenenti allo strato di uscita.



Self organizing Maps (Mappe SOM)

La **Self Organizing Map** o **mappe auto-organizzanti (SOM)** è un particolare tipo di rete neurale caratterizzato da un **apprendimento non supervisionato**.

La rete SOM analizza i dati in ingresso e li classifica ed organizza sulla base di caratteristiche comuni (**cluster**).



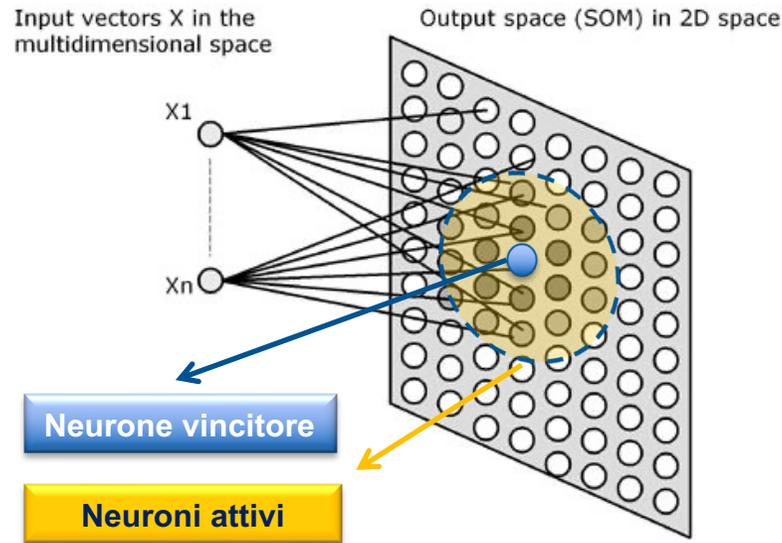
Durante la fase di apprendimento, i neuroni dello strato di Kohonen prendono parte ad un processo noto come “**winner takes all**”.

Il **neurone vincitore** è quello il cui il vettore dei pesi è più simile all’ingresso presentato.

L’apprendimento avviene modificando il peso del neurone vincitore e dei neuroni vicini e riproponendo i dati di training alla rete molte volte.

Modalità di apprendimento di una SOM

I neuroni appartenenti allo strato bi-dimensionale prendono parte a un processo di competizione detto “winner takes all” (il vincitore prende tutto).



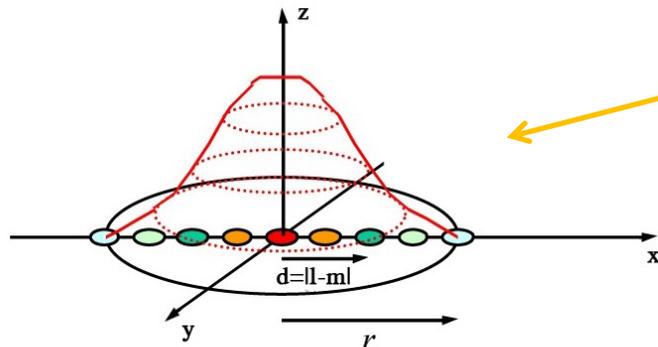
La determinazione del “neurone vincitore” avviene scegliendo il neurone con uscita massima Y . L’uscita viene rappresentata dalla seguente relazione analitica:

$$Y_j = \sum_{i=1}^n w_{ij} x_i \rightarrow \text{Vince il neurone con } Y_j \text{ massimo}$$

Il “vicinato” dei neuroni attivi si determina attraverso una funzione di vicinanza espressa da questa relazione:

$$H_{lm} = \exp\left(-\frac{\|l - m\|^2}{r^2}\right)$$

Variazione dei pesi



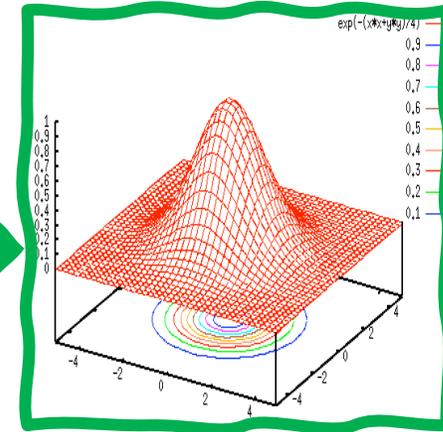
Si attivano, e quindi si modificano i pesi, solo dei neuroni contenuti entro un raggio r dal neurone vincitore.

Modalità di apprendimento di una SOM

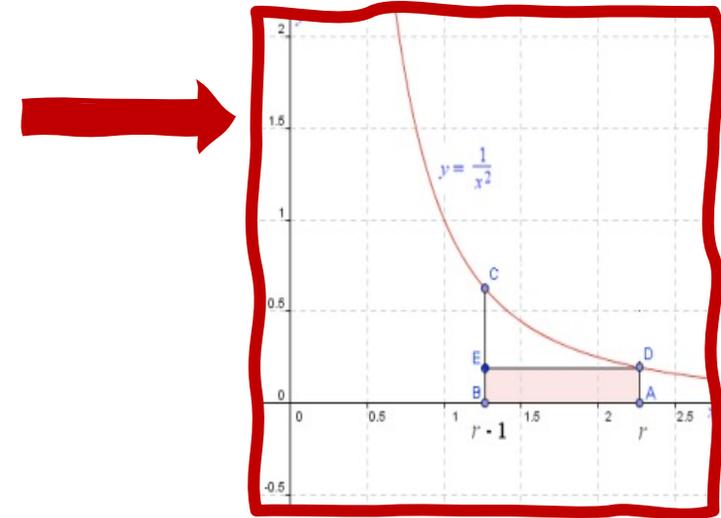
Una volta definita la mappa bidimensionale, la rete si configura tramite un processo competitivo, l'aggiornamento del peso W_n del neurone n è dato da:

$$W_n(t + 1) = W_n(t) + H(n, t) \alpha(t) (X(t) - W_n(t))$$

★ $H(n, t)$ è la funzione che definisce il vicinato



★ $\alpha(t)$ è il coefficiente di apprendimento monotono decrescente

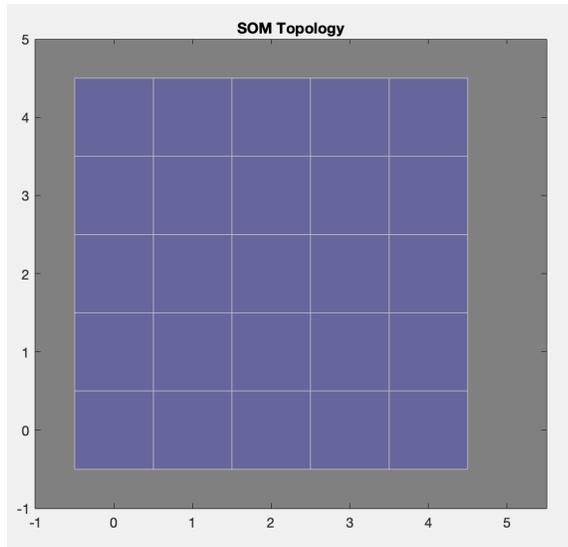


★ $[X(t) - W_n(t)]$ è la differenza tra il vettore d'ingresso ed il peso del neurone.

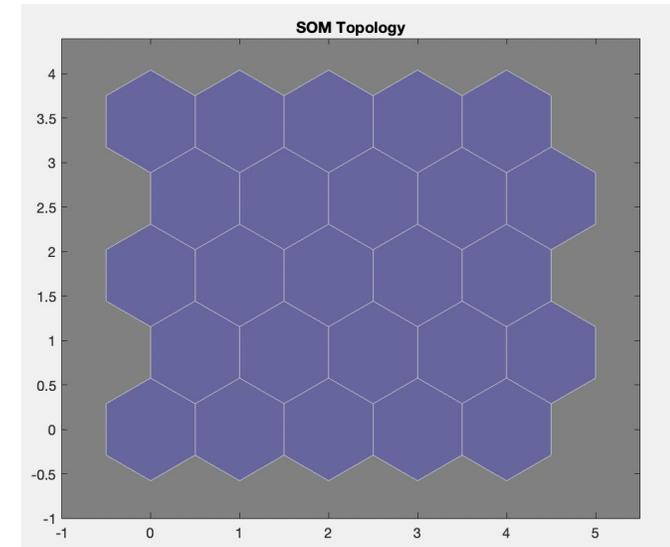
Ogni nuovo vettore d'ingresso può essere velocemente classificato collocandolo sulla mappa in uscita alla rete SOM.

Topologia di una rete SOM

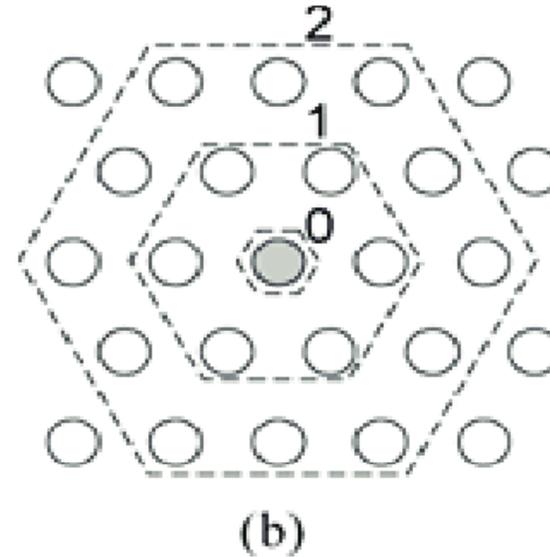
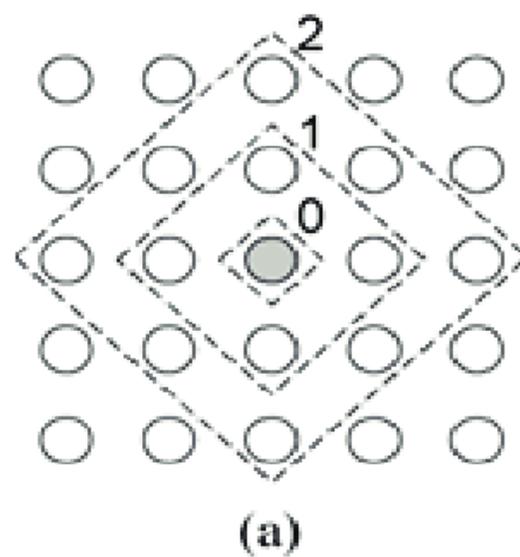
Topologia a griglia



