

Relazione Progetto IA Procelli

Alessio

Introduzione:	2
Scopo:	2
Breve introduzione:	2
Dataset:	3
Risultati	4

Introduzione:

Scopo:

In questo esercizio si implementa (in un linguaggio di programmazione a scelta) l'architettura multi-layer perceptron per il caso particolare della classificazione binaria, ed il relativo algoritmo backpropagation per il calcolo del gradiente (esposto in classe e descritto in R&N 2021 cap 21). Si implementi quindi un semplice ottimizzatore basato sul metodo del gradiente stocastico con momentum e si verifichi il corretto funzionamento del codice sul dataset dataset Zelando scegliendo a piacere due classi come positiva e negativa. Si riportino quindi gli andamenti della loss sul training set e sul test set in funzione del numero di iterazioni.

Breve introduzione:

Creo un MLP con due strati nascosti.

- Nodi input: $28 \times 28 = 784$ uno per ogni pixel.
- Nodi nascosti : 10,10 che formano gli strati nascosti.
- Nodo output: 1

Ogni nodo ha una funzione di attivazione Relu gli strati nascosti e sigmoide l'output:

- $f(x) = \max(0, x)$ funzione Rectifier
- $f(x) = \frac{1}{1-e^{-x}}$ funzione sigmoide.

Loss cross entropy con y risultato di computazione e \bar{y} valore target.

$$\mu(y, \bar{y}) = -(\bar{y} \ln(y) + (1 - \bar{y}) \ln(1 - y))$$

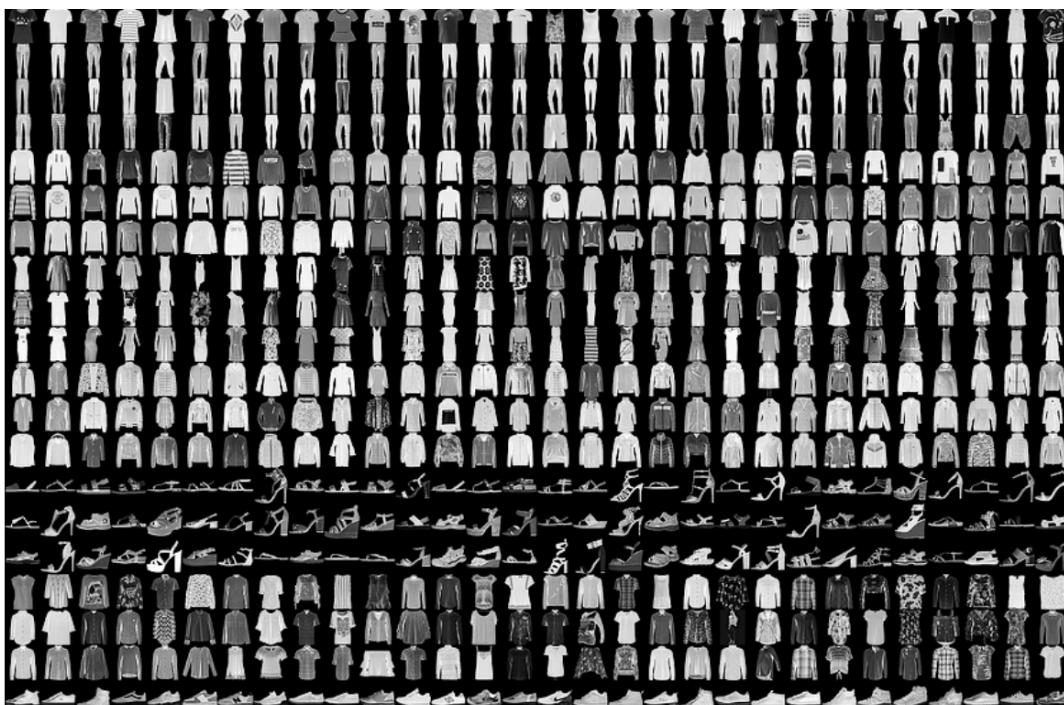
Gradiente stocastico con momentum

$$\Delta w_{t+1} = w_t - (\eta \nabla u(w_t) + \alpha \Delta(w_t))$$

dove η è il tasso di apprendimento e α tasso di rilevanza del momentum

Dataset:

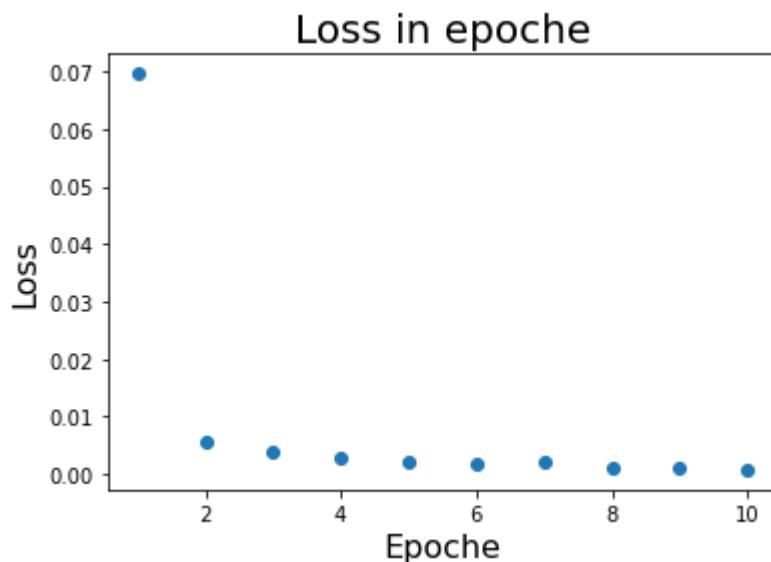
Fashion-MNIST è un set di dati delle immagini degli articoli di Zalando , costituito da un set di formazione di 60.000 esempi e un set di test di 10.000 esempi. Ogni esempio è un'immagine in scala di grigi 28x28, associata a un'etichetta di 10 classi.



Etichetta	Descrizione
0	T-shirt/top
1	Pantaloni
2	Maglione
3	Vestire
4	Cappotto
5	Sandalo
6	Camicia
7	Sneaker
8	Borsa
9	Stivaletto

Ho scelto come etichette 1 e 6. Con 1 considerato come output 0 mentre 6 con output 1.

Risultati



Come si può vedere dal grafico che mette in relazione il valore della loss con il passare delle epocha si ha che la rete già dopo la prima epoca migliora notevolmente ed continua a migliorare facendo tendere la loss a zero con il crescere delle epocha. Ci sono alcuni “rimbalzi leggeri” verso l’alto dovuti al momentum. Sotto abbiamo i risultati della computazione di “prova”.

