Machine Learning
19-Giu-2019 (90 minuti)

Matricola: ____________________
Cognome: __________ Nome: ____________

1) Nella regressione lineare (sia rispetto ai parametri sia rispetto alla variabile indipendente) i dati con cosa sono approssimati nel caso 2D e 3D?

Dispense “Regressione”

2) Quali sono le più note tecniche di riduzione di dimensionalità? Quali i loro tipici utilizzi?

Dispense “Riduzione Dimensionalità”

3) Nell’ambito di classificazione con SVM cosa si intende per pattern linearmente separabili e non linearmente separabili? Fare esempio grafico dei due casi.

Dispense “Classificazione (2)”

4) Cosa si intende per convergenza di un algoritmo di apprendimento iterativo? Accuratezza e loss come si comportano durante le iterazioni in caso di convergenza. Disegnare un semplice grafico.

Dispense “Fondamenti”

Machine Learning	**Matricola: ____________________**
19-Giu-2019 (90 minuti)	**Cognome: __________ Nome: ____________**

5) Date due distribuzioni multinormali identificate dai seguenti parametri:

$\boldsymbol{\mu}_0 = \begin{bmatrix} 4{,}91 \\ -2{,}02 \end{bmatrix}$ $\qquad$ $\boldsymbol{\mu}_1 = \begin{bmatrix} 2{,}23 \\ -3{,}00 \end{bmatrix}$

$\boldsymbol{\Sigma}_0^{-1} = \begin{bmatrix} 0{,}40 & 0{,}18 \\ 0{,}18 & 2{,}23 \end{bmatrix}$ $\qquad$ $\boldsymbol{\Sigma}_1^{-1} = \begin{bmatrix} 0{,}35 & 0{,}17 \\ 0{,}17 & 0{,}93 \end{bmatrix}$

$|\boldsymbol{\Sigma}_0| = 1{,}159678$ $\qquad$ $|\boldsymbol{\Sigma}_1| = 3{,}415143$

$P(w_0) = 0{,}40$ $\qquad$ $P(w_1) = 0{,}60$

Nell'ipotesi dell'impiego di un classificatore di Bayes multinormale, calcolare per il punto $\boldsymbol{x} = \begin{bmatrix} 3{,}80 \\ -2{,}40 \end{bmatrix}$:

- le densità di probabilità condizionali;
- le probabilità a posteriori;
- l'indice della classe restituita in output.

Si ricorda che la densità di probabilità, nel caso della distribuzione multinormale è:

$$p(\boldsymbol{x}) = \frac{1}{(2\pi)^{\frac{d}{2}} \cdot |\boldsymbol{\Sigma}|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}\cdot(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{t}}\cdot\boldsymbol{\Sigma}^{-1}\cdot(\boldsymbol{x}-\boldsymbol{\mu})}$$

<u>**Svolgimento**</u>

$$\boldsymbol{x} - \boldsymbol{\mu}_0 = \begin{bmatrix} 3{,}80 \\ -2{,}40 \end{bmatrix} - \begin{bmatrix} 4{,}91 \\ -2{,}02 \end{bmatrix} = \begin{bmatrix} -1{,}11 \\ -0{,}38 \end{bmatrix}$$

$$p(\boldsymbol{x}|w_0) = \frac{1}{2\pi \cdot \sqrt{1{,}159678}} \cdot e^{-\frac{1}{2}\cdot[-1{,}11 \quad -0{,}38]\cdot\begin{bmatrix} 0{,}40 & 0{,}18 \\ 0{,}18 & 2{,}23 \end{bmatrix}\cdot\begin{bmatrix} -1{,}11 \\ -0{,}38 \end{bmatrix}}$$

$$\begin{bmatrix} 0{,}40 & 0{,}18 \\ 0{,}18 & 2{,}23 \end{bmatrix} \cdot \begin{bmatrix} -1{,}11 \\ -0{,}38 \end{bmatrix} = \begin{bmatrix} 0{,}40 \cdot (-1{,}11) + 0{,}18 \cdot (-0{,}38) \\ 0{,}18 \cdot (-1{,}11) + 2{,}23 \cdot (-0{,}38) \end{bmatrix} = \begin{bmatrix} -0{,}5124 \\ -1{,}0472 \end{bmatrix}$$

$$[-1{,}11 \quad -0{,}38] \cdot \begin{bmatrix} -0{,}5124 \\ -1{,}0472 \end{bmatrix} = (-1{,}11) \cdot (-0{,}5124) + (-0{,}38) \cdot (-1{,}0472) = 0{,}9667$$