Topologia esagonale

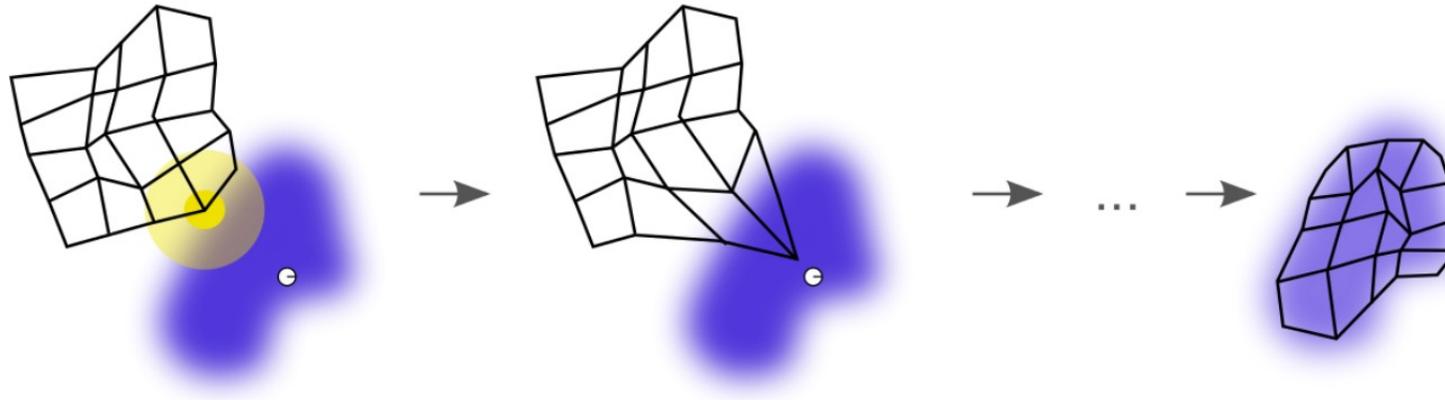


Vicinanza → scelta dei neuroni vicini



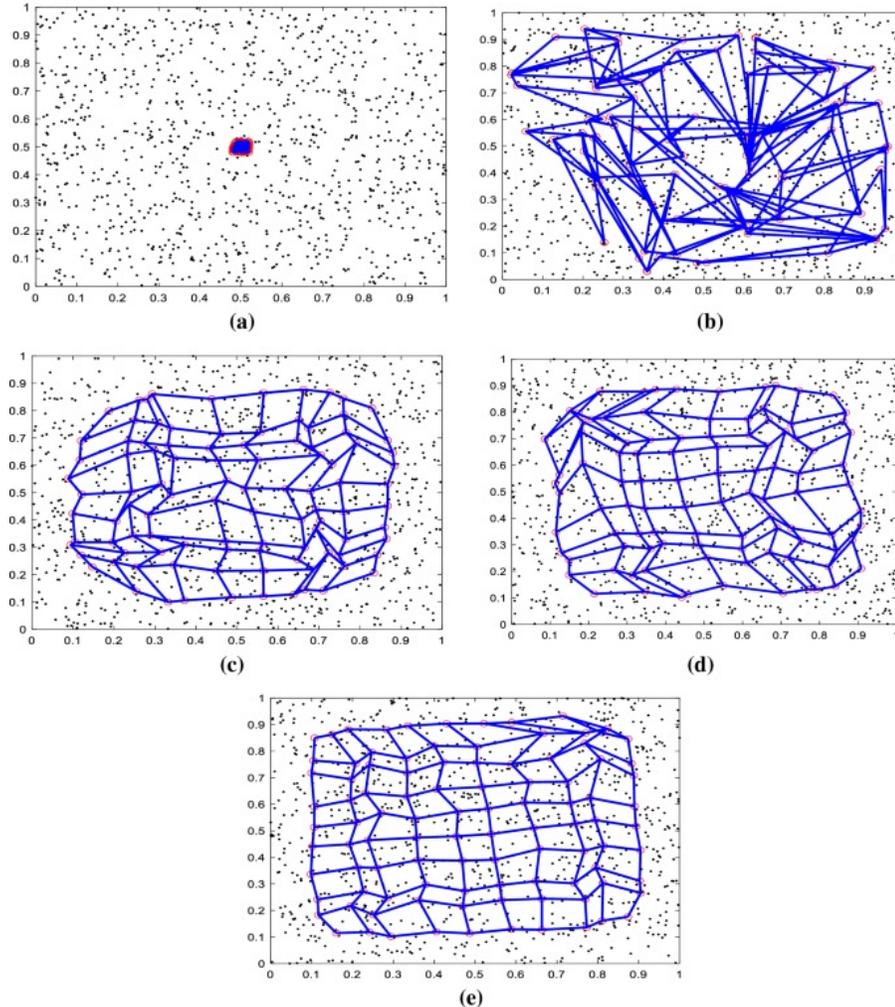
Modalità di apprendimento di una SOM

I pesi vengono modificati in base alla vicinanza dal neurone vincitore, facendo sì che i neuroni distanti abbiano aggiornamenti minori. L' algoritmo SOM si interrompe quando la mappa smette di cambiare o cambia poco.



Modalità di apprendimento di una SOM

I pesi vengono modificati in base alla vicinanza dal neurone vincitore, facendo sì che i neuroni lontani abbiano aggiornamenti minori. L' algoritmo SOM si interrompe quando la mappa smette di cambiare o cambia poco.



Neural Network Training (09-Nov-2022 16:07:45)

Network Diagram

Training Results
Epoch: 845 of 1000

Training Progress

Unit	Initial Value	Current Value	Target Value
Epoch	0	845	1000
Elapsed Time	-	00:00:18	-

Training Algorithms
Data Division: Batch Weight/Bias Rule `trainbu`
Performance: Mean Squared Error `mse`
Calculations: MATLAB

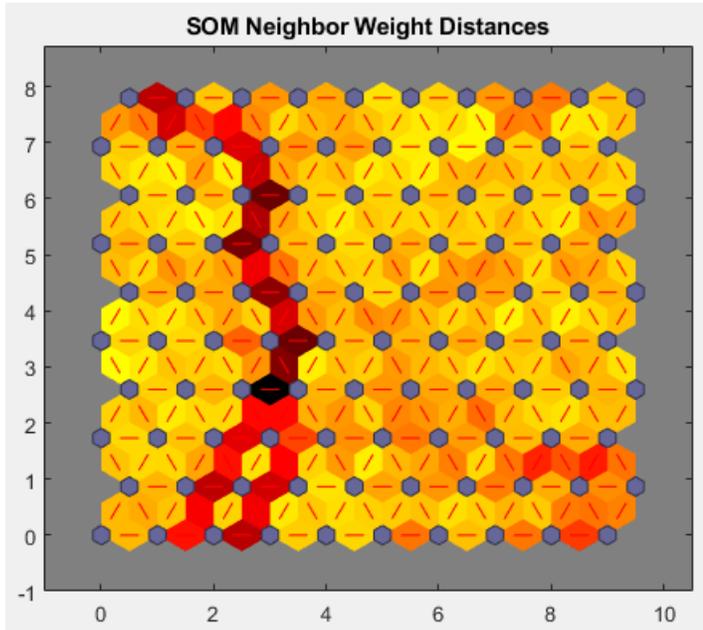
Training Plots

SOM Topology SOM Neighbor Connections
SOM Neighbor Distances SOM Input Planes
SOM Sample Hits SOM Weight Positions

Mappa di uscita di una SOM e mappa delle Hit

Esempio: topologia è una griglia 10x10

Esempio di mappa in uscita alla rete SOM



In questa figura gli **esagoni grigi** rappresentano i **neuroni**.

Le **linee rosse** collegano i neuroni vicini.

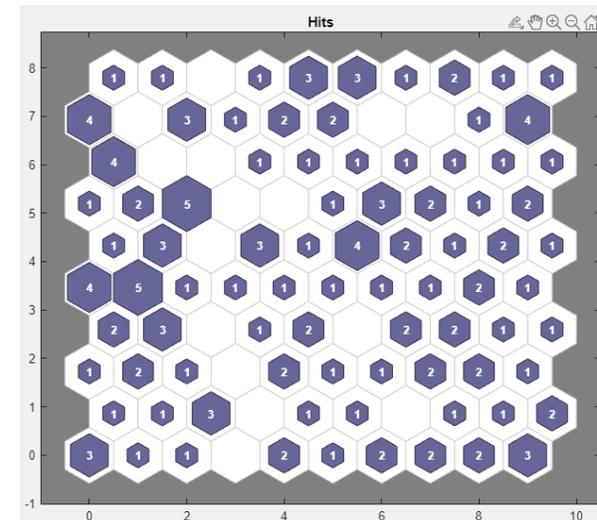
I colori nelle regioni contenenti righe rosse indicano le distanze tra i neuroni.

I colori più scuri rappresentano distanze maggiori, mentre **i colori più chiari rappresentano distanze minori**.

Nell'esempio: la mappa è attraversata da una fascia di segmenti scuri. La rete SOM sembra avere raggruppato i dati in due grandi gruppi distinti.

Questa figura mostra le posizioni dei neuroni all'interno della topologia e indica **quante delle osservazioni/dati sono associate ad ogni neurone**. Il numero massimo di hit associati a qualunque neurone è 5. Pertanto ci sono 5 vettori di input in quel cluster.

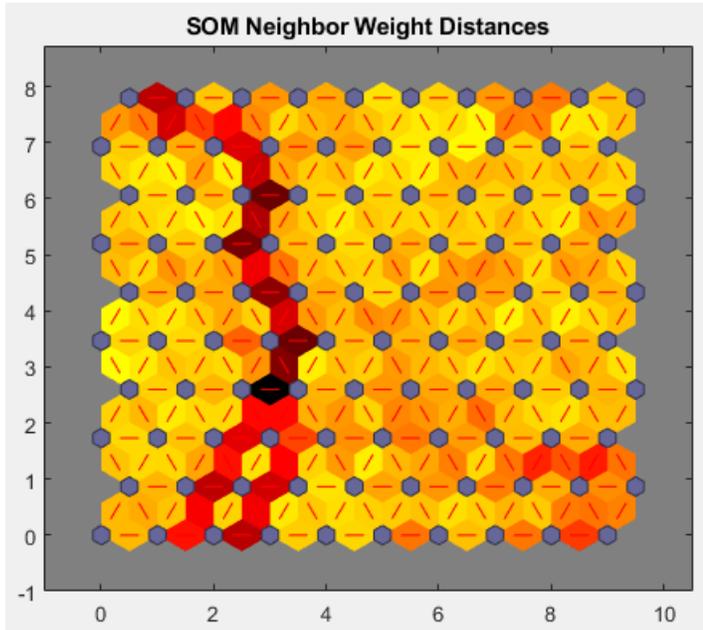
Esempio di mappa delle Hit



Mappa di uscita di una SOM e mappa delle Hit

Esempio: topologia è una griglia 10x10

Esempio di mappa in uscita alla rete SOM



In questa figura gli **esagoni grigi** rappresentano i **neuroni**.

Le **linee rosse** collegano i neuroni vicini.

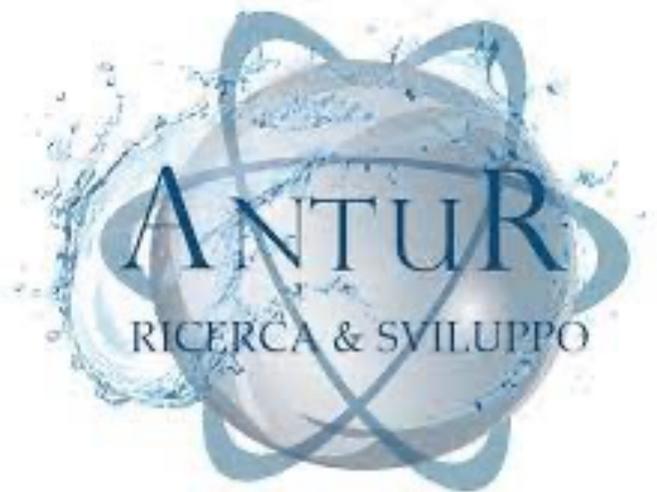
I colori nelle regioni contenenti righe rosse indicano le distanze tra i neuroni.

I colori più scuri rappresentano distanze maggiori, mentre **i colori più chiari rappresentano distanze minori**.

Nell'esempio: la mappa è attraversata da una fascia di segmenti scuri. La rete SOM sembra avere raggruppato i dati in due grandi gruppi distinti.

Se non si è soddisfatti delle prestazioni della rete, si può procedere in uno dei seguenti modi:

- **Addestrare nuovamente la rete** → Ogni addestramento ha pesi iniziali della rete diversi e può produrre una rete migliore a seguito del riaddestramento.
- **Aumentare il numero di neuroni** aumentando le dimensioni della mappa.
- **Usare un set più ampio di dati di addestramento.**



Analisi del campione Antur e applicazione di una rete SOM dedicata



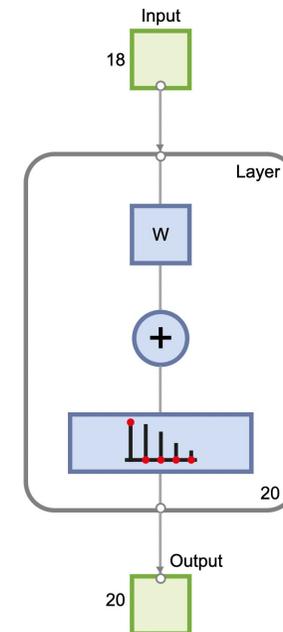
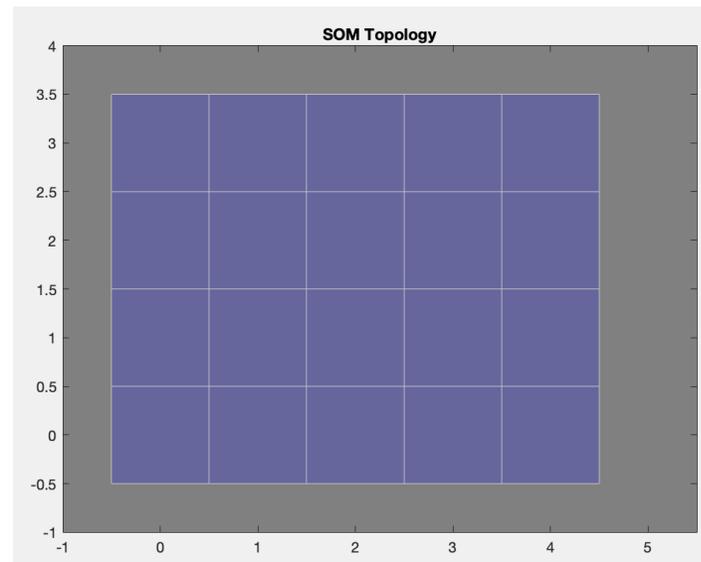
Caratteristiche generali del campione di training e rete SOM utilizzata

	Uomini	Donne	Totale
Prima visita	497	1383	1880
Totale visite	1461	4571	6032

Il campione è stato separato in un campione di training composto dal 66% dei dati e in un campione di test composto dal 34% dei dati.

