

RICONOSCIMENTO IMMAGINI CON KERAS E TENSORFLOW

Gabriele Baldi

https://github.com/bnznamco/keras_model_and_slides_notabugconf

RICONOSCIMENTO IMMAGINI

Attività tipica del cervello umano

Task molto difficile per una macchina

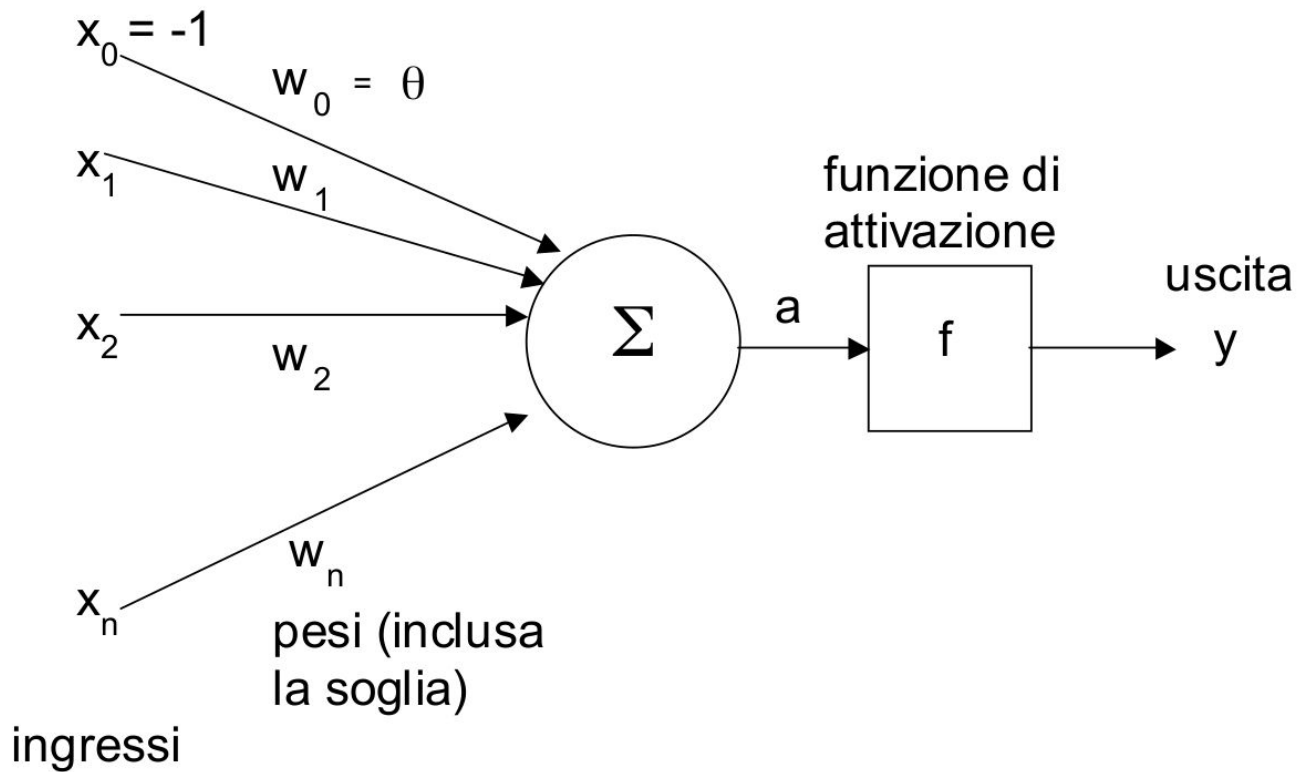
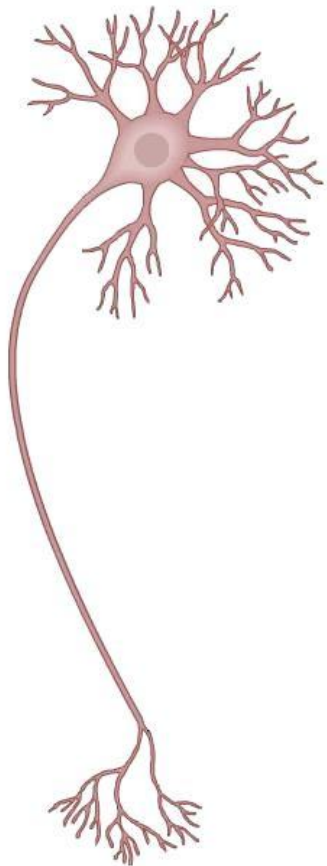


APPRENDIMENTO E GENERALIZZAZIONE

INSEGNARE
ALLE
MACCHINE A
IMPARARE???



