CAFFÈ BELTRAMI

UNIVERSITÀ DEGLI STUDI DI PAVIA

Caffè Beltrami presenta...

Massimiliano Ghiotto

Massimiliano Ghiotto è dottorando del XXXIX ciclo presso il Dipartimento di Matematica dell'Università di Pavia. La sua ricerca si concentra nell'approssimazione di PDEs tramite operatori neurali, con particolare attenzione allo sviluppo di operatori neurali applicabili a problemi definiti su geometrie multipatch.

CAFFÈ BELTRAMI

UNIVERSITÀ DEGLI STUDI DI PAVIA

Operatori neurali

Deep Learning per matematici

Deep Learning

Applicazioni interessanti

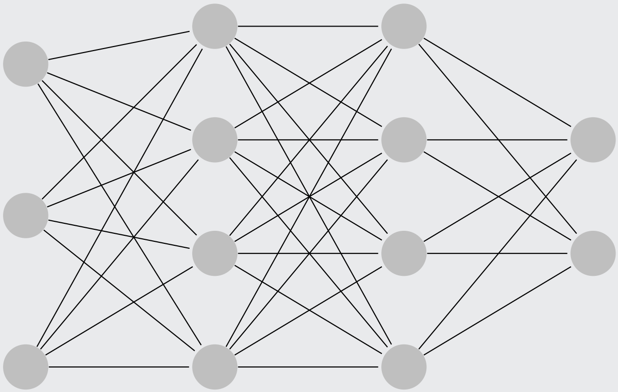
Operatori neurali

Approsimazione di equazioni differenziali

Deep Learning

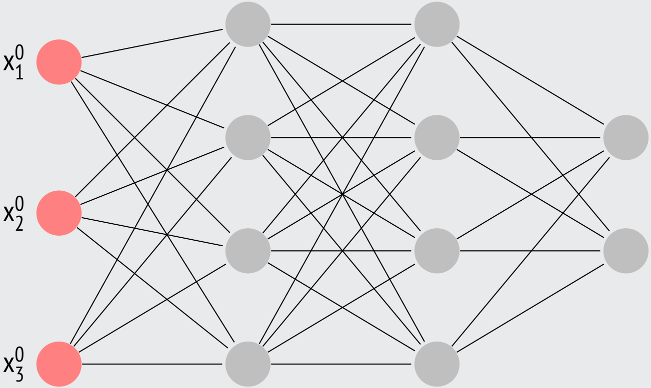
Percettrone

Multi-layer Perceptron



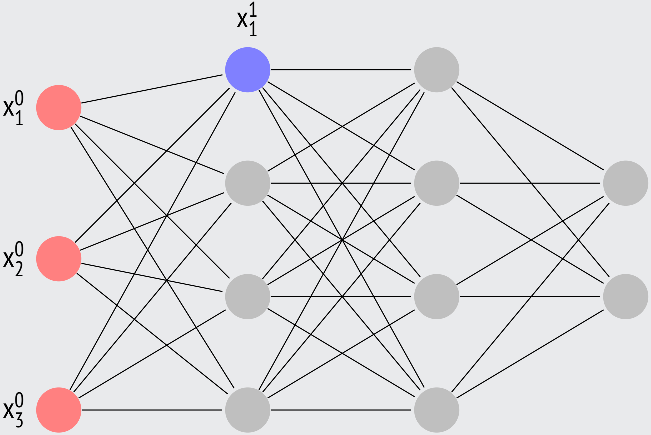
Esempio di architettura di un Multi-layer Perceptron con 3 neuroni di input, 4 neuroni nel primo hidden layer, 4 neuroni nel secondo hidden layer e 2 neuroni di output.

Multi-layer Perceptron



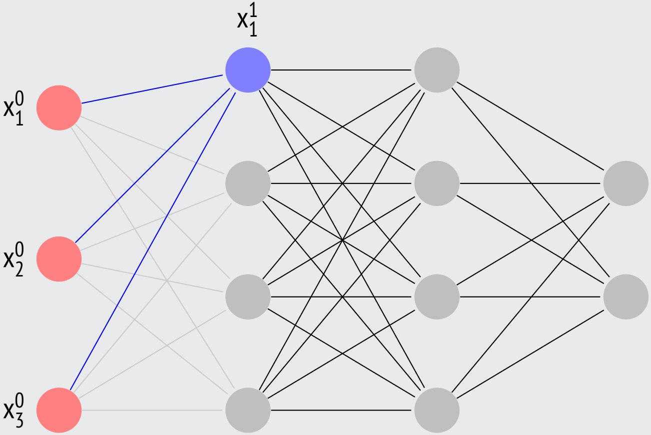
Layer di input con 3 neuroni.