Rete SOM utilizzata:

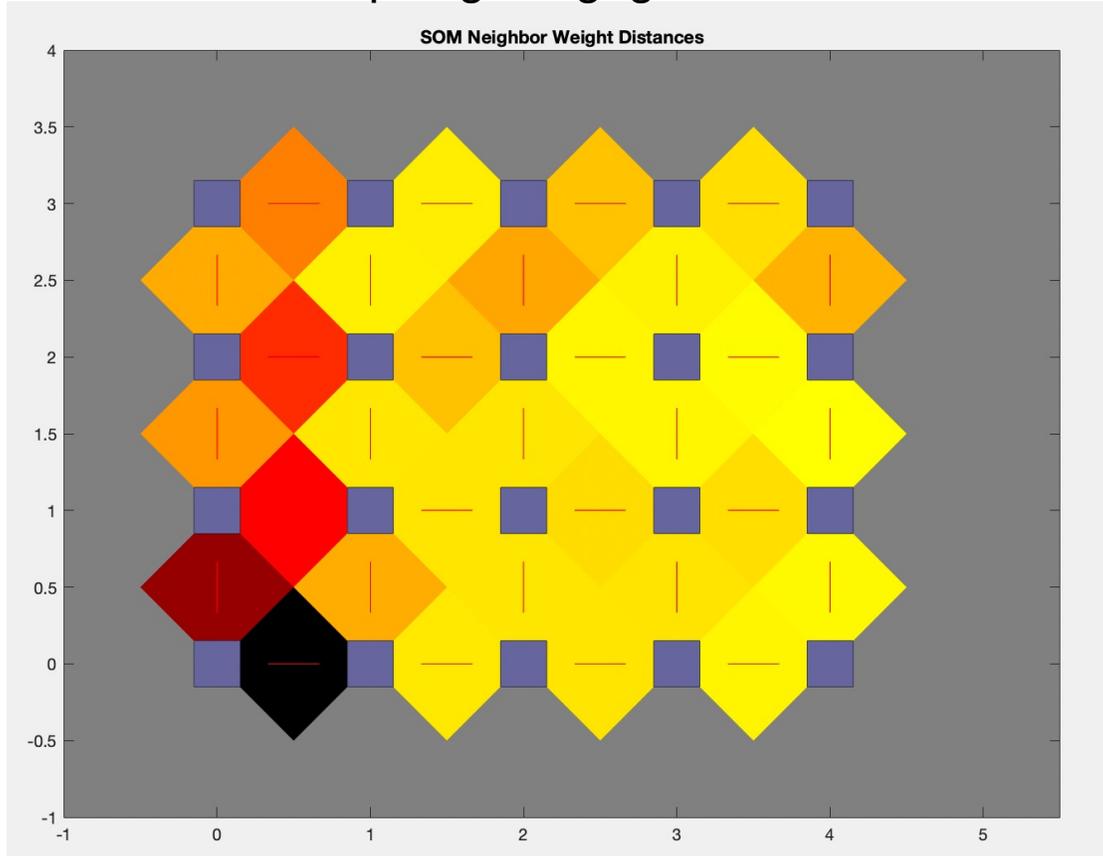
- topologia a griglia 5x4 → 20 neuroni
- funzione di distanza euclidea, modifica pesi fino ad una distanza massima di 3 nodi
- 5000 epoche di training
- 18 variabili in ingresso



Mappa di uscita di una SOM e definizione dei cluster

Risultati sul campione di training

topologia a griglia 5x4

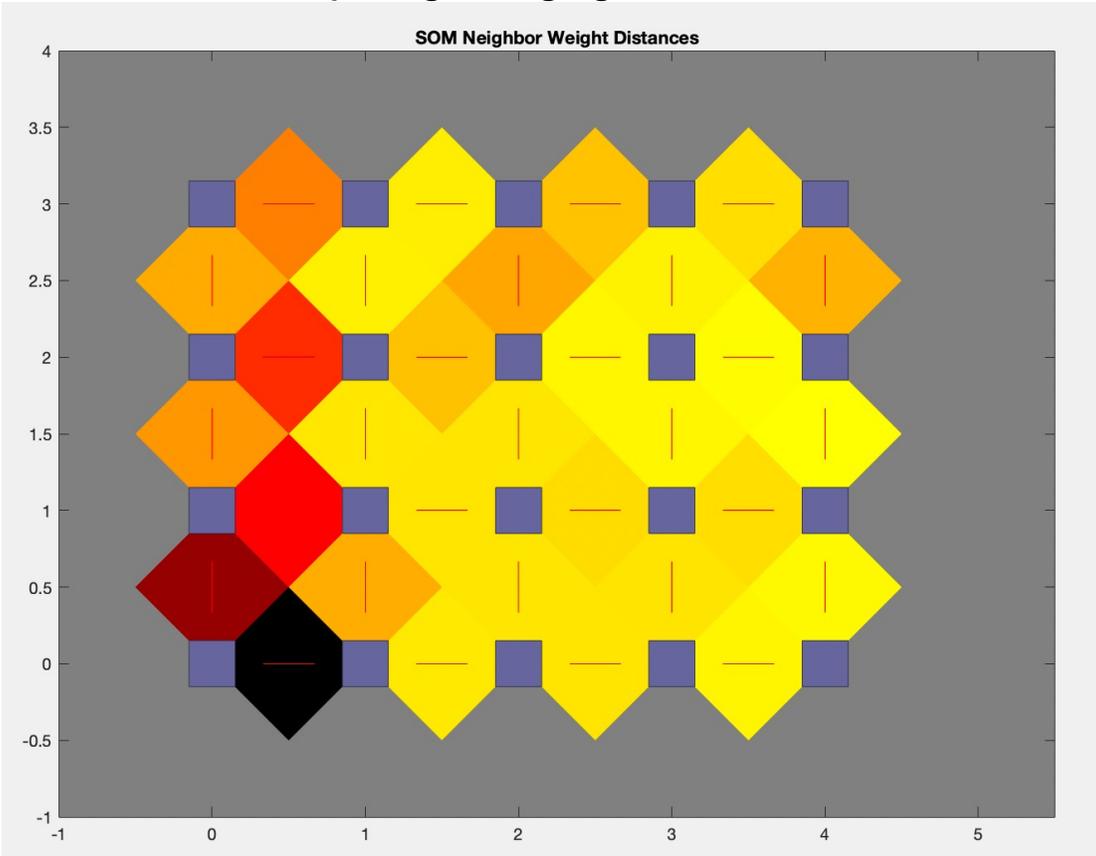


Mappa di uscita di una SOM e definizione dei cluster

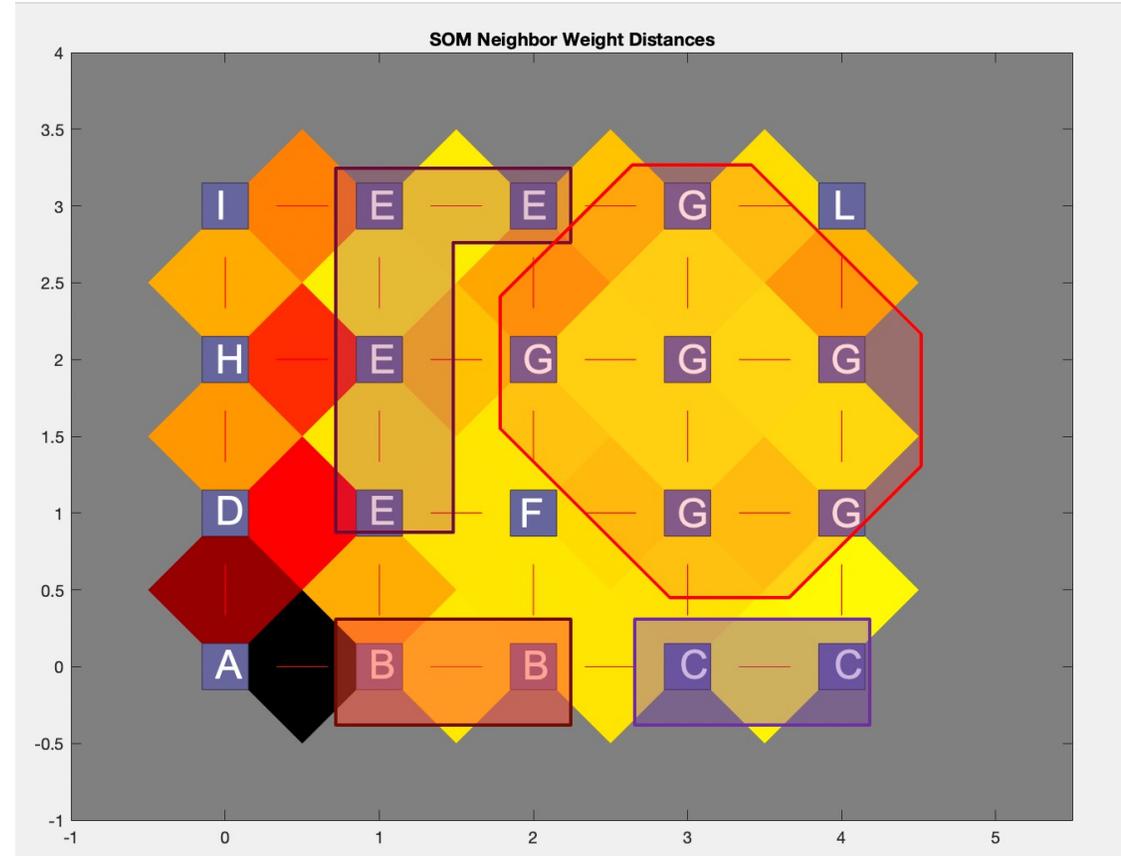
Risultati sul campione di training

6032 entries => informazioni relative a 6032 visite

Topologia a griglia 5x4



Definizione dei cluster



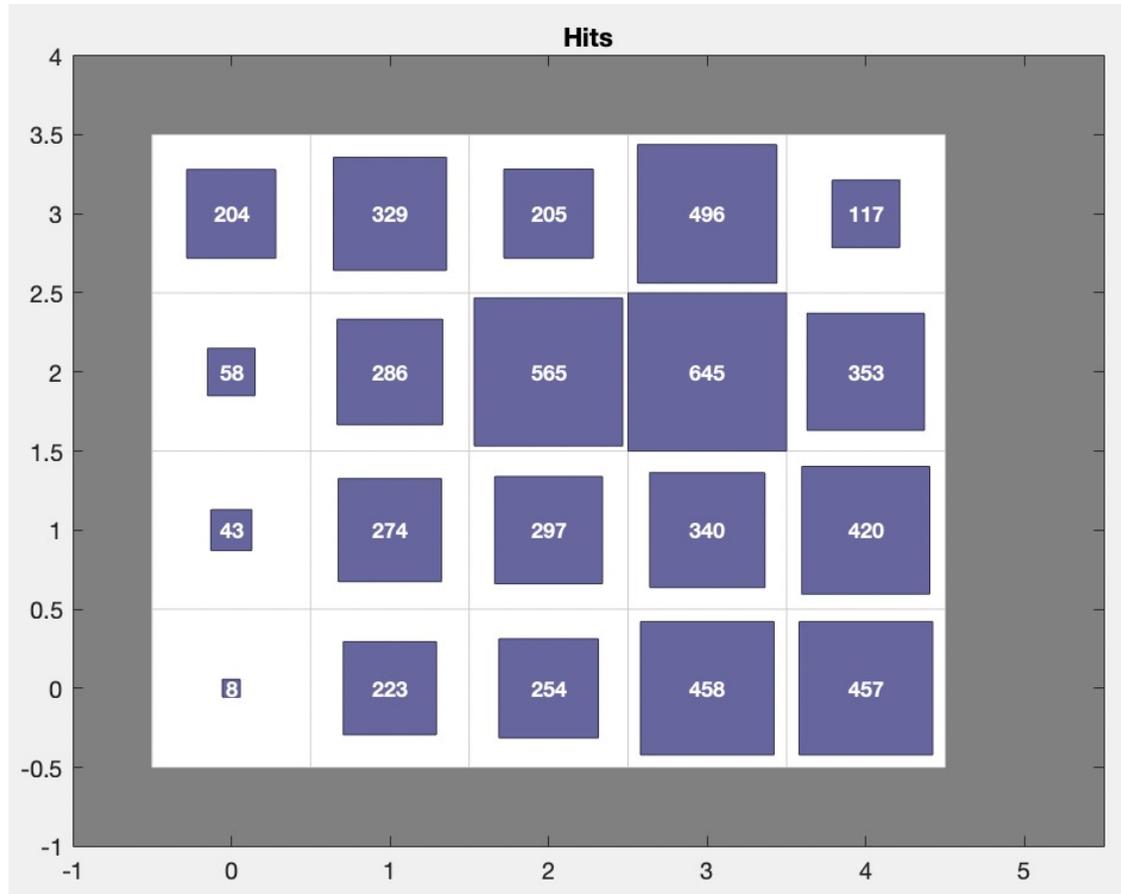
Identificati 10 cluster con caratteristiche simili

Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

Mappa di uscita delle Hit e definizione dei cluster

Risultati sul campione completo

6032 entries => informazioni relative a 6032 visite

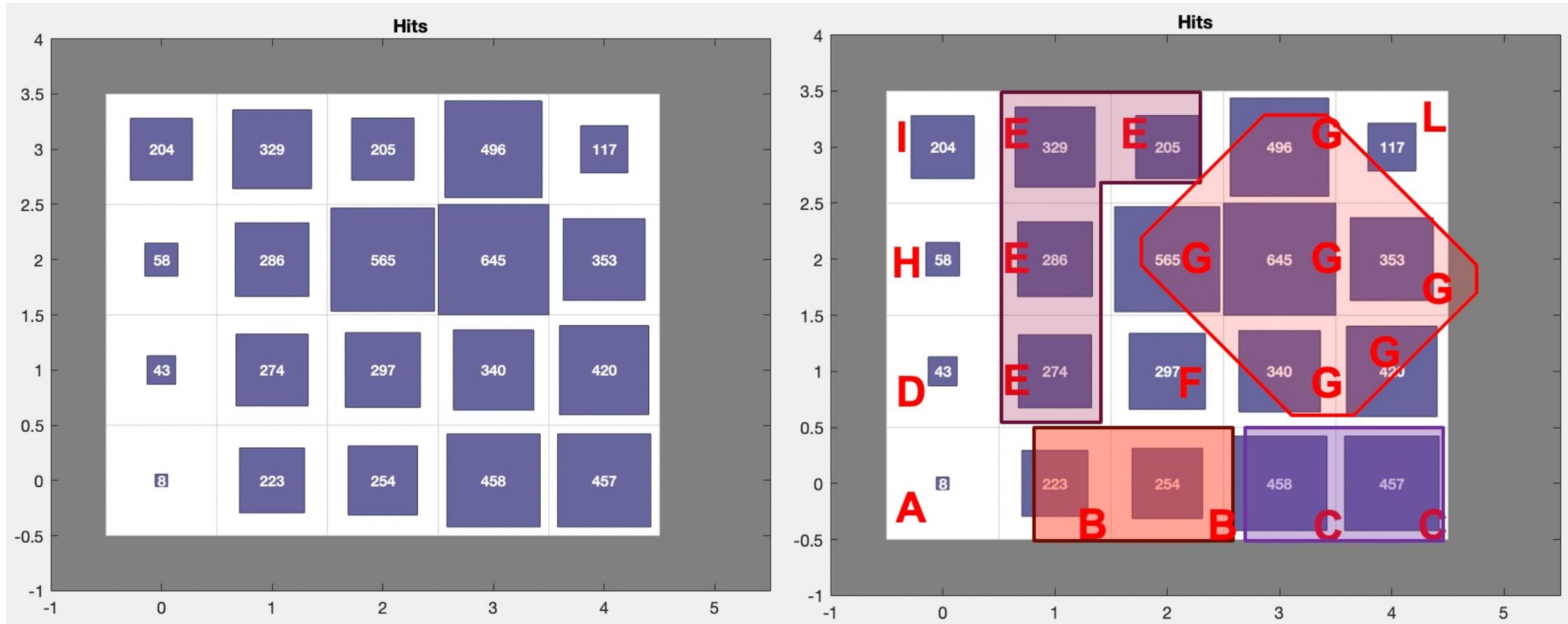


Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

Mappa di uscita delle Hit e definizione dei cluster

Risultati sul campione completo

6032 entries => informazioni relative a 6032 visite



Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

Mappa di uscita delle Hit e definizione dei cluster

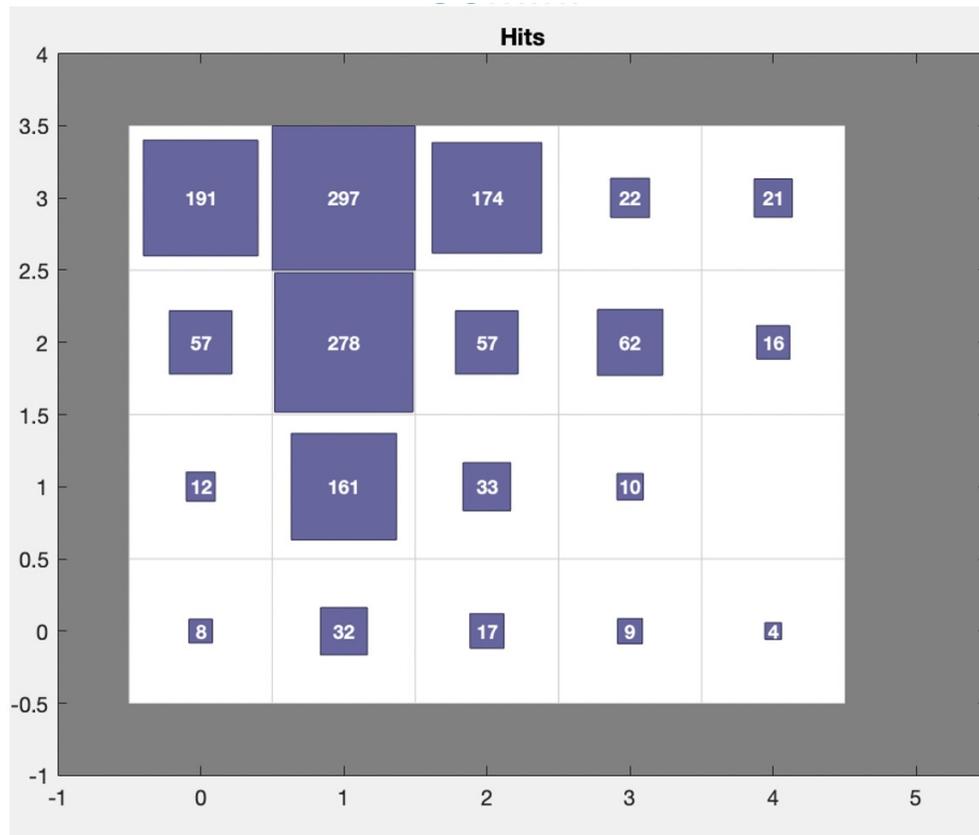
Risultati sul campione di training

6032 entries => informazioni relative a 6032 visite

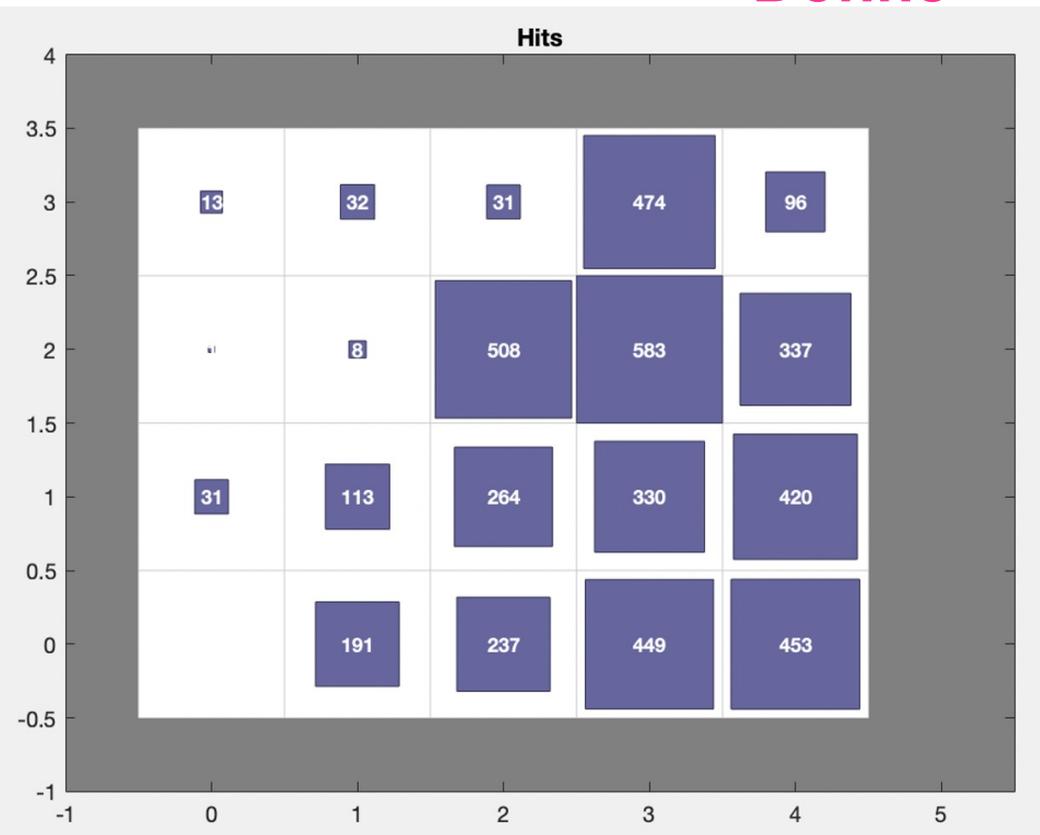
Uomini: 1461 visite

Donne: 4571 visite

Uomini



Donne



Mappa di uscita delle Hit e definizione dei cluster

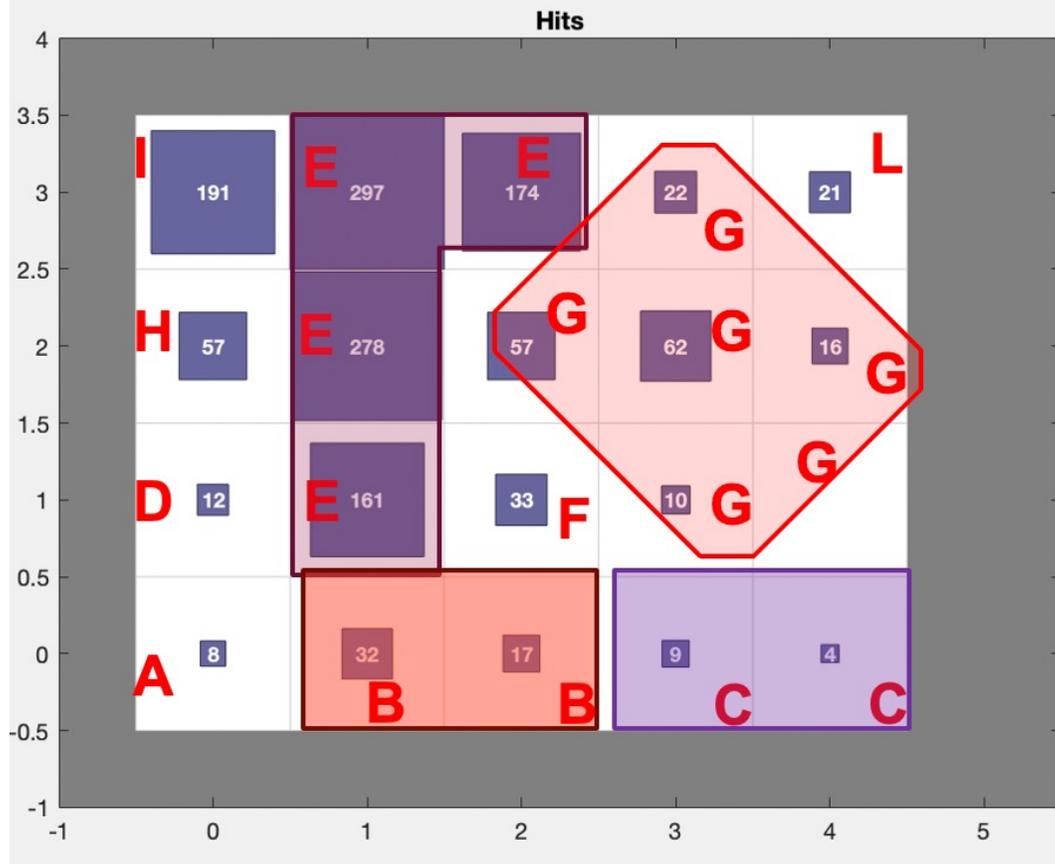
Risultati sul campione di training

6032 entries => informazioni relative a 6032 visite

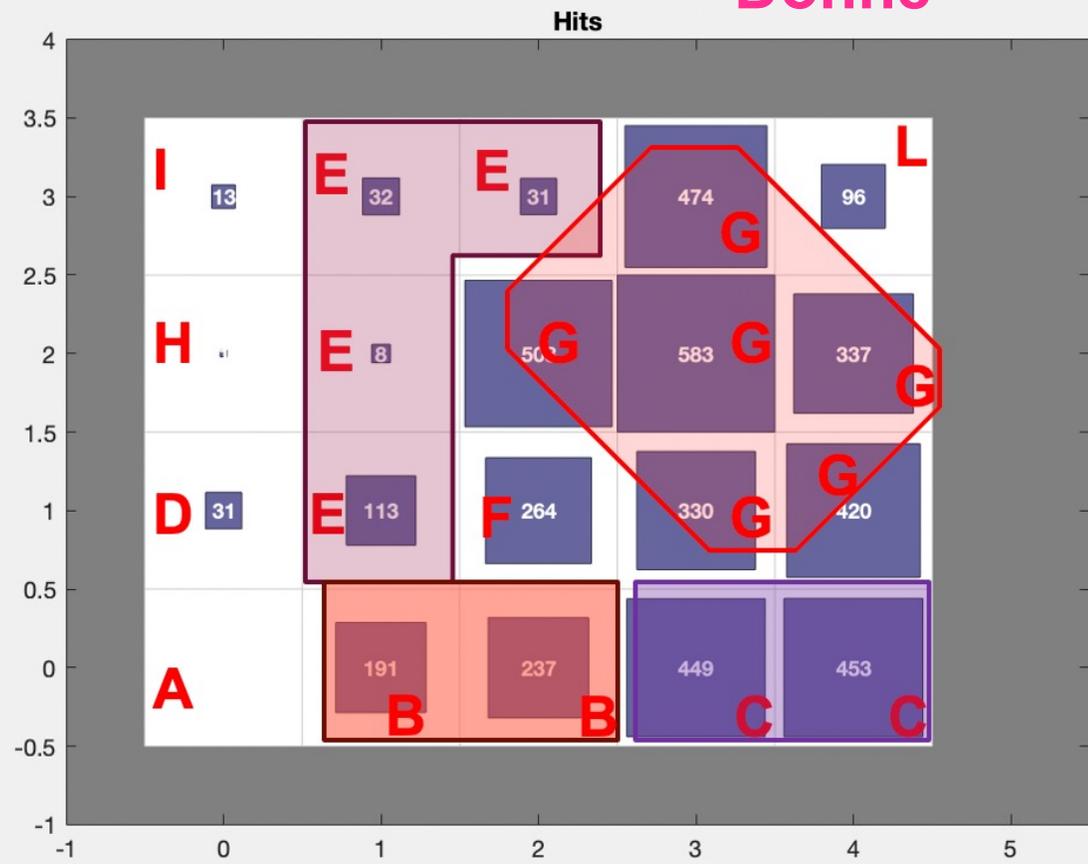
Uomini: 1461 visite

Donne: 4571 visite

Uomini



Donne



Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

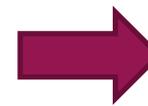
Caratteristiche generali del campione

Risultati sul campione di training

Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

Media per il cluster:	1	2	3	4	5	6	7	8	9	10
Sesso	0.00	0.90	0.99	0.73	0.18	0.90	0.94	0.03	0.06	0.69
Altezza	182.00	160.52	158.07	160.00	174.19	161.13	161.16	179.31	178.64	151.48
Età	42.20	44.59	54.57	43.02	38.47	27.48	34.69	38.87	36.49	25.01
Peso	186.26	101.72	76.80	126.21	86.57	87.24	67.40	133.05	110.29	43.82
