

## Machine Learning

- Razionale: usare machine learning come mezzo per affiancare, all'output dell'inversione (immagine), un "allarme".
- Ipotesi nulla ( $H_0$ )  $\rightarrow$  cervello sano
- Ipotesi alternativa ( $H_1$ )  $\rightarrow$  presenza di **emorragia**. (E gli ictus ischemici? si può provare, ma serve un modo di modellarli, come abbiamo modellato gli emorragici con un'inclusione di sangue)
- In letteratura si fa, ma, (**as far as I know!** non ho più controllato la letteratura da quando abbiamo chiuso la relativa parte del deliverable ) si utilizzano tecniche statistiche:
  - Raccolgo dati (segnale nel nostro caso) di interesse
  - Trovo un insieme di statistiche significative, nel senso che possano caratterizzare e discriminare le due ipotesi.
  - Uso una quantità opportuna di training per trovare, su queste statistiche, delle soglie/criteri di decisione.
  - Per classificare il campione sconosciuto, ne ricavo analogamente le suddette statistiche e le confronto con le soglie/criteri trovate con il training
- Nostra proposta: usare una rete neurale. Di fatto è un approccio simile.
  - Raccolgo dati (segnale nel nostro caso) di interesse
  - Se possibile ne ricavo *feature* "furbe" (es. simmetria)
  - Uso una quantità opportuna di training per addestrare la rete (qui ne serve molto, possibilmente i metodi statistici ne richiedono meno..)
  - Campione sconosciuto: lo infilo nella rete neurale che in uscita ha due neuroni, uno per  $H_0$  e uno per  $H_1$ : ciascuno sarà attivato in certa misura (uno scalare).
  - Per classificarlo confronto le attivazioni dei neuroni in uscita: se non ho priori, come al solito, devo considerare quello che è maggiormente attivato
- Riassumendo: la rete neurale di fatto cerca di trovare il "modo migliore" di combinare i dati in ingresso per far sì che i neuroni output si accendano correttamente.
- Il vantaggio è che è forse più facile come approccio (non devo pensare a complesse tecniche statistiche, trasformate, algebra lineare, sottospazi eccetera), e che soprattutto trova da sola la via "migliore", sempre che i dati da darle in pasto siano ben posti, mentre con tecniche statistiche forse si perde un po' di questa flessibilità.

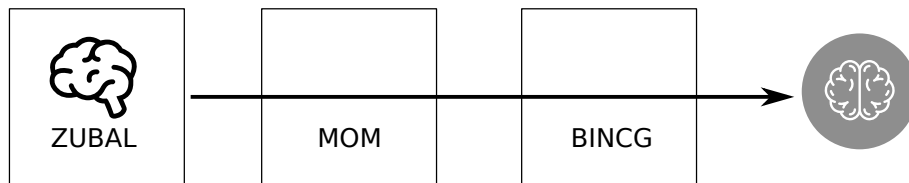


Figura 1:

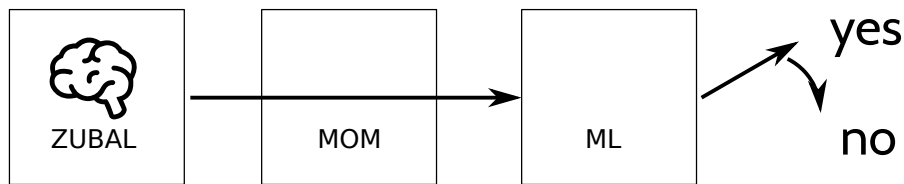


Figura 2:

## La Baracca<sup>TM</sup>

Ho usato parte del simulatore di Randazzo/Fedeli. In origine è come in fig 1, io lo uso come in fig 2.

- MOM: a partire da
  - il modello di cervello 3D
  - che fetta di questo modello considero
  - quante antenne
  - dove sono posizionate
  - a che frequenza emettono
  - quali sono i valori e.m. dei tessuti
  - etc

mi restituisce **le misure** elettromagnetiche (**2D**, si considera una sola “fetta” del modello 3D)

- Lavora a singola frequenza (1 GHz)
- **Misure**, cosa sono: quanto forte e con quale fase, ogni antenna riceve la sinusoide a 1GHz.
- **Misure**, quante sono: Se ho  $N$  antenne, ogni volta che una trasmette,  $N - 1$  ricevono. Totale  $N(N - 1)$  antenne,  $\rightarrow N(N - 1)$  numeri complessi (intensità e fase).