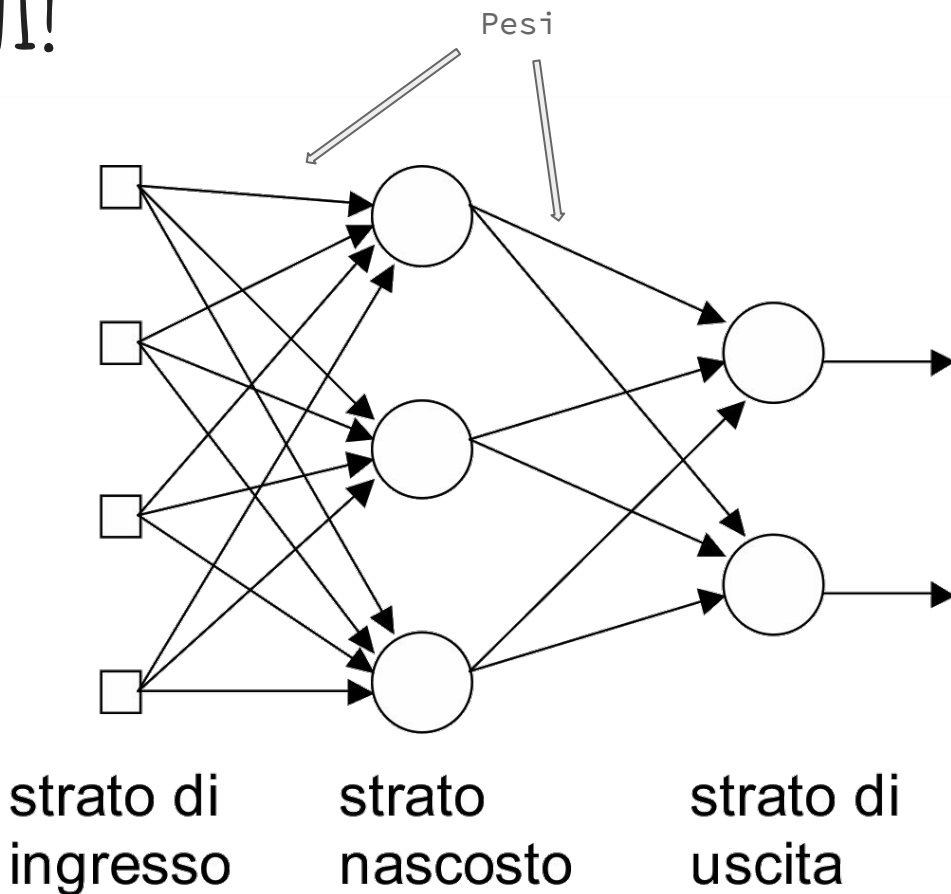
NEURONI



CONNETTIAMO I NEURONI!

RETI STRATIFICATE COMPLETAMENTE CONNESSE

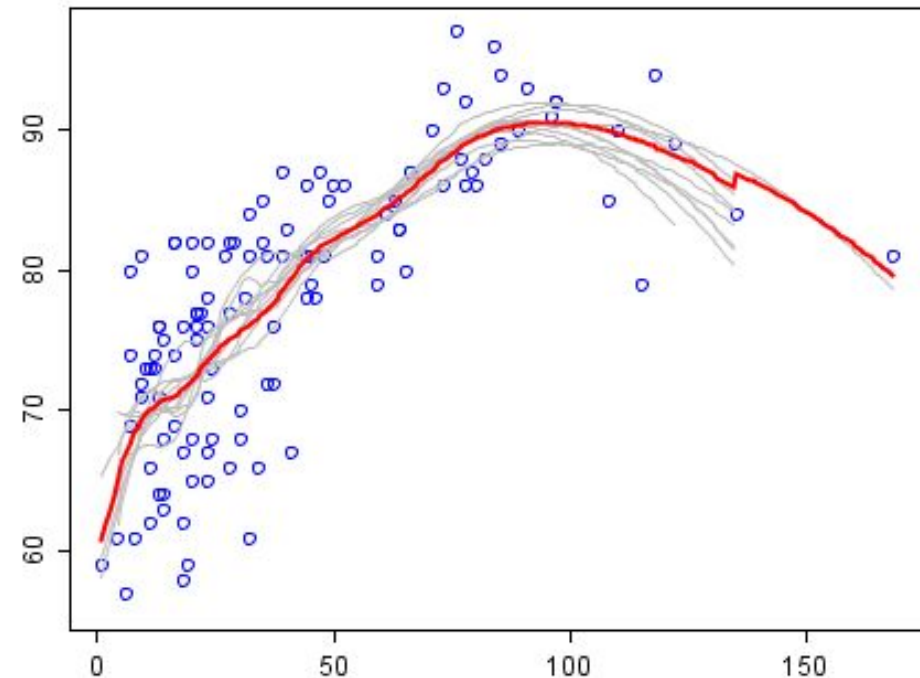
- Le connessioni tra i neuroni di una rete stratificata sono rappresentate mediante tante matrici quante sono le coppie di strati adiacenti.
- Ogni matrice contiene i pesi delle connessioni tra le coppie di neuroni di due strati adiacenti.



The background is a vibrant yellow field filled with a complex network of thin, dark lines connecting numerous small, dark circular nodes. Some nodes are larger than others, and the connections form a dense, web-like structure that spans the entire frame. In the center, there is a white rectangular box containing the text.

COME POSSIAMO
USARE LA RETE?

REGRESSIONE E CLASSIFICAZIONE



REGRESSIONE

Permette di trovare una relazione funzionale tra variabili misurate sulla base di dati campionati.

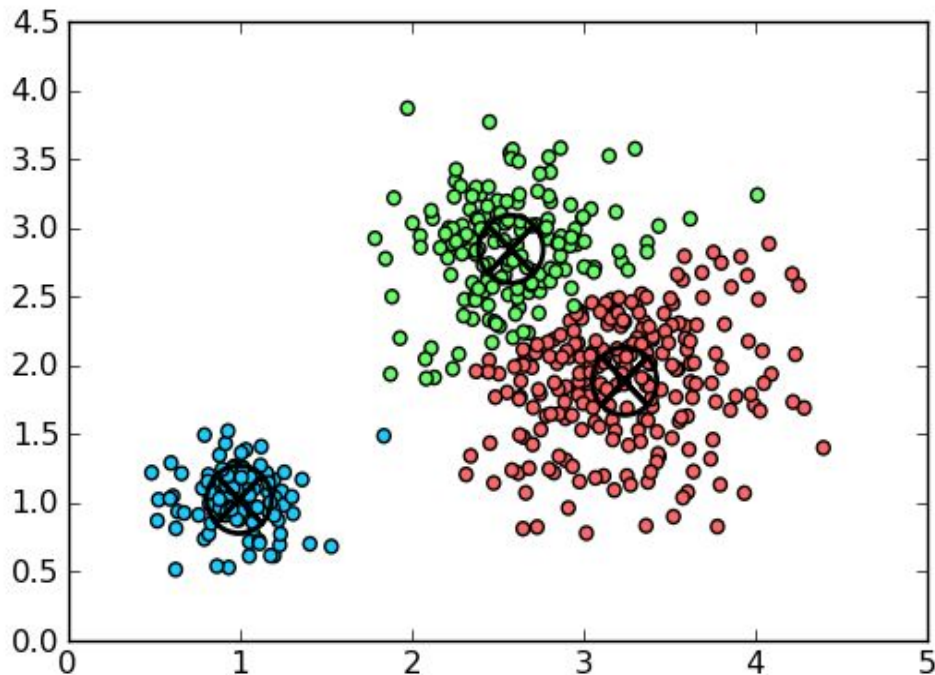
- Trovare correlazioni tra dati apparentemente non correlati
- Prevedere un valore continuo

REGRESSIONE E CLASSIFICAZIONE

CLASSIFICAZIONE

Permette di selezionare e raggruppare elementi omogenei in base alla loro somiglianza (concepita in termini di distanza in uno spazio multidimensionale).