BMI	56.32	39.51	30.75	50.03	28.59	33.63	25.94	41.37	34.60	19.03

- Cluster 1 costituito da soggetti affetti da **obesità grave**
- Cluster 8 costituito da soggetti **molto sovrappeso**
- Cluster 9 costituito da soggetti **sovrappeso**
- Cluster 5 costituito da soggetti **leggermente sovrappeso**
- Tutti i suddetti cluster sono composti da soggetti mediamente giovani



Gruppi prevalentemente popolati da uomini

Maggiori dettagli sul campione e sui risultati nella presentazione di Agostino

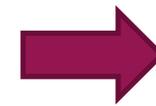
Caratteristiche generali del campione

Risultati sul campione di training

Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

Media per il cluster:	1	2	3	4	5	6	7	8	9	10
Sesso	0.00	0.90	0.99	0.73	0.18	0.90	0.94	0.03	0.06	0.69
Altezza	182.00	160.52	158.07	160.00	174.19	161.13	161.16	179.31	178.64	151.48
Età	42.20	44.59	54.57	43.02	38.47	27.48	34.69	38.87	36.49	25.01
Peso	186.26	101.72	76.80	126.21	86.57	87.24	67.40	133.05	110.29	43.82
BMI	56.32	39.51	30.75	50.03	28.59	33.63	25.94	41.37	34.60	19.03

- Cluster 4 costituito da soggetti affetti da **obesità grave di età media di 47 anni**
- Cluster 2 costituito da soggetti **sovrappeso e meno giovani**
- Cluster 6 costituito da soggetti **sovrappeso e giovani**
- Cluster 7 costituito da soggetti **mediamente normopeso e giovani**
- Cluster 3 costituito da soggetti **mediamente normopeso e meno giovani**
- Cluster 10 costituito da soggetti **sottopeso e molto sottopeso e molto giovani**

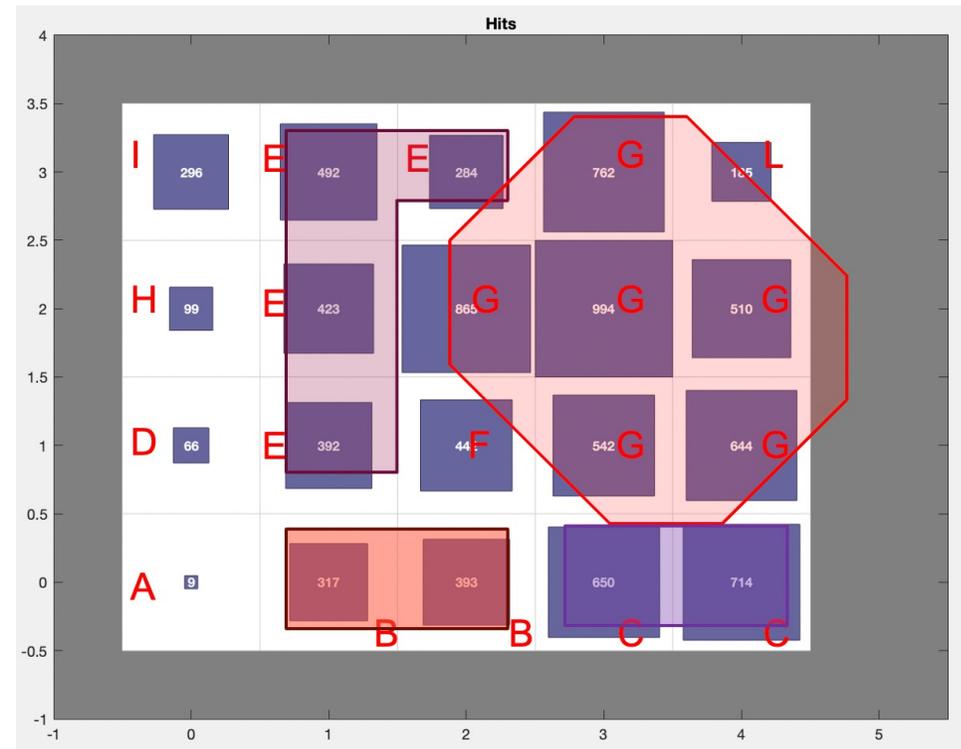
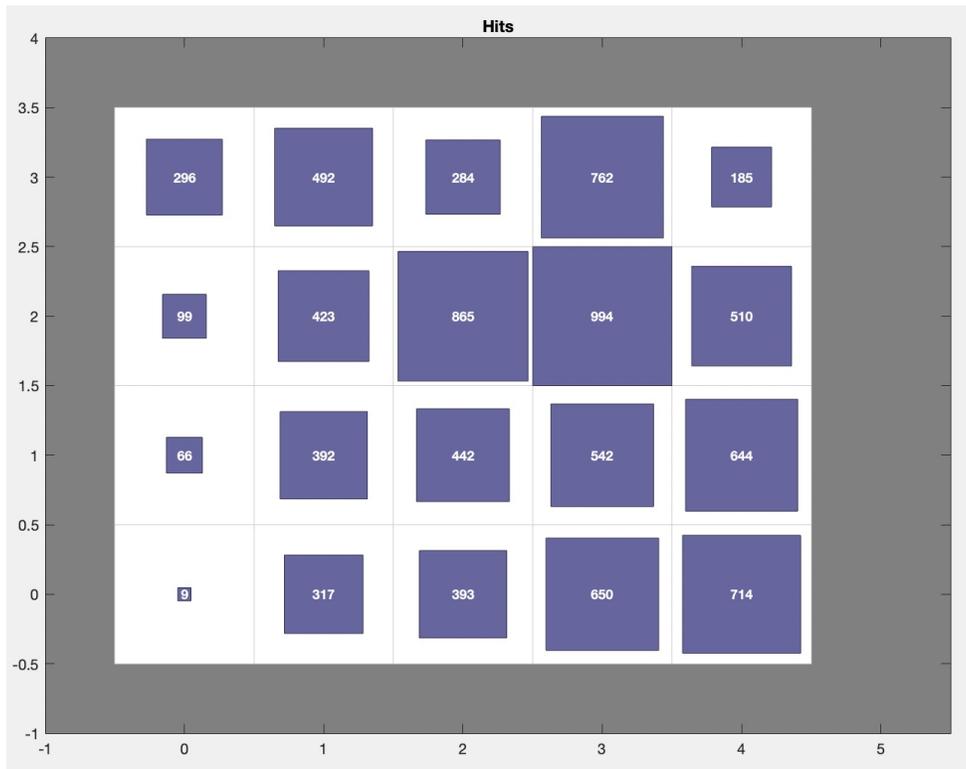