Multi-layer Perceptron



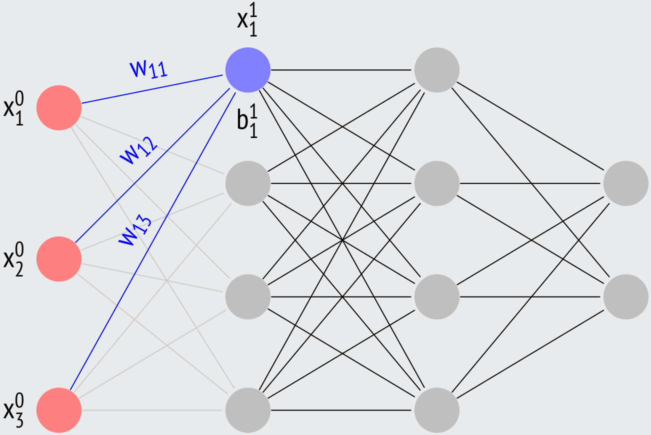
Vogliamo definire il valore del neurone x_1^1.

Multi-layer Perceptron



Evidenziamo le connesioni con i neuroni nel layer precedente.

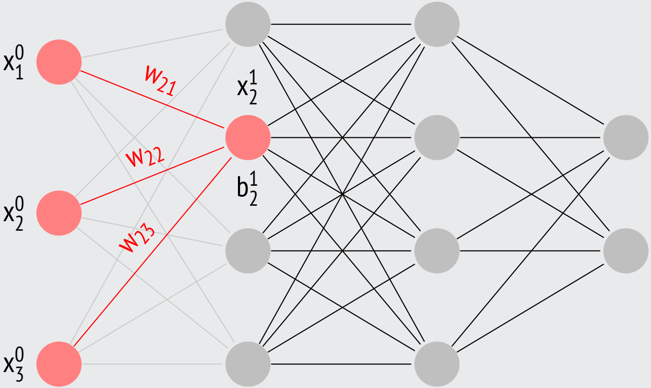
Multi-layer Perceptron



$$x_1^1 = \sigma\left(b_1^1 + \sum_{i=1}^{3} w_{1i} x_i^0\right)$$

CAFFÈ BELTRAMI

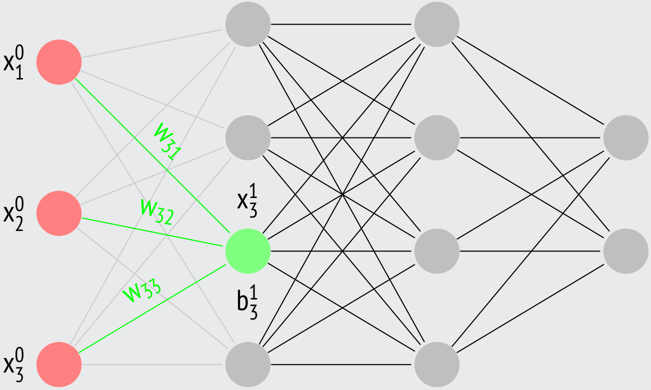
Multi-layer Perceptron



$$x_1^1 = \sigma\left(b_1^1 + \sum_{i=1}^{3} w_{1i} x_i^0\right)$$

$$x_2^1 = \sigma\left(b_2^1 + \sum_{i=1}^{4} w_{2i} x_i^0\right)$$

Multi-layer Perceptron

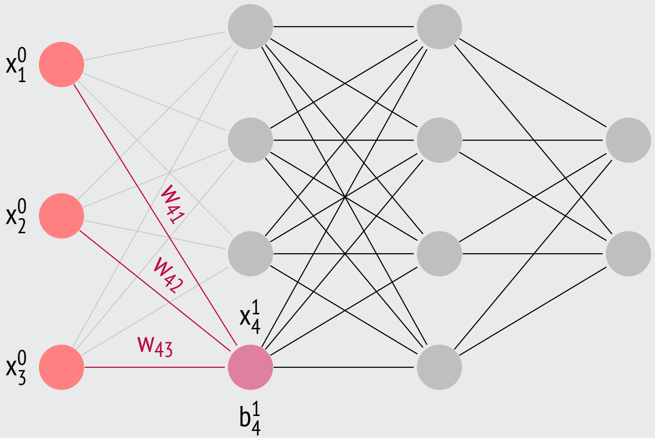


$$x_1^1 = \sigma\left(b_1^1 + \sum_{i=1}^{3} w_{1i} x_i^0\right)$$

$$x_2^1 = \sigma\left(b_2^1 + \sum_{i=1}^{4} w_{2i} x_i^0\right)$$

$$x_3^1 = \sigma\left(b_3^1 + \sum_{i=1}^{4} w_{3i} x_i^0\right)$$

Multi-layer Perceptron



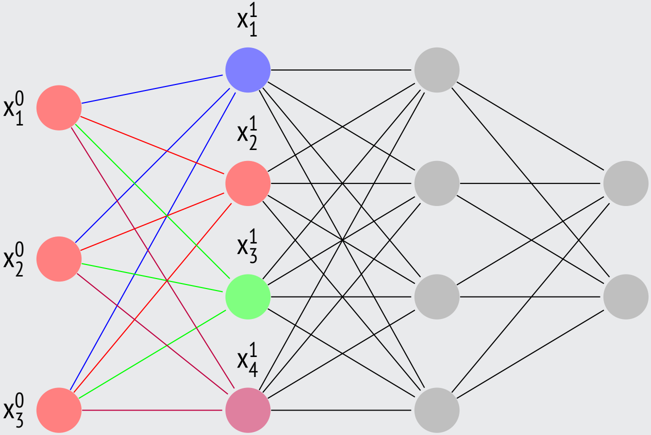
$$x_1^1 = \sigma\left(b_1^1 + \sum_{i=1}^{3} w_{1i}x_i^0\right)$$

$$x_2^1 = \sigma\left(b_2^1 + \sum_{i=1}^{4} w_{2i}x_i^0\right)$$

$$x_3^1 = \sigma\left(b_3^1 + \sum_{i=1}^{4} w_{3i}x_i^0\right)$$

$$x_4^1 = \sigma\left(b_4^1 + \sum_{i=1}^{2} w_{4i}x_i^0\right)$$