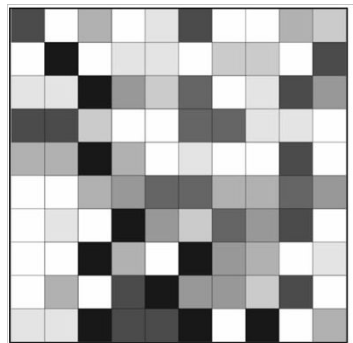
- Distanza euclidea nello spazio delle caratteristiche



CHE TIPO DI PROBLEMA È
IL RICONOSCIMENTO
IMMAGINI?



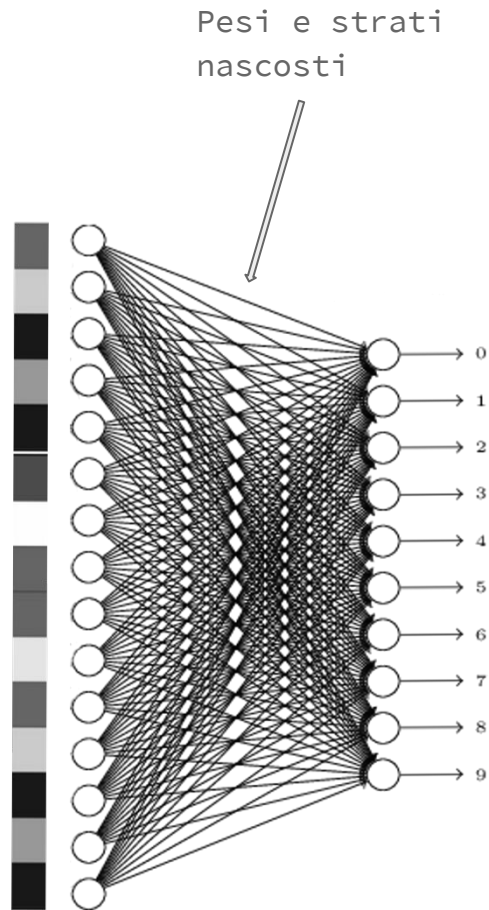
INPUT E OUTPUT



[INPUT] Feature:

Un'immagine è un array di pixels.

Il numero di pixel può essere interpretato come il numero di feature.



[OUTPUT] Category:

1. Car 🚗
2. Human 🤖
3. Rocket 🚀
4. Cat 🐱
5. Bicycle 🚲
6. Toilet 🚽
7. Boat 🚢
8. Clock 🕒
9. Cup ☕

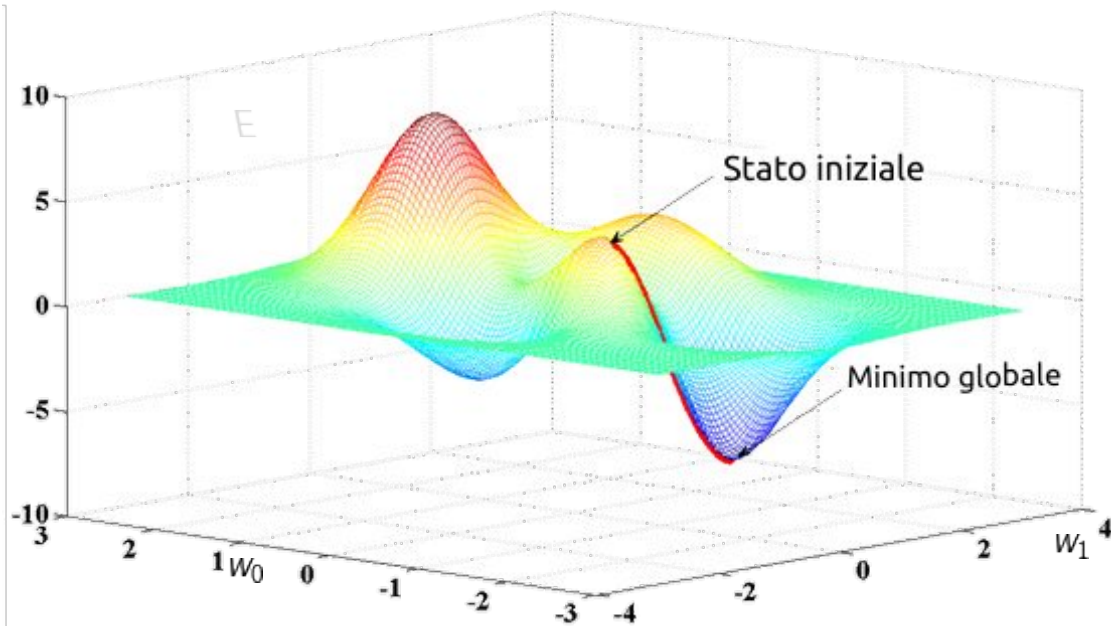
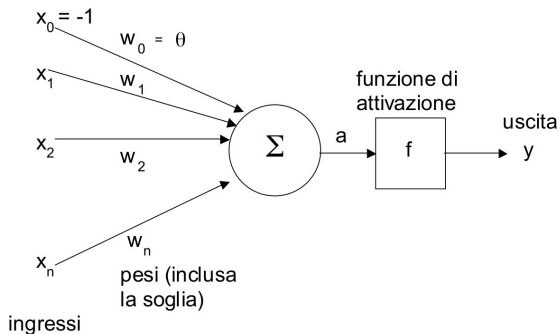
ERRORE DELLA RETE