Gruppi prevalentemente popolati da donne

Maggiori dettagli sul campione e sui risultati nella presentazione di Agostino

Caratteristiche generali del campione completo e rete SOM utilizzata

	Uomini	Donne	Totale
Prima visita	651	1765	2416
Totale visite	1461	6993	9080



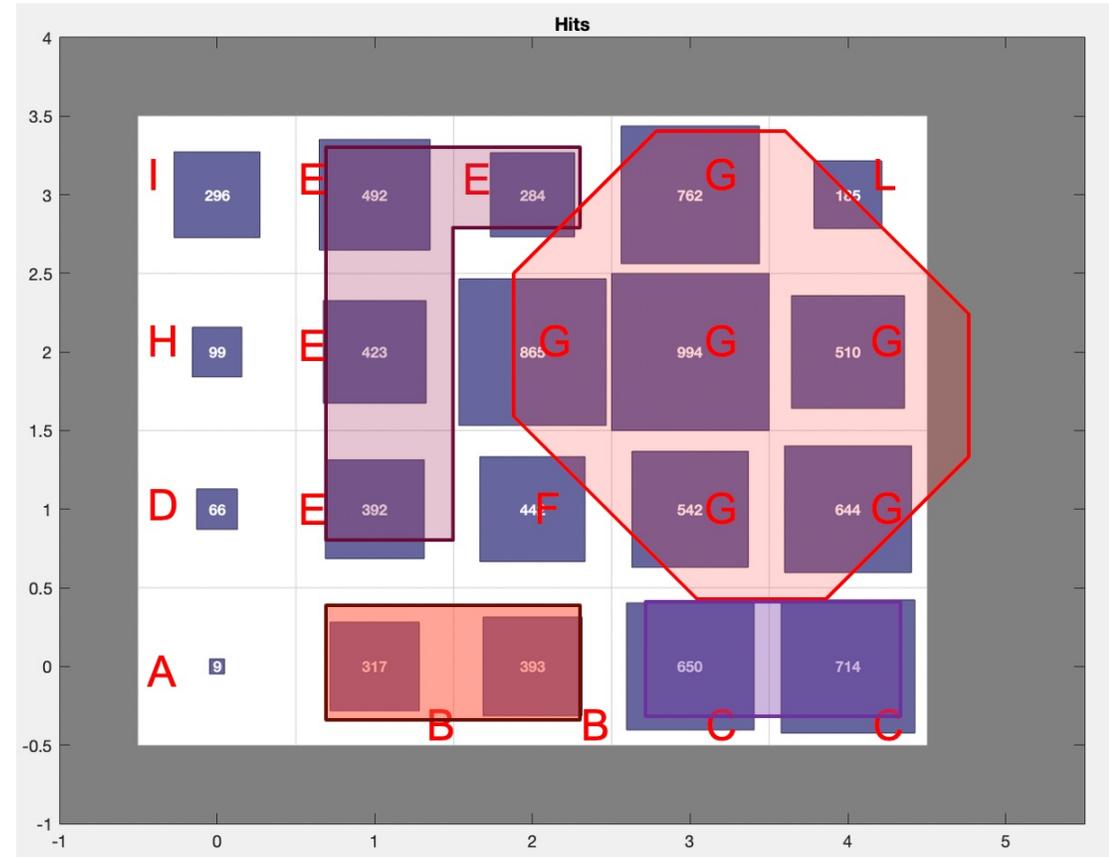
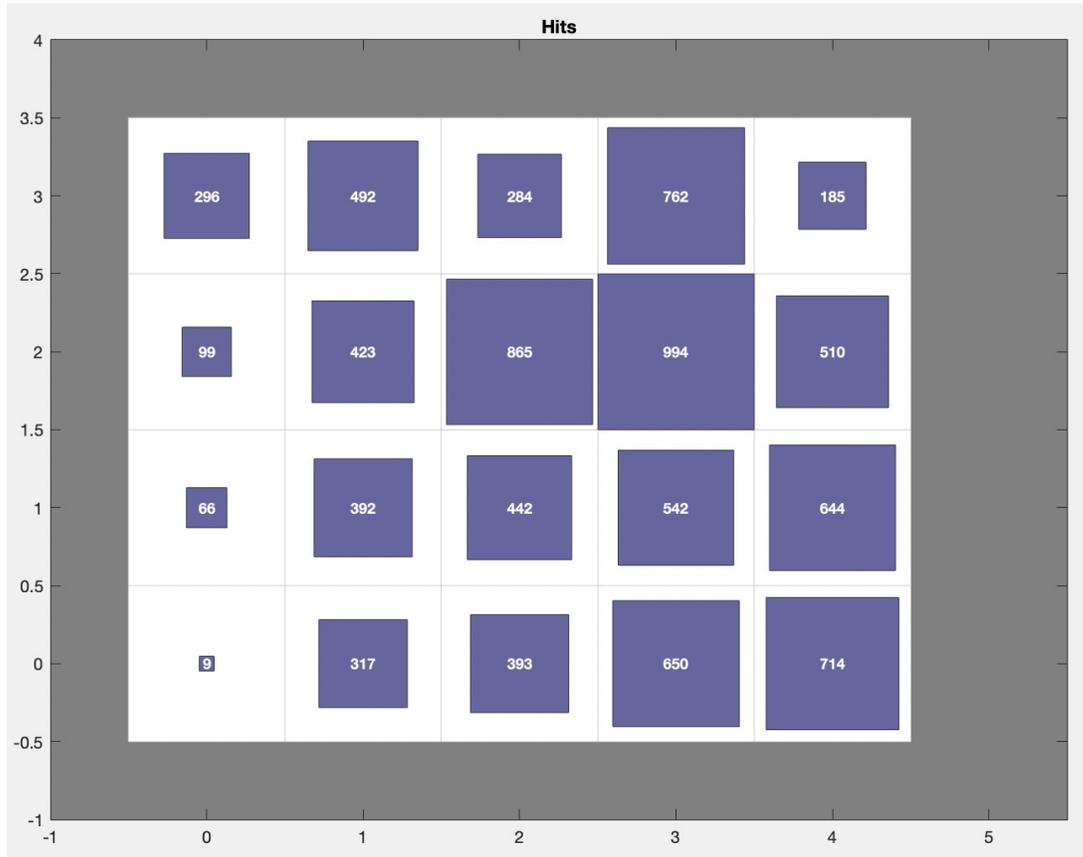
Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

L'applicazione della rete sul campione di test conferma e valida la caratterizzazione dei cluster

Mappa di uscita delle Hit e definizione dei cluster

Risultati sul campione completo

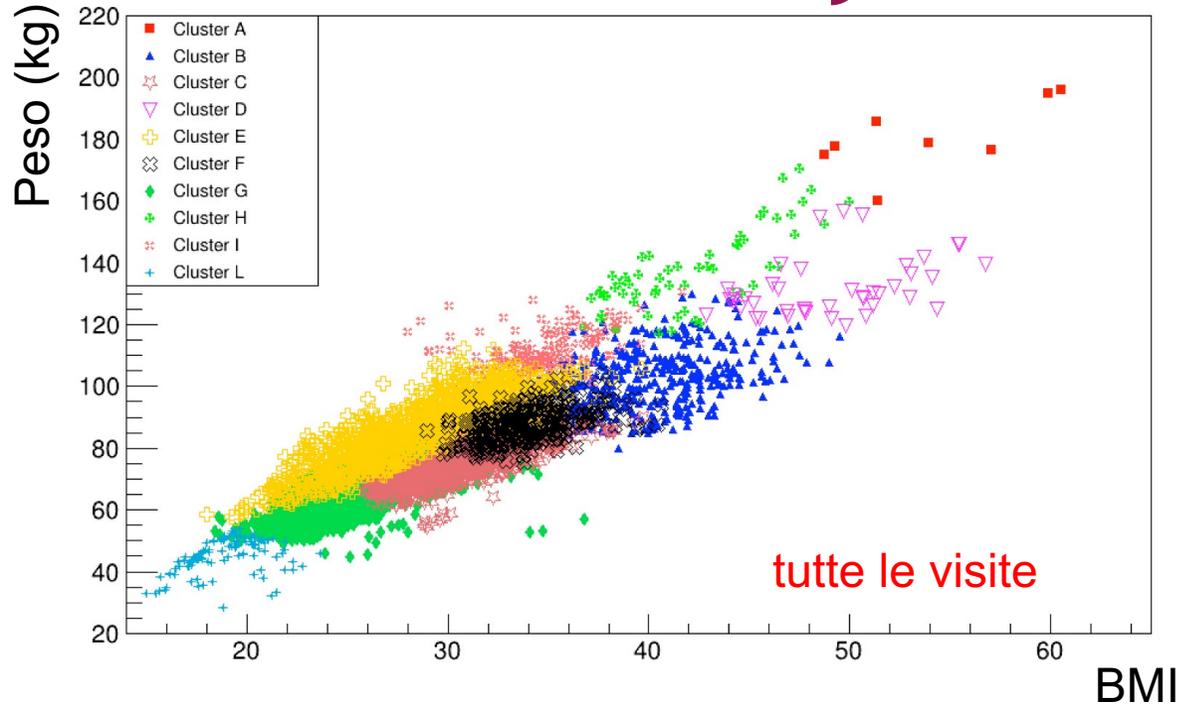
9080 entries => informazioni relative a 9080 visite



Cluster: A=1; B=2; C=3; D=4; E=5; F=6; G=7; H=8; I=9; L=10

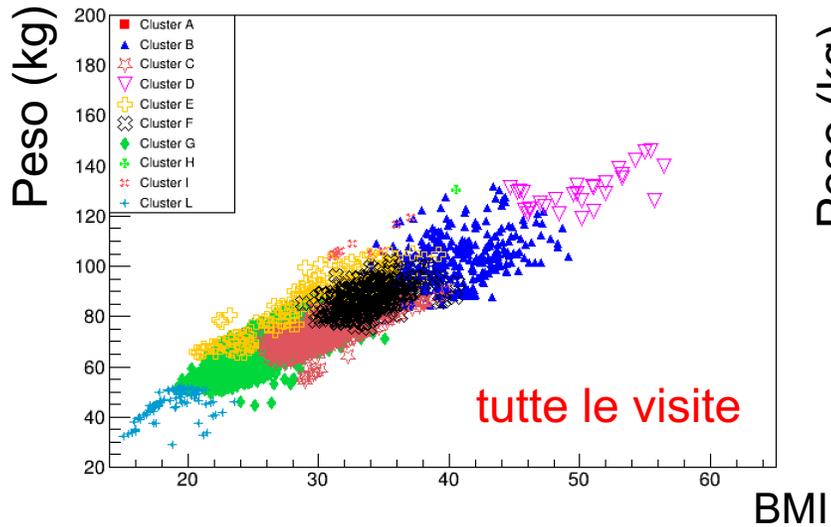
L'applicazione della rete sul campione di test conferma e valida la caratterizzazione dei cluster

Peso vs. Body Mass Index (BMI)

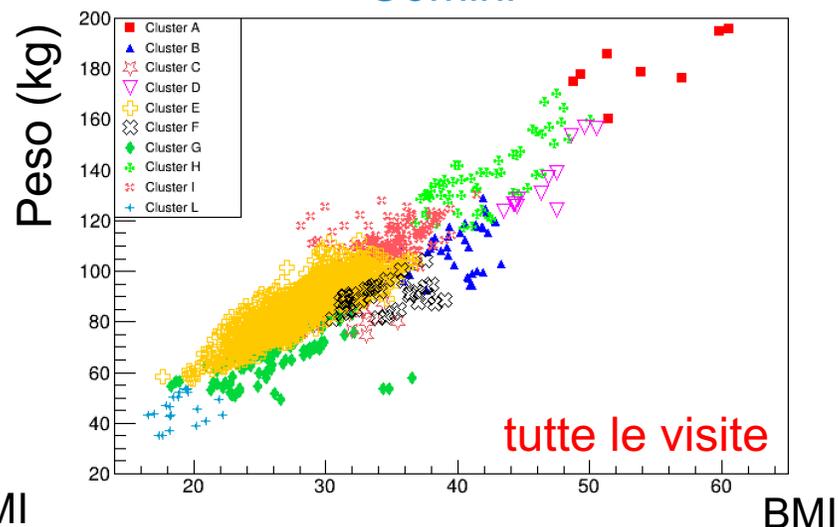


BMI = Peso (Kg)/[Altezza (m)]²	
Valori di riferimento	
Sottopeso	≤ 18.5
Normopeso	18.6 - 24.9
Sovrappeso	25.0 – 29.9
Obesità 1° grado	30.0 – 34.9
Obesità 2° grado	35.0 – 39.9
Obesità morbigena	≥ 40