## Parametri usati e come ho prodotto il Dataset

A rigore occorrerebbero molti esempi di cervello diversi. Problema: ne abbiamo uno (Zubal). Come ho “risolto” il problema: inserisco “rumore” sulle

caratteristiche dei tessuti. Esempio: se il sangue ha  $\epsilon_b$  e  $\sigma_b$ , considero tante simulazioni in cui ciascuna ha  $\epsilon = \epsilon_b * (1 - n)$  e  $\sigma = \sigma_b * (1 - n)$  con  $n$  v.a. Considero anche diverse fette di cervello, così da avere configurazioni dei tessuti leggermente diverse.

## Rete neurale

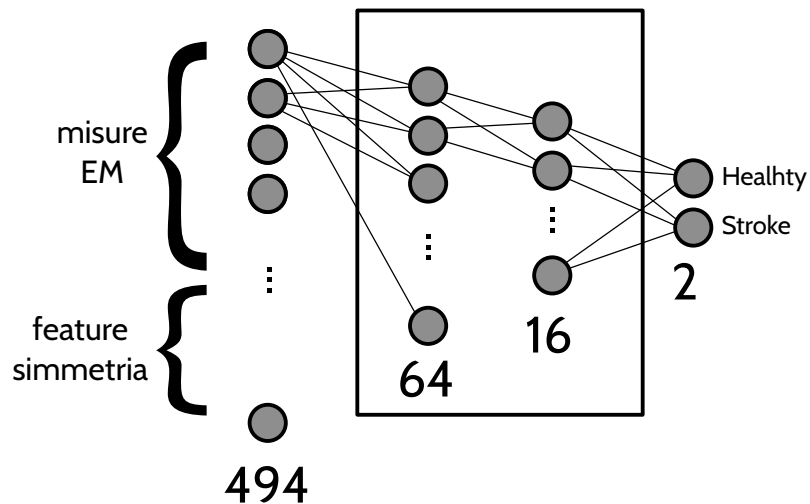


Figura 3:

In pasto gli do le misure elettromagnetiche: 16 antenne, quindi  $16 \cdot 15 \cdot 2$  misure, più 14, misure di simmetria su modulo e fase (vedi fig. 4)

Semplicissimo multilayer perceptron. Due layer, uno da 64 e uno da 16. Numeri euristicissimi.

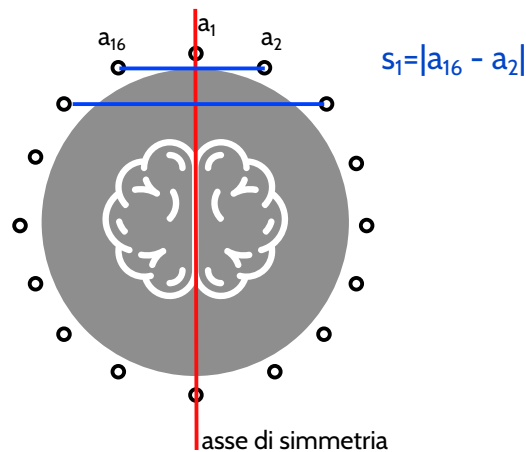


Figura 4:

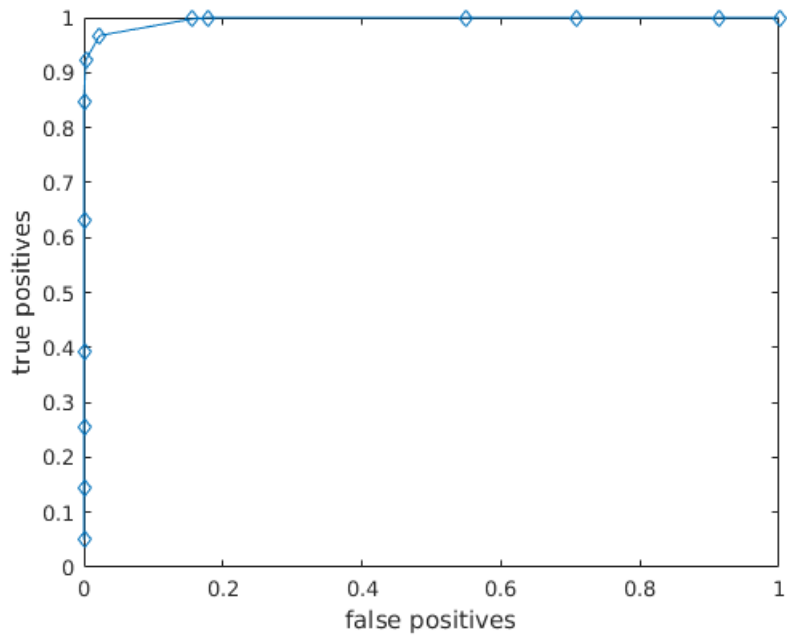


Figura 5:

## Risultati

### Dataset 4

Tot: 2000 campioni, 1600 di training e 400 di test.

- 16 antenne
- rumore post MOM (SNR = v.a. intera uniforme tra -10 e 30 dB)
- variabilità su posizione (v.a. uniforme in cerchio di raggio 5cm) e dimensione (ellisse, assi v.a. uniformi in  $[1,4]$  cm) stroke
- rumore su tissueprop (ogni valore scalato di percentuale gaussiana,  $\sim 95\%$  contenuta in  $\pm 10\%$ )
- variabilità su raggio antenne (uniforme 11-12cm)

Risultati:

- AUC = 0.9937
- accuracy: 97%

### Dataset 5

Tot: 1000 campioni, 700 di training e 300 di test.

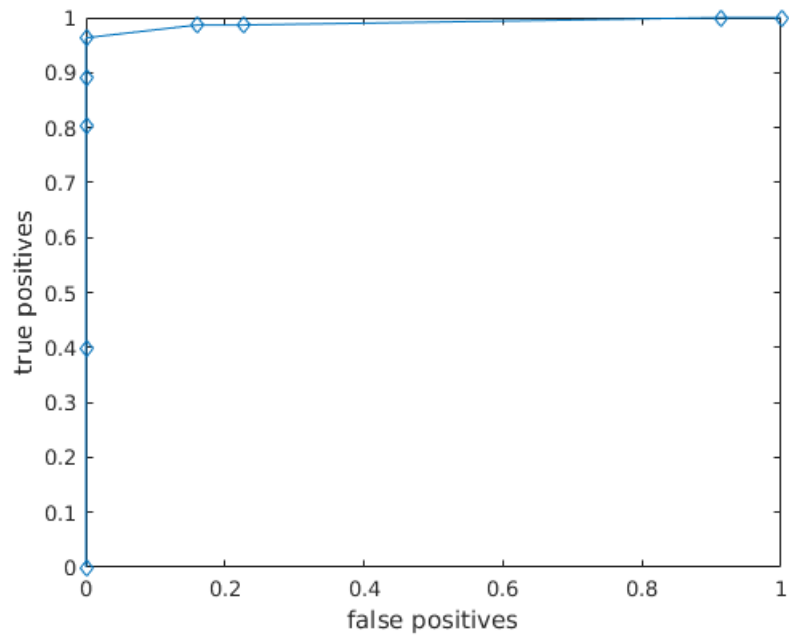


Figura 6:

- 16 antenne
- rumore post MOM (SNR = v.a. intera uniforme tra -10 e 30 dB)
- variabilità su posizione (v.a. uniforme in cerchio di raggio 5cm) e dimensione (ellisse, assi v.a. uniformi in [1,4] cm) stroke
- rumore su tissueprop (ogni valore scalato di percentuale gaussiana,  $\sim 95\%$  contenuta in  $\pm 10\%$ )
- variabilità su raggio antenne (uniforme 11-12cm)
- numero di slice zupal casuale tra 39 e 41

Risultati:

- AUC=0.9841
- accuracy: 98%