FUNZIONE DI LOSS

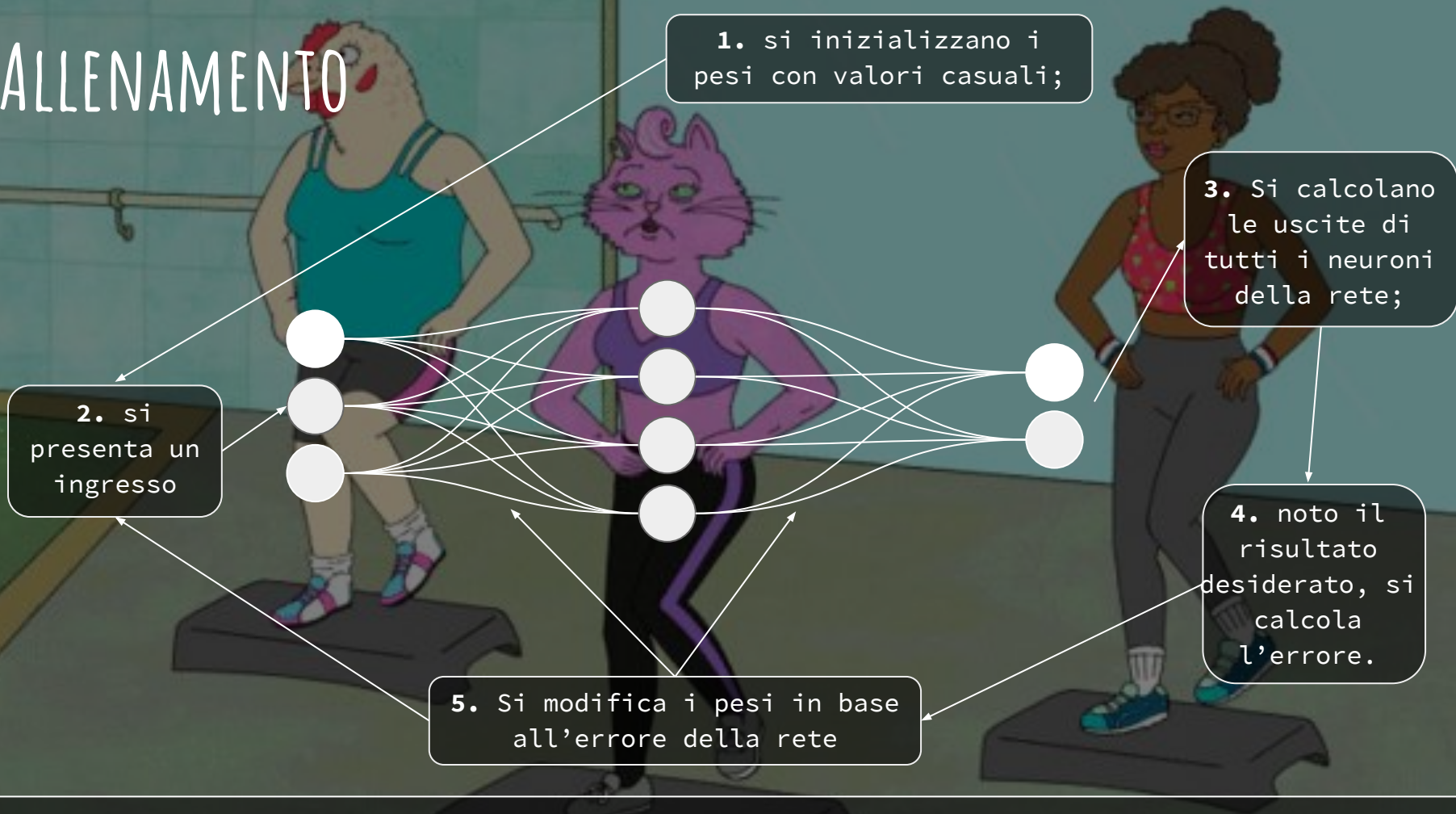
Quantifica il concetto di errore. Es: $\text{Errore} = \text{risultato atteso} - \text{risultato della rete}$

COME RIDUCIAMO L'ERRORE?

Modificando i pesi



ALLENAMENTO



Un ciclo di presentazione degli esempi del training set è detto **epoca**.



TUTTO BELLO,
MA È COSÌ SEMPLICE
CON LE IMMAGINI?