Donne

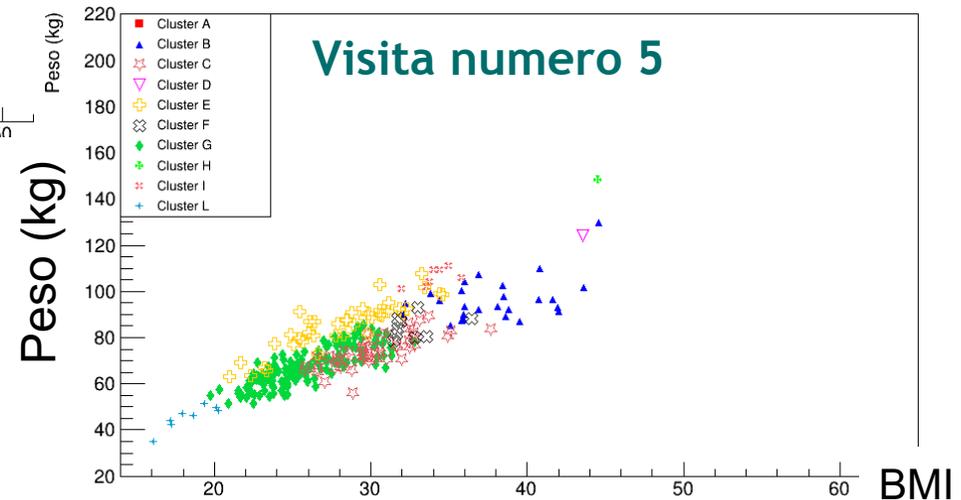
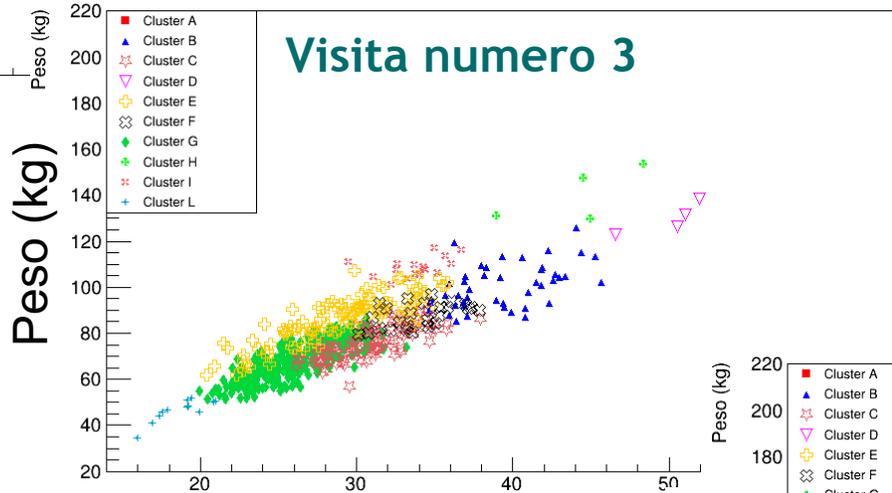
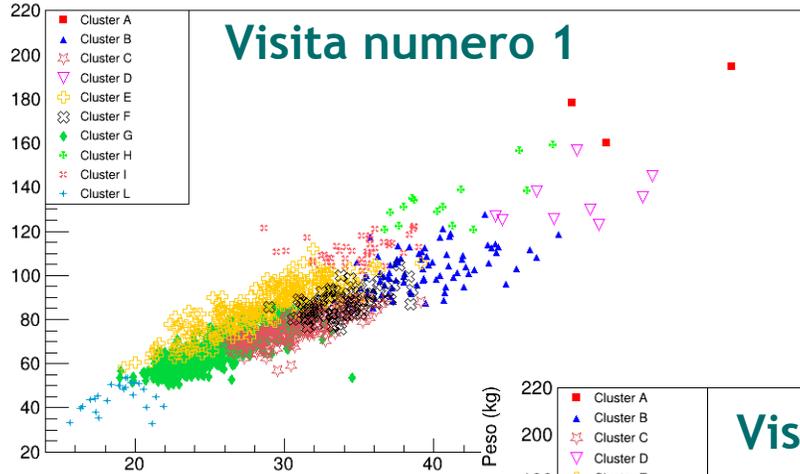


Uomini



Peso vs. Body Mass Index (BMI)

Peso (kg)



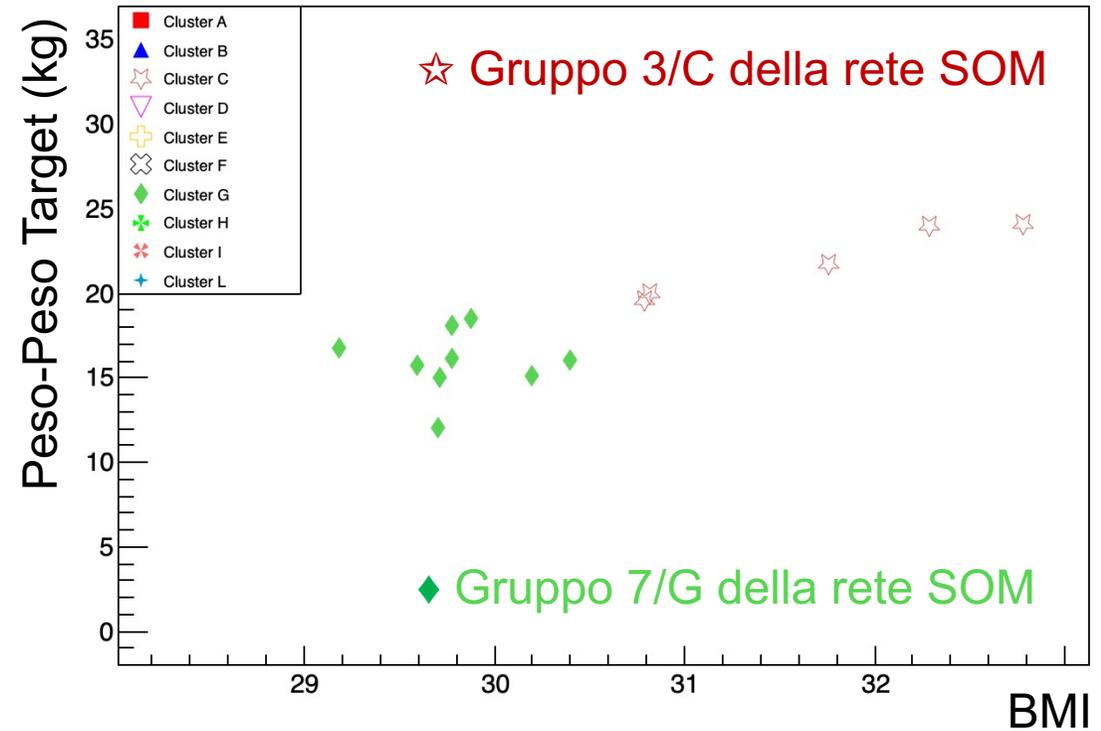
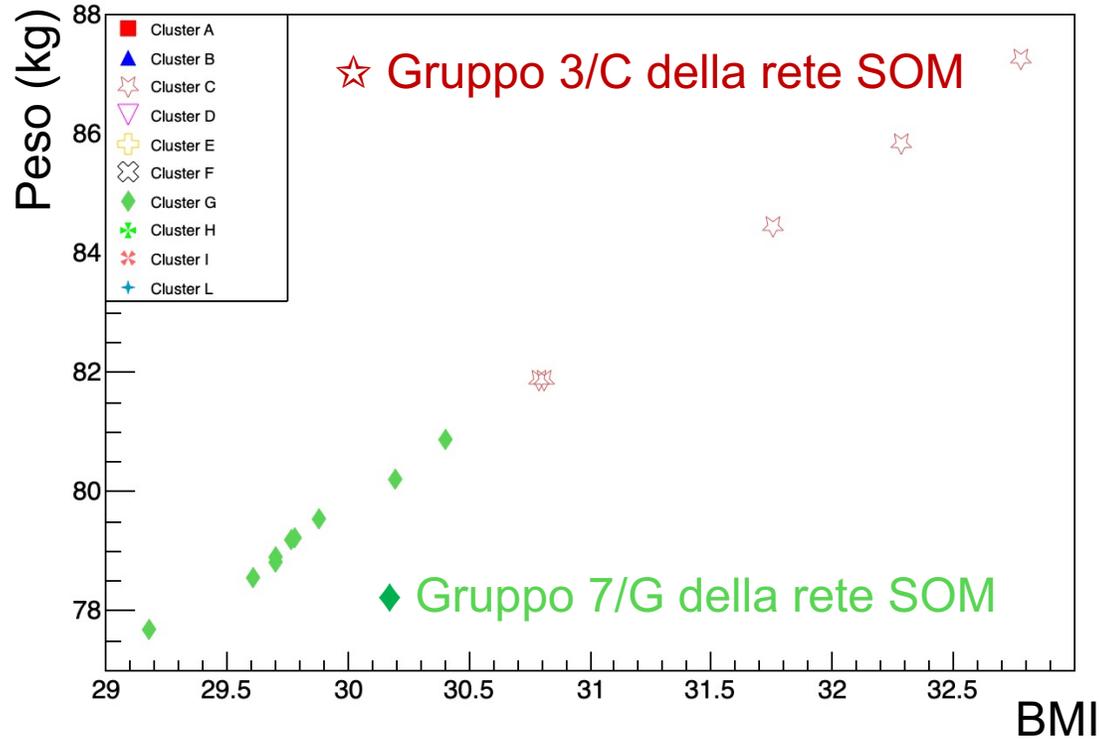
$$\text{BMI} = \text{Peso (Kg)} / [\text{Altezza (m)}]^2$$

Valori di riferimento

Sottopeso	≤ 18.5
Normopeso	18.6 - 24.9
Sovrappeso	25.0 – 29.9
Obesità 1° grado	30.0 – 34.9
Obesità 2° grado	35.0 – 39.9
Obesità morbigena	≥ 40

Height=163cm
Gender=1 (Donna)

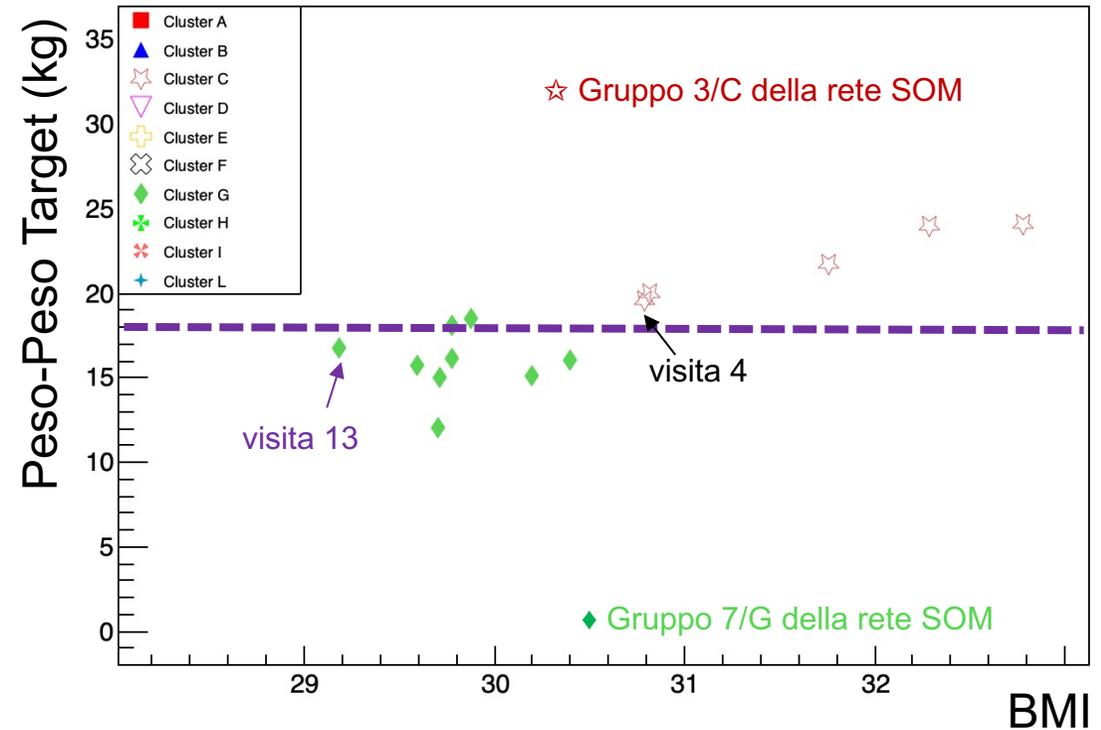
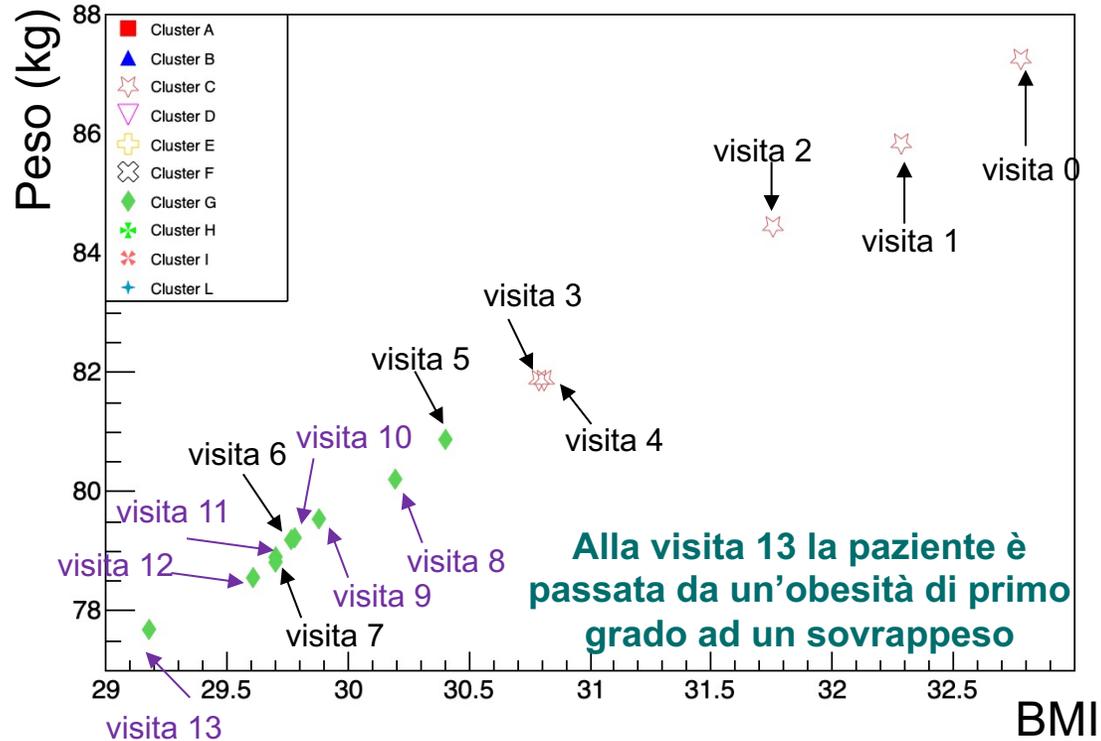
"Local_ID==610"