Multi-layer Perceptron

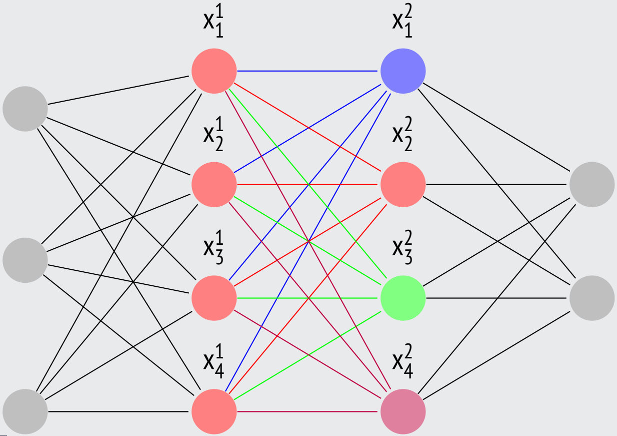


Dato $\mathbf{x}^0 \in \mathbb{R}^3$

$$\mathbf{x}^1 := \sigma\left(\mathbf{W}_1\mathbf{x}^0 + \mathbf{b}_1\right) \in \mathbb{R}^4$$

con $\mathbf{W} \in \mathbb{R}^{4\times 3}$, $\mathbf{b} \in \mathbb{R}^4$ e σ funzione di attivazione non lineare applicata componente per componente.

Multi-layer Perceptron



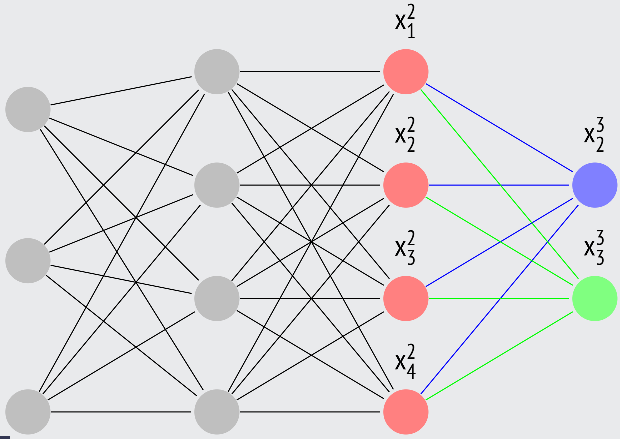
Dato $\mathbf{x}^0 \in \mathbb{R}^3$

$$\mathbf{x}^1 := \sigma\left(\mathbf{W}_1\mathbf{x}^0 + \mathbf{b}_1\right) \in \mathbb{R}^4$$

$$\mathbf{x}^2 := \sigma\left(\mathbf{W}_2\mathbf{x}^1 + \mathbf{b}_2\right) \in \mathbb{R}^4$$

con $\mathbf{W}_2 \in \mathbb{R}^{4\times 4}$ e $\mathbf{b} \in \mathbb{R}^4$.

Multi-layer Perceptron



Dato $\mathbf{x}^0 \in \mathbb{R}^3$

$$\mathbf{x}^1 := \sigma\left(\mathbf{W}_1\mathbf{x}^0 + \mathbf{b}_1\right) \in \mathbb{R}^4$$

$$\mathbf{x}^2 := \sigma\left(\mathbf{W}_2\mathbf{x}^1 + \mathbf{b}_2\right) \in \mathbb{R}^4$$

$$\mathbf{x}^3 := \sigma\left(\mathbf{W}_3\mathbf{x}^2 + \mathbf{b}_3\right) \in \mathbb{R}^2$$

con $\mathbf{W}_3 \in \mathbb{R}^{2\times 4}$ e $\mathbf{b} \in \mathbb{R}^2$.