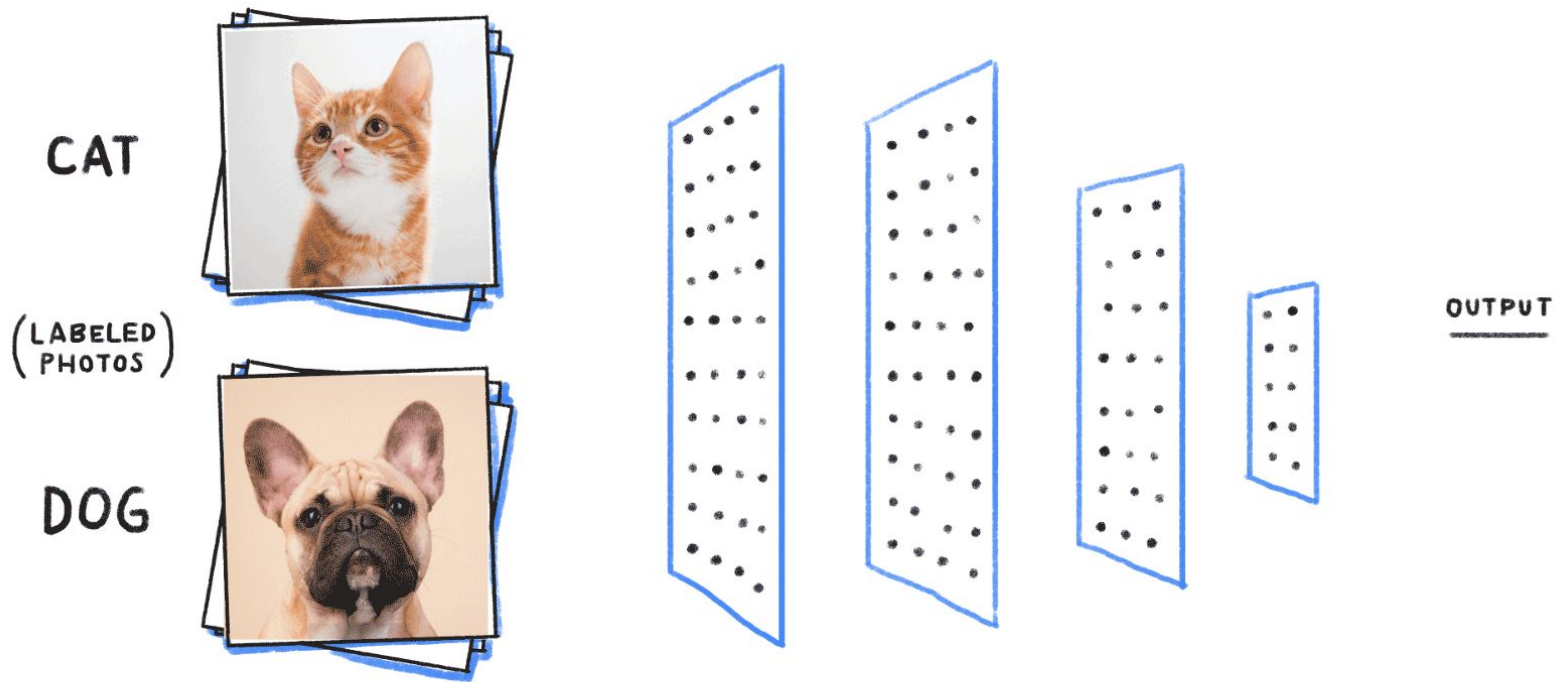
RICONOSCIMENTO IMMAGINI

1. Indipendenza dal contesto.
2. Indipendenza dalla posizione.

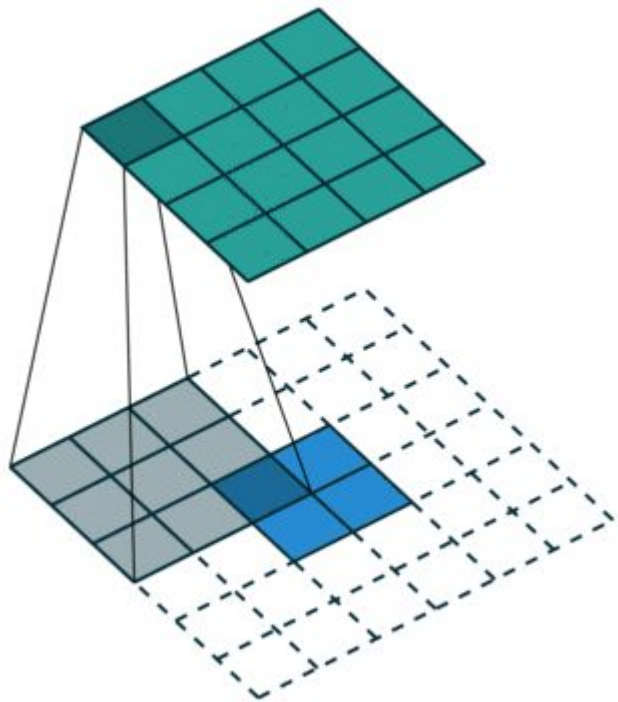
Scelta delle feature



RETI CONVOLUZIONALI



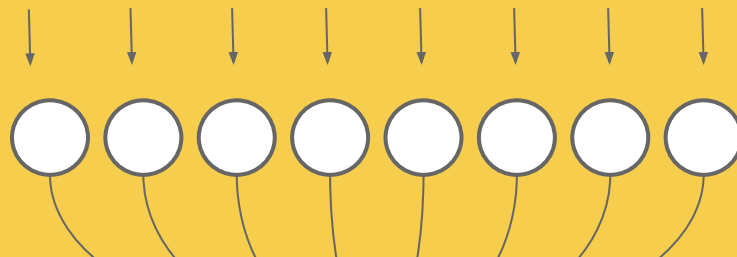
3 TRUCCHI



Conv2d - Convoluzione

Max Pooling

ReLu - Rectified Linear unit



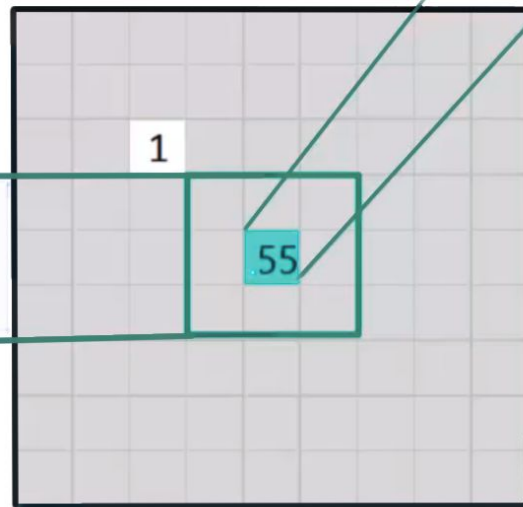
CONV2D

1	-1	-1
-1	1	-1
-1	-1	1

1	1	-1
1	1	1
-1	1	1

$$\frac{1 + 1 - 1 + 1 + 1 + 1 - 1 + 1 + 1}{9} = .55$$

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1



CONV2D

1	-1	-1
-1	1	-1
-1	-1	1

1	-1	1
-1	1	-1
1	-1	1

-1	-1	1
-1	1	-1
1	-1	-1

-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1



0.33	-0.11	0.55	0.33	0.11	-0.11	0.77
0.33	-0.55	0.11	-0.11	0.11	-0.55	0.33
-0.11	-0.55	0.77	-0.11	0.11	0.33	0.55
0.55	0.11	-0.11	1.00	-0.11	0.33	-0.11
0.33	-0.11	0.11	-0.11	1.00	-0.33	0.11
0.11	-0.11	0.33	0.33	-0.33	0.55	-0.33
-0.11	0.11	0.55	-0.11	0.11	-0.33	1.00
-0.55	0.11	-0.11	0.11	-0.33	1.00	-0.11
0.77	-0.55	-0.11	0.11	-0.11	0.33	-0.11
0.33	0.33	-0.11	0.55	0.11	-0.11	0.77