BMI = Peso (Kg)/[Altezza (m)] ²	
Valori di riferimento	
Sottopeso	≤ 18.5
Normopeso	18.6 - 24.9
Sovrappeso	25.0 – 29.9
Obesità 1° grado	30.0 – 34.9
Obesità 2° grado	35.0 – 39.9
Obesità morbigena	≥ 40

Height=163cm
Gender=1 (Donna)

"Local_ID==610"



BMI = Peso (Kg)/[Altezza (m)] ²	
Valori di riferimento	
Sottopeso	≤ 18.5
Normopeso	18.6 - 24.9
Sovrappeso	25.0 – 29.9
Obesità 1° grado	30.0 – 34.9
Obesità 2° grado	35.0 – 39.9
Obesità morbigena	≥ 40

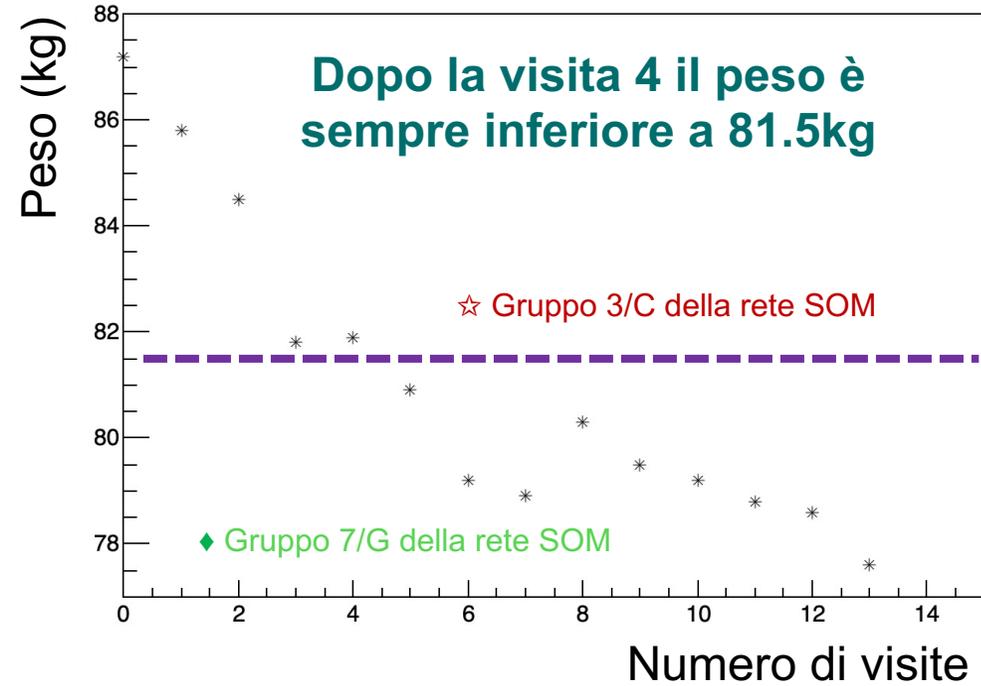
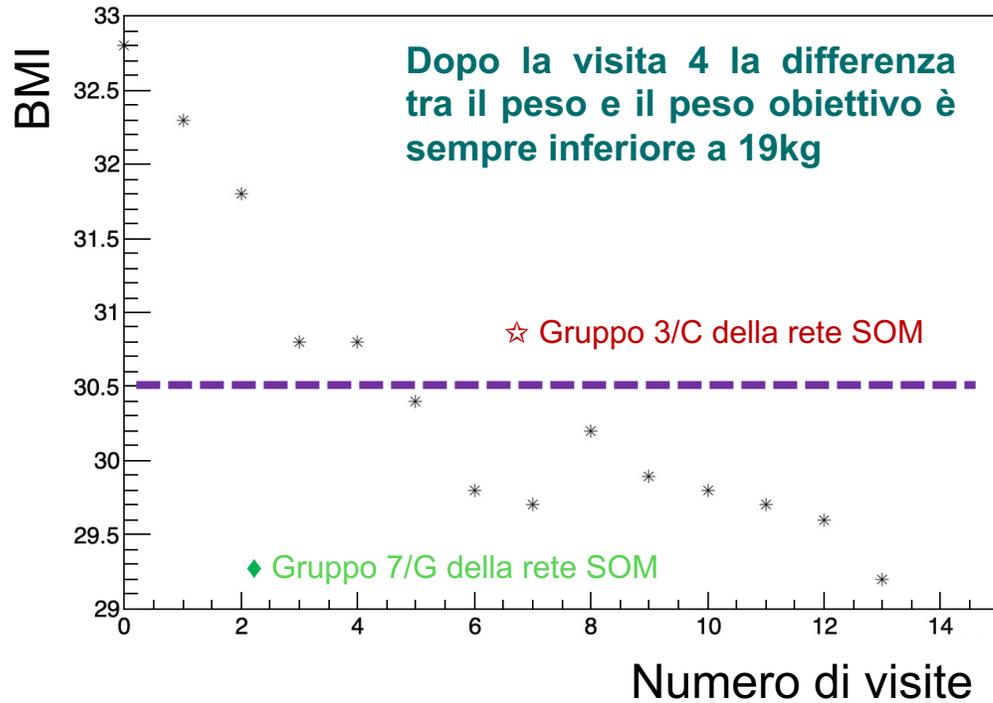
Dopo la visita 4 la differenza tra il peso e il peso obiettivo è sempre inferiore a 19kg

Alla visita 13 la differenza tra il peso e il peso obiettivo è di 16.7kg

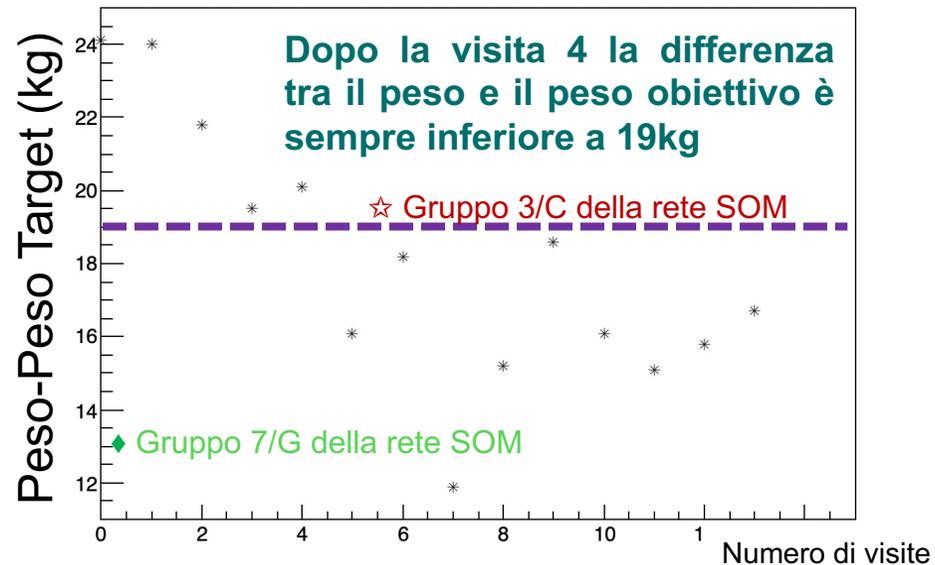
Il peso obiettivo è stato adattato alla motivazione della paziente e ai risultati ottenuti

Height=163cm
Gender=1 (Donna)

"Local_ID==610"



BMI = Peso (Kg)/[Altezza (m)]²	
Valori di riferimento	
Sottopeso	≤ 18.5
Normopeso	18.6 - 24.9
Sovrappeso	25.0 – 29.9
Obesità 1° grado	30.0 – 34.9
Obesità 2° grado	35.0 – 39.9
Obesità morbigena	≥ 40

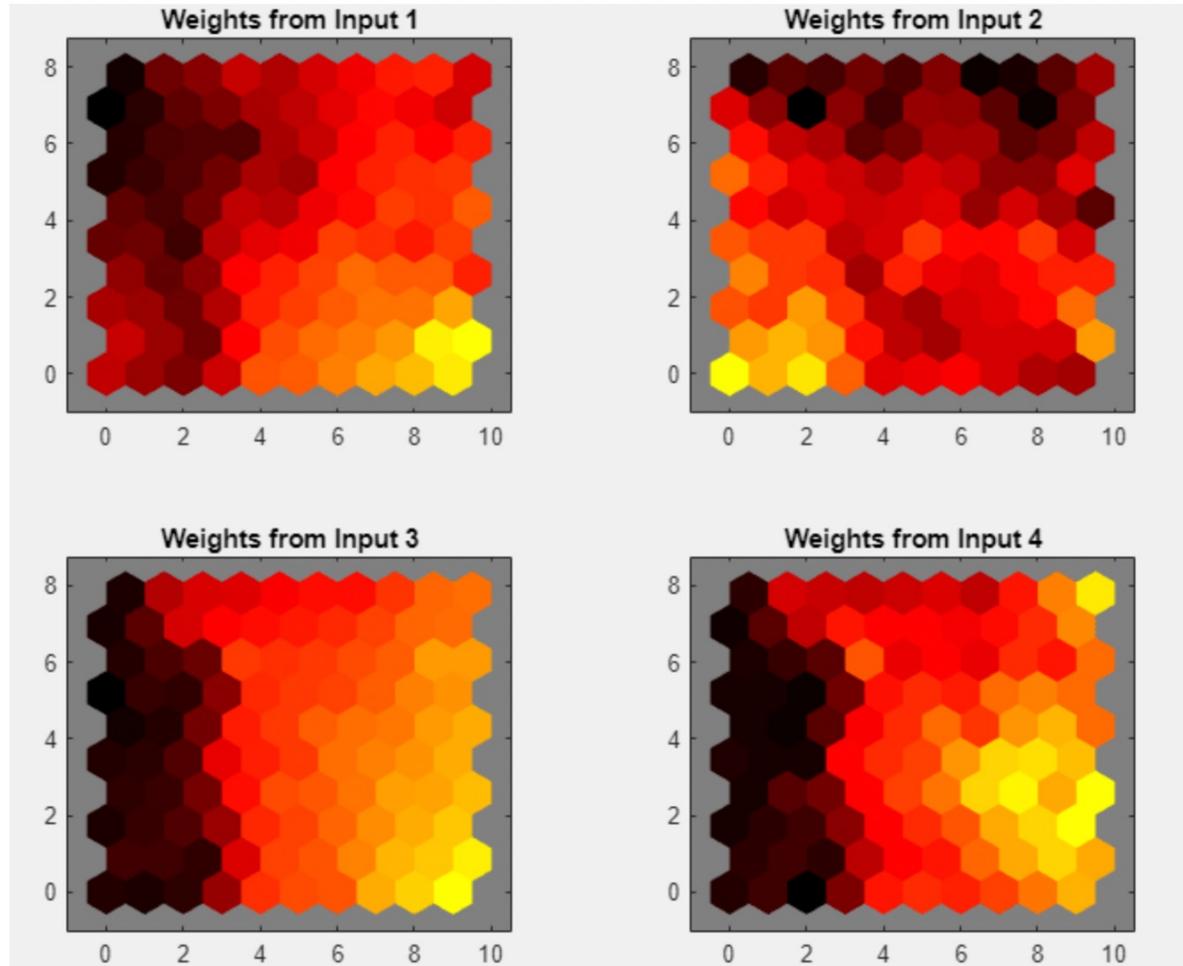


BMI

backup

Grafici dei pesi

Il grafico mostra i pesi che collegano ogni input a ciascuno dei neuroni



4 variabili o feature di input

Questa figura mostra un piano del peso per ciascun elemento delle feature di input (quattro nell'esempio).

I colori più scuri rappresentano i pesi più grandi.

Se i modelli di connessione di due feature sono molto simili, si può presupporre che le feature siano altamente correlate.