Multi-layer Perceptron

Definizione

Multi-layer Perceptron Un Multi-layer Perceptron (MLP) con input $\mathbf{x}^0 \in \mathbb{R}^{d_0}$, L hidden layers con $d_1, \dots, d_L$ neuroni e output $\mathbf{x}^L \in \mathbb{R}^{d_L}$ è definito come

$$
\begin{aligned}
\mathbf{x}^1 &:= \sigma\left(\mathbf{W}_1 \mathbf{x}^0 + \mathbf{b}_1\right) \in \mathbb{R}^{d_1} \\
&\vdots \\
\mathbf{x}^L &:= \sigma\left(\mathbf{W}_L \mathbf{x}^{L-1} + \mathbf{b}_L\right) \in \mathbb{R}^{d_L}
\end{aligned}
$$

dove $\mathbf{W}_\ell \in \mathbb{R}^{d_\ell \times d_{\ell-1}}$ e $\mathbf{b}_\ell \in \mathbb{R}^{d_\ell}$ per $\ell = 1, \dots, L$ e σ è una funzione di attivazione non lineare applicata componente per componente.

Metodo del gradiente

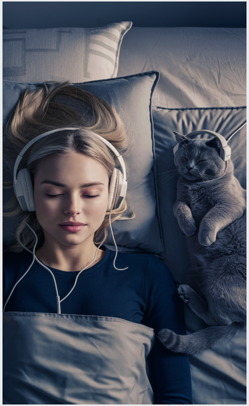
Architetture diverse per problemi diversi

Transformer

Applicazioni interessanti

Applicazioni alle immagini

Siti per generazione d'immagini: DALL-E3 DALL·E, Midjourney e Ideogram.

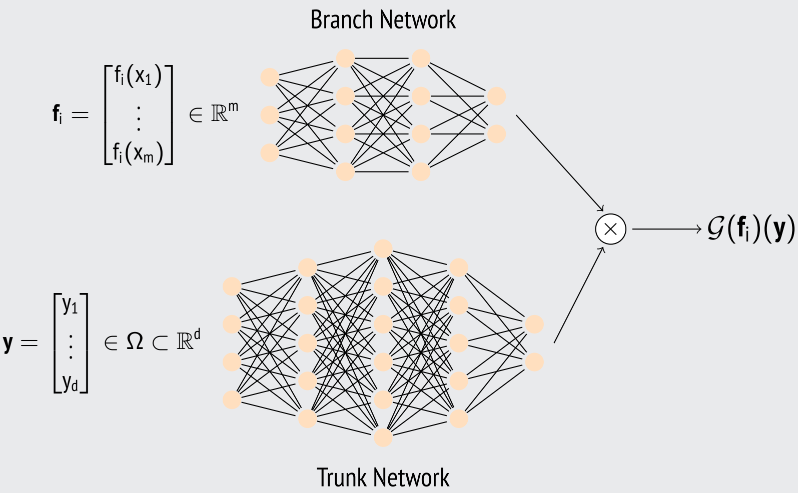


Applicazioni ai video

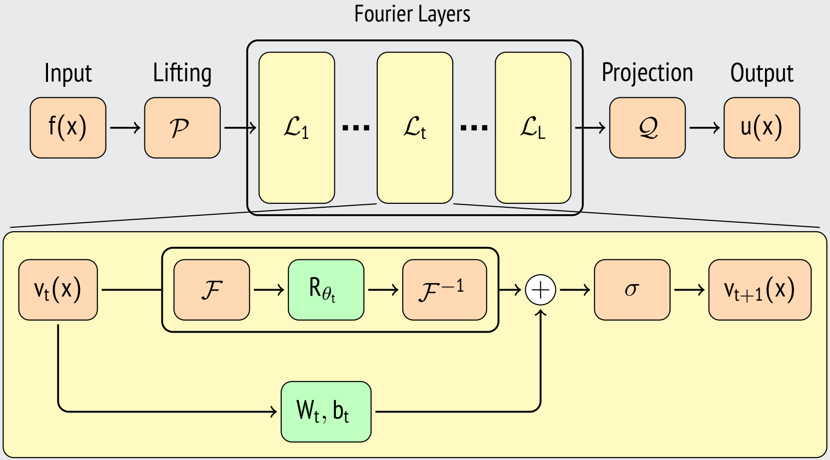
Deep learning per la generazione di video SORA o per modificare video anche in tempo reale NVIDIA Broadcast .

Operatori neurali

Deep Operator Network



Fourier Neural Operator



Approsimazione di equazioni differenziali