CONV2D

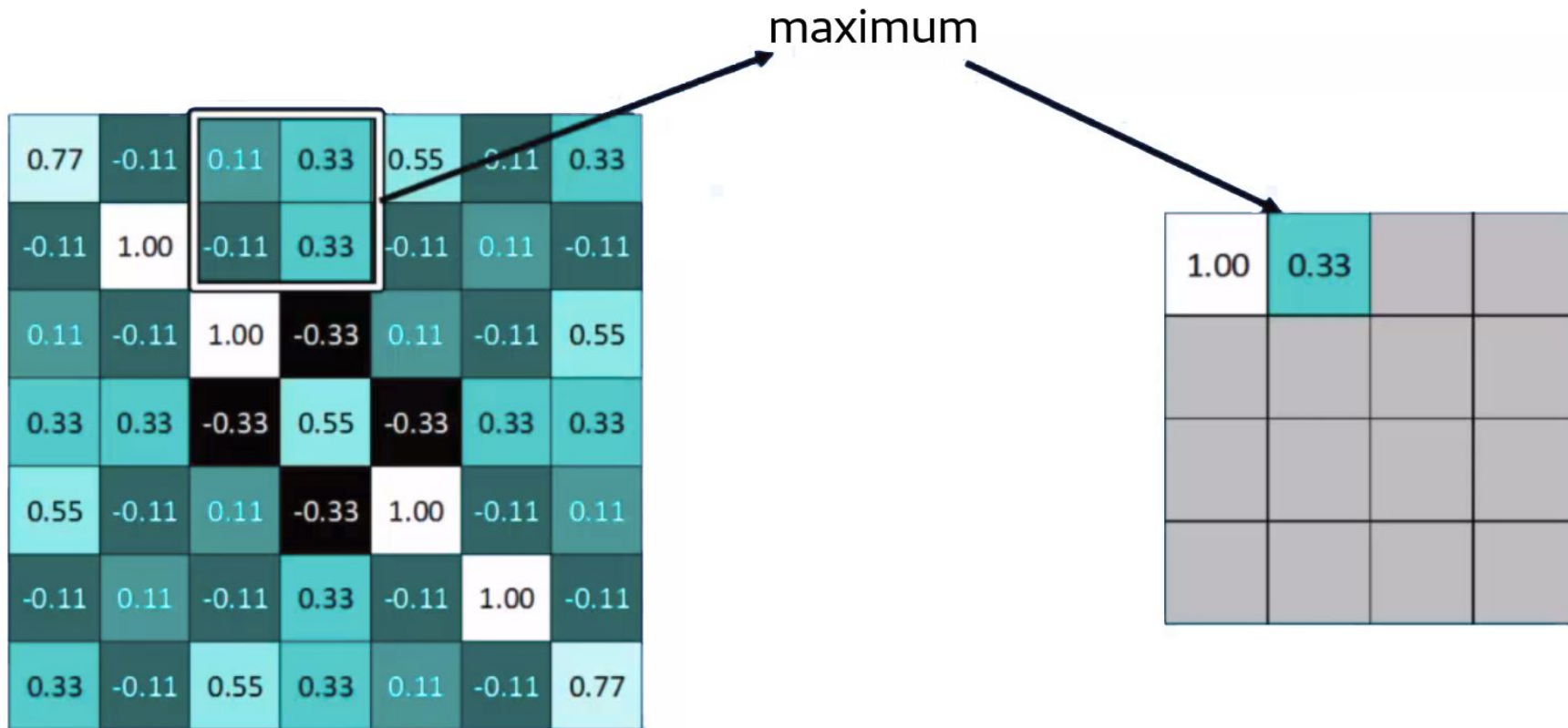
Permette alla rete di cercare anche solo somiglianze parziali.

Trasforma l'immagine in uno stack, creando una mappa di occorrenza di ogni feature.

Bisogna scegliere:

1. Lo shape delle feature
2. Il numero di feature
3. Lo stride (spostamento della finestra)

MAX POOLING



MAX POOLING

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

max pooling



1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

MAX POOLING

È molto conveniente ridurre le dimensioni dei pattern per snellire i calcoli.

Rende le feature meno sensibili alla posizione in cui vengono trovate.

Bisogna scegliere:

1. Lo shape del pool di partenza
2. Lo stride (movimento della finestra) che di default è uguale allo shape

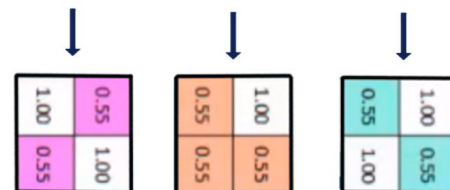
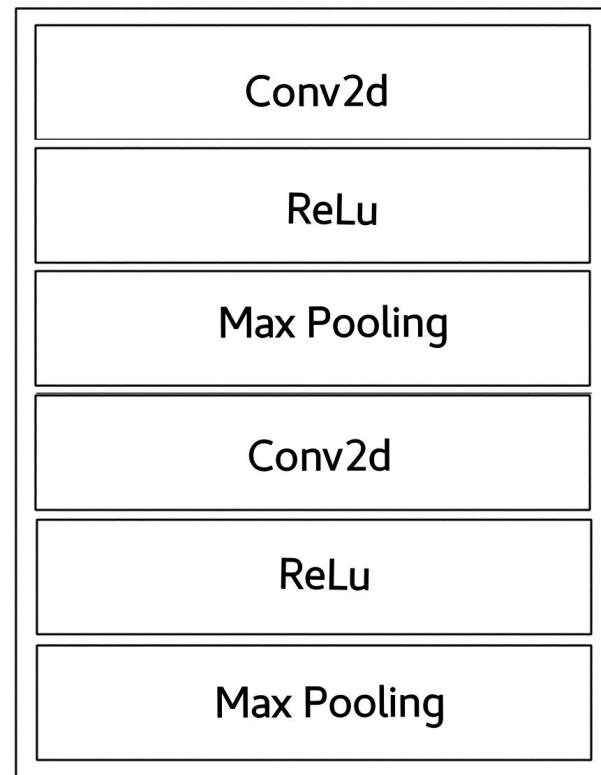
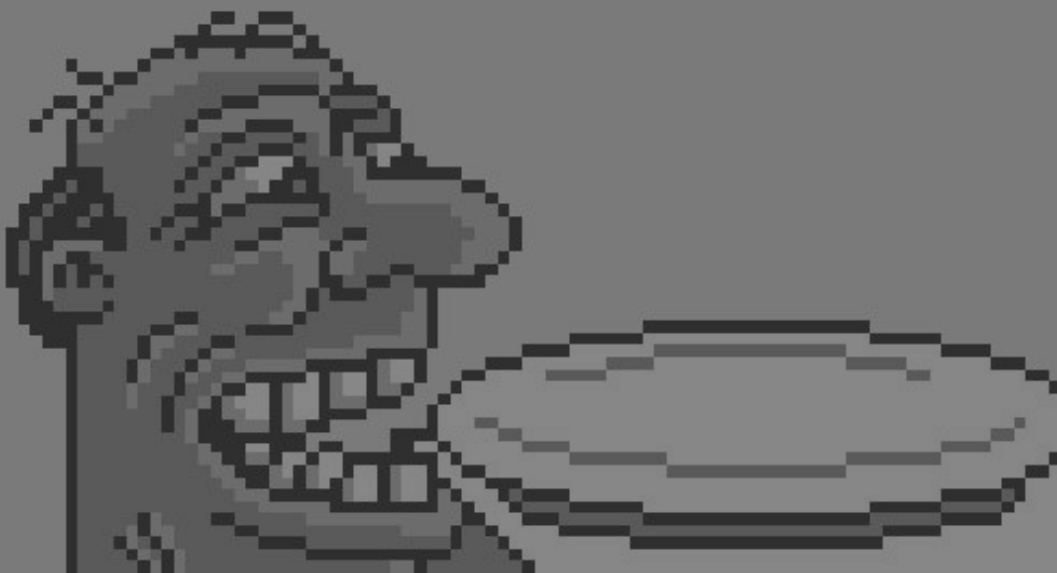
RECTIFIED LINEAR UNIT - RELU

