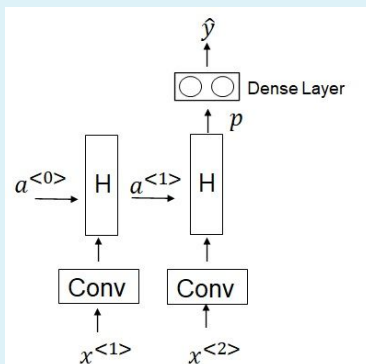


Domanda 1

Risposta non data

Punteggio max.: 1,00

Suppose your input is a 20x20 grayscale image, and you use the following network:



- A Convolutional layer with 1 filters 3x3, option "same"
- A Recurrent Neural Networks with an hidden layer equal to 50 units
- A Dense layer with 2 units
- An output unit with Sigmoid activation function.

How many parameters does this network have (including the bias parameters)?

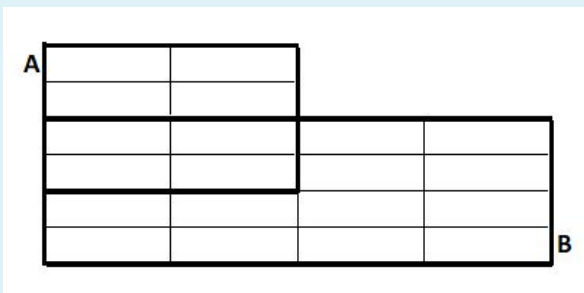
Risposta: ❌

Domanda 2

Risposta non data

Punteggio max.: 1,00

What is the IoU between these two boxes? The upper-left box is 4x2, and the lower-right box is 4x4.



Round the figure to two digits after the decimal point. Use comma as separator.

Risposta: ❌

Domanda **3**

Risposta non data

Punteggio max.: 1,00

Suppose you have an input volume of dimension 32x32x50. Which of the following layer you should use to get an output volume of 16x16x20?

- ☐ A. A 1x1 Convolutional Layer with 20 filter, stride of 2 and no padding
- ☐ B. Maxpooling Layer with filter 2x2, stride of 1 and no padding
- ☐ C. A 1x1 Convolutional Layer with 20 filter, stride of 1 and no padding
- ☐ D. Maxpooling Layer with filter 2x2, stride of 2 and no padding
- ☐ E. A Convolutional Layer with 20 filter 3x3, stride of 1 and no padding
- ☐ F. A Convolutional Layer with 20 filter 3x3, stride of 2 and no padding

Domanda **4**

Risposta non data

Punteggio max.: 1,00

Suppose you are building a Deep Learning system for Image Classification with localization of three classes of objects for images. How many units/neurons in the last layer does this network have?

Risposta:



Domanda **5**

Risposta non data

Punteggio max.: 1,00

Suppose you learn a word embedding for a vocabulary of 10000 words. Then the embedding vectors should be 10000 dimensional, so as to capture the full range of variation and meaning in those words.

- ☐ A. False
- ☐ B. True

Domanda **6**

Risposta non data

Punteggio max.: 1,00

After training a neural network with Batch Norm, at test time, to evaluate the neural network on a new example you should:

- ☐ A. Use the most recent mini-batch's value of μ and σ^2 to perform the needed normalizations.
- ☐ B. Perform the needed normalizations, use μ and σ^2 estimated using an exponentially weighted average across mini-batches seen during training.
- ☐ C. Skip the step where you normalize using and since a single test example cannot be normalized.
- ☐ D. If you implemented Batch Norm on mini-batches of (say) 256 examples, then to evaluate on one test example, duplicate that example 256 times so that you're working with a mini-batch the same size as during training.