$p(\boldsymbol{x}|w_0) = 0{,}147792 \cdot e^{-\frac{0{,}9667}{2}} = 0{,}091145$ (Densità di probabilità condizionale di $\boldsymbol{x}$ data w_0)

$$\boldsymbol{x} - \boldsymbol{\mu}_1 = \begin{bmatrix} 3{,}80 \\ -2{,}40 \end{bmatrix} - \begin{bmatrix} 2{,}23 \\ -3{,}00 \end{bmatrix} = \begin{bmatrix} 1{,}57 \\ 0{,}60 \end{bmatrix}$$

$$p(\boldsymbol{x}|w_1) = \frac{1}{2\pi \cdot \sqrt{3{,}415143}} \cdot e^{-\frac{1}{2}\cdot[1{,}57 \quad 0{,}60]\cdot\begin{bmatrix} 0{,}35 & 0{,}17 \\ 0{,}17 & 0{,}93 \end{bmatrix}\cdot\begin{bmatrix} 1{,}57 \\ 0{,}60 \end{bmatrix}}$$

$$\begin{bmatrix} 0{,}35 & 0{,}17 \\ 0{,}17 & 0{,}93 \end{bmatrix} \cdot \begin{bmatrix} 1{,}57 \\ 0{,}60 \end{bmatrix} = \begin{bmatrix} 0{,}35 \cdot 1{,}57 + 0{,}17 \cdot 0{,}60 \\ 0{,}17 \cdot 1{,}57 + 0{,}93 \cdot 0{,}60 \end{bmatrix} = \begin{bmatrix} 0{,}6515 \\ 0{,}8249 \end{bmatrix}$$

$$[1{,}57 \quad 0{,}60] \cdot \begin{bmatrix} 0{,}6515 \\ 0{,}8249 \end{bmatrix} = 1{,}57 \cdot 0{,}6515 + 0{,}60 \cdot 0{,}8249 = 1{,}517795$$

$p(\boldsymbol{x}|w_1) = 0{,}086122 \cdot e^{-\frac{1{,}517795}{2}} = 0{,}040321$ (Densità di probabilità condizionale di $\boldsymbol{x}$ data w_1)

$p(\boldsymbol{x}) = 0{,}091145 \cdot 0{,}40 + 0{,}040321 \cdot 0{,}60 = 0{,}060651$ (Densità di probabilità assoluta dato $\boldsymbol{x}$)

$p(w_0|\boldsymbol{x}) = \frac{0{,}091145\cdot 0{,}40}{0{,}060651} = 0{,}601$ (Probabilità a posteriori di w_0 dato $\boldsymbol{x}$)

$p(w_1|\boldsymbol{x}) = \frac{0{,}040321\cdot 0{,}60}{0{,}060651} = 0{,}399$ (Probabilità a posteriori di w_1 dato $\boldsymbol{x}$)

Indice della classe restituita: **0**

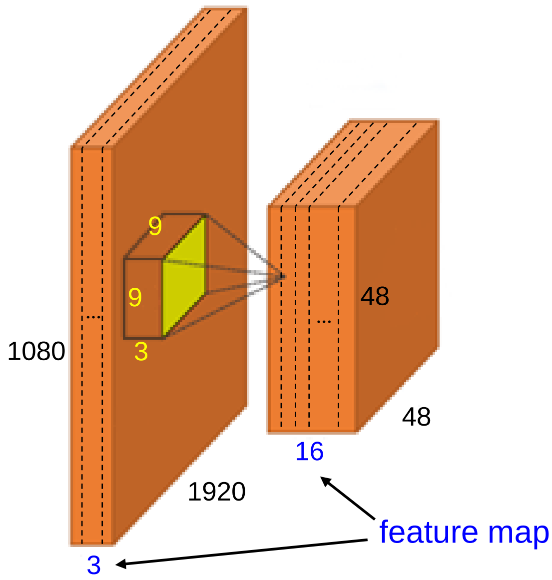
6) Dati un volume di input ed uno di output relativi a un livello di convoluzione in una CNN, aventi le seguenti dimensioni:

- *Volume Input*: $3 \times 1920 \times 1080$
- *Volume Output*: $16 \times 48 \times 48$

Considerando che ciascun filtro abbia dimensioni:

- *Dimensione Filtro*: $3 \times 9 \times 9$

Si calcoli il numero totale di connessioni e di pesi del livello (<u>considerando i bias</u>) motivando la risposta.