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77



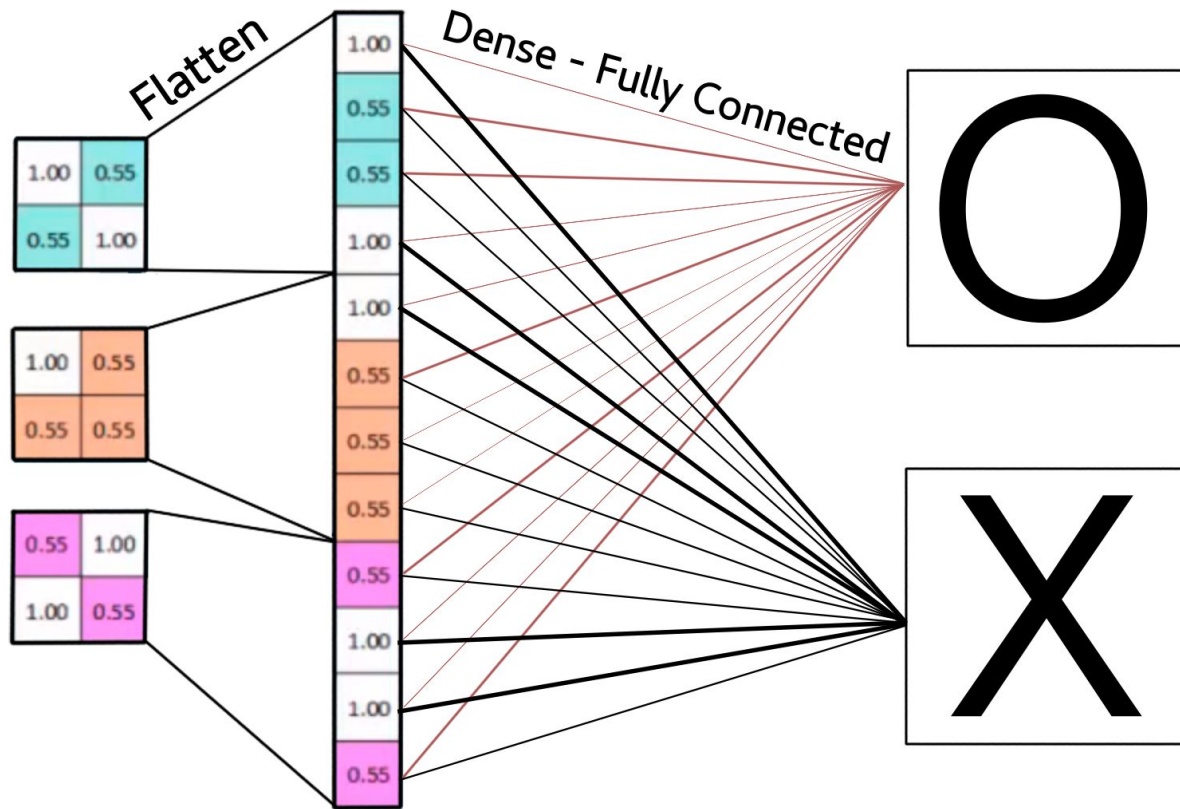
0.77	0	0.11	0.33	0.55	0	0.33
0	1.00	0	0.33	0	0.11	0
0.11	0	1.00	0	0.11	0	0.55
0.33	0.33	0	0.55	0	0.33	0.33
0.55	0	0.11	0	1.00	0	0.11
0	0.11	0	0.33	0	1.00	0
0.33	0	0.55	0.33	0.11	0	0.77

MULTISTRATO



FLATTEN E FULLY CONNECTED

Ogni feature risulterà più o meno significativa nel prevedere la categoria.



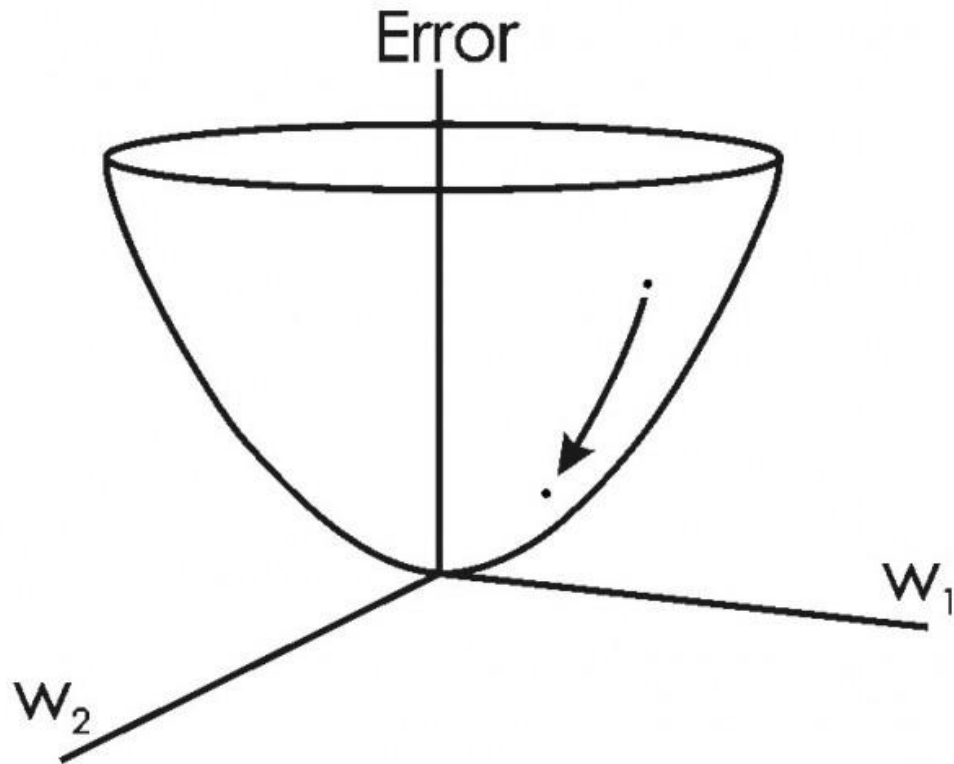
LOSS FUNCTION E OPTIMIZER

Loss Function:

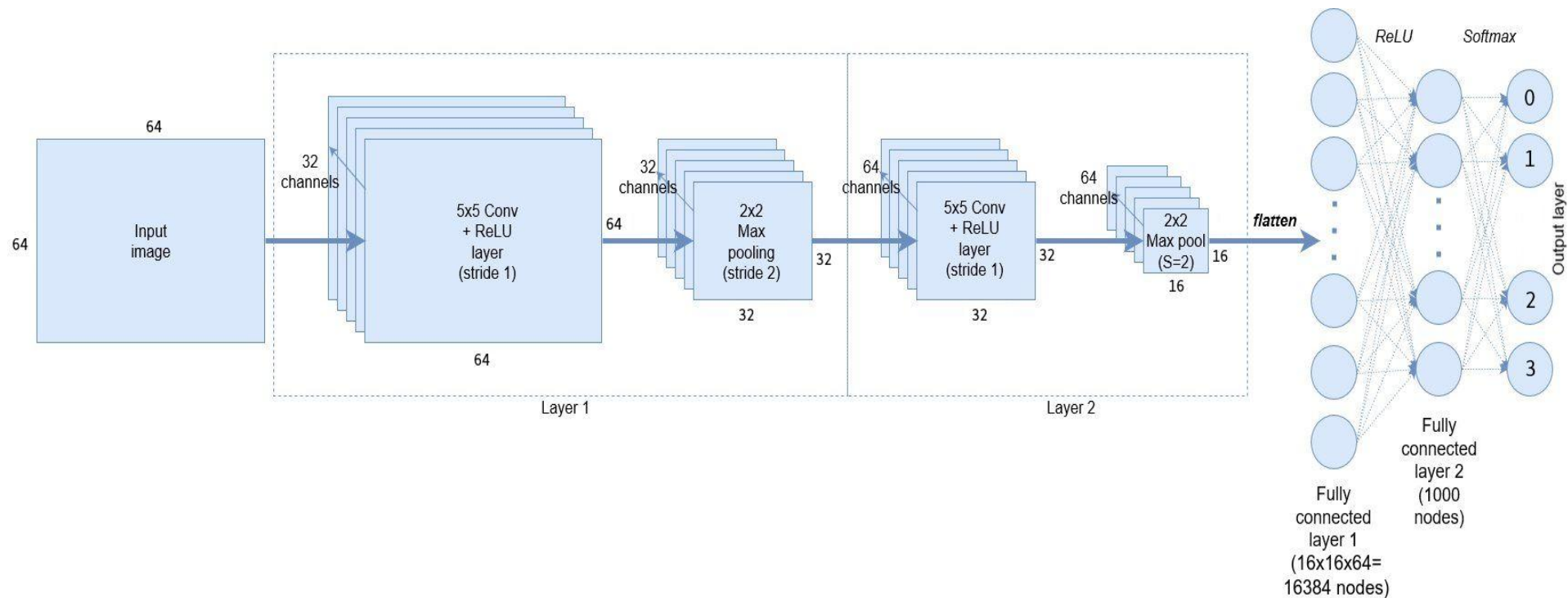
È la funzione con cui si calcola l'errore della rete.

Optimizer:

Fornisce un gradiente per il learning rate.



RETE COMPLETA



COSTRUIAMOLA!

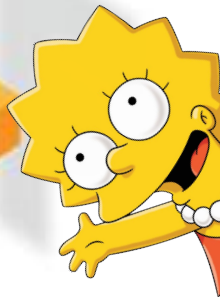


<https://keras.io>



<https://www.tensorflow.org>

K



TensorFlow

```
Epoch 2/5
2018-06-11 14:40:22.658460: I tensorflow/stream_executor/cuda/cuda_gpu_executor.cc:898] successful NUMA node read from SysFS had negative value (-1), but there must be at least one NUMA node, so returning NUMA node zero
2018-06-11 14:40:22.658817: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1344] Found device 0 with properties: Name: GeForce MX150 major: 6 minor: 1 memoryClockRate(GHz): 1.341
pciBusID: 0000:01:00:00
totalMemory: 1.96GiB freeMemory: 1.92GiB
2018-06-11 14:40:22.658836: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1423] Adding visible gpu devices: 0
2018-06-11 14:40:23.591179: I tensorflow/core/common_runtime/gpu/gpu_device.cc:911] Device interconnect bandwidth optimized for device with strength 1 edge matrix:
2018-06-11 14:40:23.591241: I tensorflow/core/common_runtime/gpu/gpu_device.cc:917] 0
2018-06-11 14:40:23.591252: I tensorflow/core/common_runtime/gpu/gpu_device.cc:930] 0: N
2018-06-11 14:40:23.591878: I tensorflow/core/common_runtime/gpu/gpu_device.cc:1041] Created TensorFlow device (/job:localhost/replica:0/task:0/device:GPU:0 with 1684 MB memory) -> physical GPU (device: 0, name: GeForce MX150, pciBusID: 0000:01:00:00, compute capability: 6.1)
37 [=====] - 98s 105ms/step - loss: 0.8271 - acc: 0.6434 - val_loss: 0.8271 - val_acc: 0.7674
2/5
37 [=====] - 97s 104ms/step - loss: 0.5376 - acc: 0.7834 - val_loss: 0.5376 - val_acc: 0.8275
Epoch 3/5
37 [=====] - 91s 98ms/step - loss: 0.4307 - acc: 0.8325 - val_loss: 0.4307 - val_acc: 0.8924
4/5
37 [=====] - 85s 91ms/step - loss: 0.3677 - acc: 0.8744 - val_loss: 0.3677 - val_acc: 0.9062
5/5
37 [=====] - 87s 92ms/step - loss: 0.3320 - acc: 0.8744 - val_loss: 0.3320 - val_acc: 0.9097
dict_keys(['val_loss', 'val_acc', 'loss', 'acc'])
^C
```



DOMANDE?



thanks

`gabriele.baldi.01@gmail.com`

`https://github.com/bnznamco/keras_model_and_slides_notabugconf`

@giurassicparc



@bnznamco

