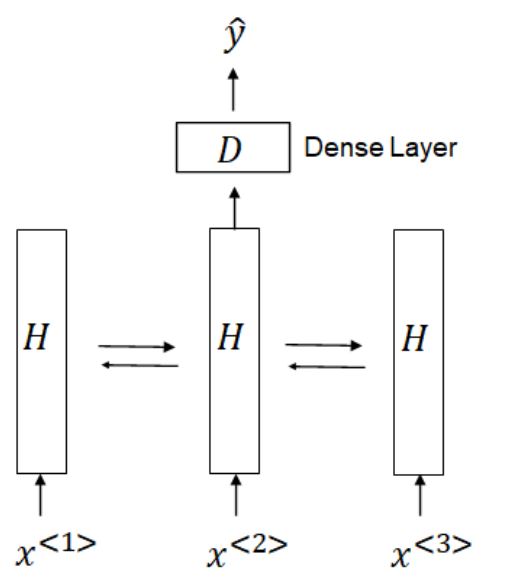
Domanda 1

Risposta non data

Punteggio max.: 1,00

Suppose your input is a $x^{<i>}$ vector with three elements and you use the following Recurrent Neural Network (RNN) for a regression task, with which you want to predict a single value. This Network consists in a Bidirectional RNN with each hidden layer of 5 units, and a Dense layer with 2 units.



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

Remember Bidirectional recurrent neural networks(RNN) are two independent RNNs putting together.

Risposta:

Domanda **2**

Risposta non data

Punteggio max.: 1,00

If your Neural Network model seems overfit, what of the following would be promising thing to try (check all that apply.)

- [ ] a. Make the Neural Network deeper
- [ ] b. Use Data Augmentation
- [ ] c. Get more training data
- [ ] d. Get more test data
- [ ] e. Increase the numer of epochs
- [ ] f. Apply weight decay

Domanda **3**

Risposta non data

Punteggio max.: 1,00

You are building a Deep Learning system based on Computer Vision for recognizing the quality of tomato. The system has to predict 1 when you have a tomato with top quality (y=1) and zero in the opposite case (y=0). Which of the following activation functions would you recommend using for the intermediate layers? (Check all the apply.)

- [ ] A. SoftMax
- [ ] B. LeakyReLU
- [ ] C. Sigmoid
- [ ] D. Relu

Domanda **4**

Risposta non data

Punteggio max.: 1,00

Let's suppose you are using a Convolutional Networks to address a multiclass classification tasks and you are using a Softmax Activation in the last layer.

Considering the following z vector compute the second element of the output vector of the ConvNet:

$$z = \begin{bmatrix} 3 \\ 2 \\ -2 \end{bmatrix}$$

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

Risposta: ✖

Domanda **5**

Risposta non data

Punteggio max.: 1,00

Image scaling can be interpreted as a form of image resampling or image reconstruction. Upsampling to a bigger image from a smaller image can be done with the following techniques: (check all that apply.)

- a. Nearest-neighbor interpolation
- b. Transposed Convolution
- c. Max Pooling
- d. Bilinear algorithm
- e. Convolution
- f. Sigmoid Activation function

Domanda **6**

Risposta non data

Punteggio max.: 1,00

Which ones of the following statements on The Bidirectional Encoder Representations from Transformers (BERT) are true? (Check all the apply.)

- a. BERT consists of an Encoder and a Decoder module
- b. Masked Language Model is one of the techniques used to train the Networks
- c. The size of the input and the output volume of a encoder block is different
- d. The encoder block uses the key concept of the ResNet
- e. The size of the input and the output volume of a encoder block is the same
- f. The training BERT strategy requires a large human annotated dataset

Domanda **7**

Risposta non data

Punteggio max.: 1,00

Let's suppose to have an input volume I (with size 3x3x3), and apply a 1x1 Convolution layer. This layer consists in a 1x1 Convolutivo filter F with padding = 0, stride=1, bias = 2 ans Activation Function=ReLU.

I

C1

5	2	4
3	2	3
6	1	8

C2

4	1	5
5	1	1
3	6	2

C3

1	5	3
7	3	3
7	5	2

C1, C2 and C3 are the channels of the input volume respectively.

F

1	-2	1

Compute the output volume O the value $O_{(2,2)}$.

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

Risposta: ✖