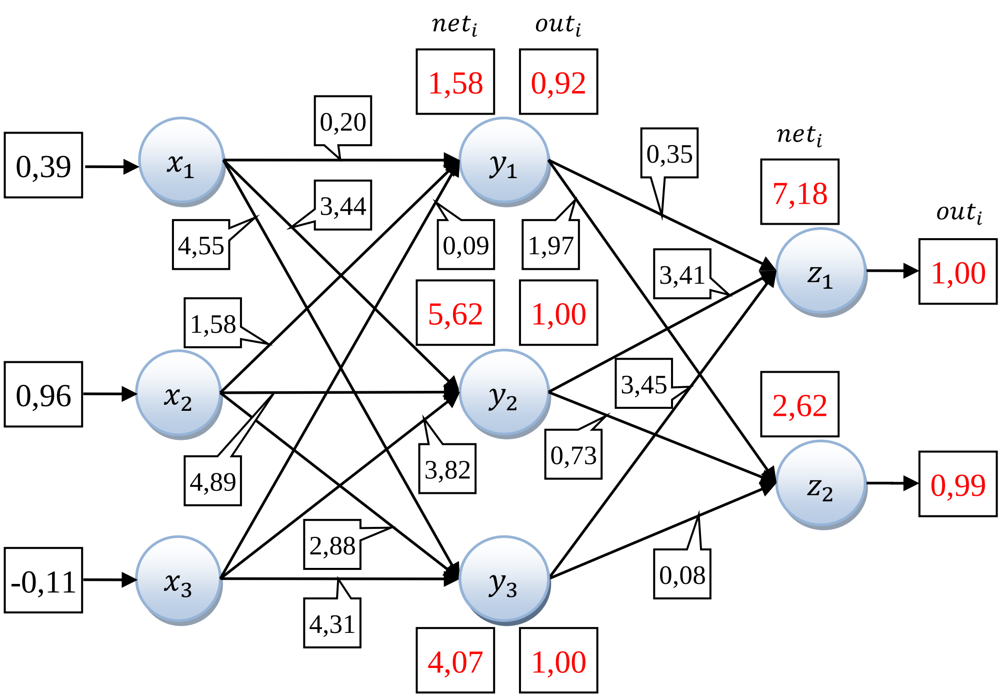


Svolgimento

Ogni neurone del livello di output ($16 \times 48 \times 48$) è connesso con tanti neuroni del livello di input pari alla dimensione del filtro ($3 \times 9 \times 9$).
Pertanto **il numero totale di connessioni** è $(16 \times 48 \times 48) \cdot (3 \times 9 \times 9) = 8957952$.

Il **numero totale di pesi**, invece, risulta molto più piccolo giacché in una CNN i pesi di ciascun filtro sono condivisi da tutti i neuroni contenuti in una stessa feature map. Visto che il numero di feature map è uguale a 16, ed il numero di input per ciascun filtro è pari a $(3 \times 9 \times 9)$, il numero totale di pesi (considerando il bias) è $\left((3 \times 9 \times 9) + 1\right) \times 16 = 3904$.

7) Data la seguente rete neurale, calcolare net_i e out_i di ogni neurone al seguito del passo forward propagation del pattern di input, utilizzando la *tangente iperbolica* ($\text{Tanh}(net)$) come funzione di attivazione.



Svolgimento

$$net(y_1) = 0{,}20 \cdot 0{,}39 + 1{,}58 \cdot 0{,}96 + 0{,}09 \cdot (-0{,}11) = 1{,}58 \qquad out(y_1) = 0{,}92$$
$$net(y_2) = 3{,}44 \cdot 0{,}39 + 4{,}89 \cdot 0{,}96 + 3{,}82 \cdot (-0{,}11) = 5{,}62 \qquad out(y_2) = 1{,}00$$
$$net(y_3) = 4{,}55 \cdot 0{,}39 + 2{,}88 \cdot 0{,}96 + 4{,}31 \cdot (-0{,}11) = 4{,}07 \qquad out(y_3) = 1{,}00$$

$$net(z_1) = 0{,}35 \cdot 0{,}92 + 3{,}41 \cdot 1{,}00 + 3{,}45 \cdot 1{,}00 = 7{,}18 \qquad out(z_1) = 1{,}00$$
$$net(z_2) = 1{,}97 \cdot 0{,}92 + 0{,}73 \cdot 1{,}00 + 0{,}08 \cdot 1{,}00 = 2{,}62 \qquad out(z_2) = 0{,}99$$