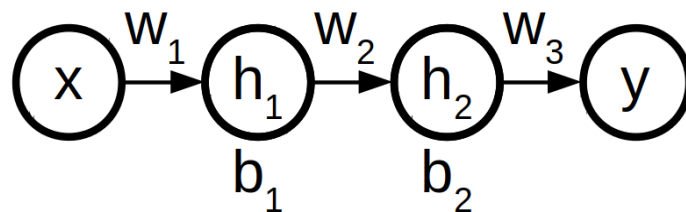


Data Mining – Exam

Deep Learning

1h – Axel Carrier
(documents are allowed)

Question 0 – Neural Network forward pass



A neural network is defined with the following values:

$$h_1 = \text{ReLU}(w_1 * x + b_1)$$

$$h_2 = \text{ReLU}(w_2 * h_1 + b_2)$$

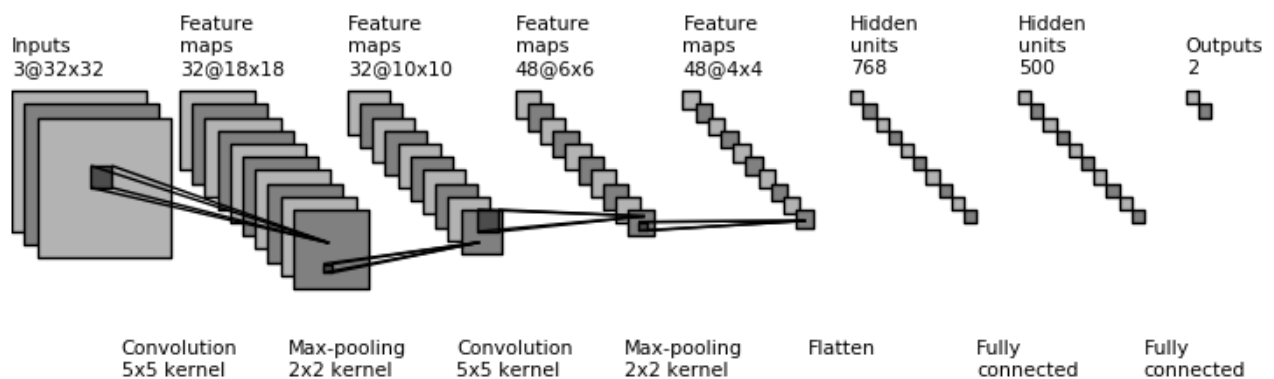
$$y = w_3 * h_2$$

Assuming $b_1 = -2$, $b_2 = 5$, $w_1 = -1$, $w_2 = 3$ and $w_3 = 2$, compute the y values associated to:

- $x_1 = 0$
- $x_2 = -2$
- $x_3 = 3$

Question 1 – Neural Network Analysis

Here is a schematic representation of a particular neural network.



1.1) What is the depth of this neural network?

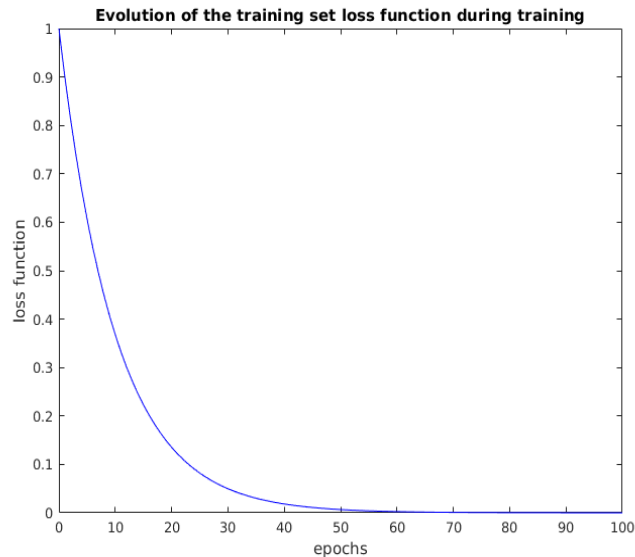
1.2) What type of input do you think this neural network can process?

1.3) What is the number of parameters of this neural network? Please detail your computations.

1.4) What are the reasons why a convolutional neural network might be preferable to a fully connected neural network?

Question 2 - Learning

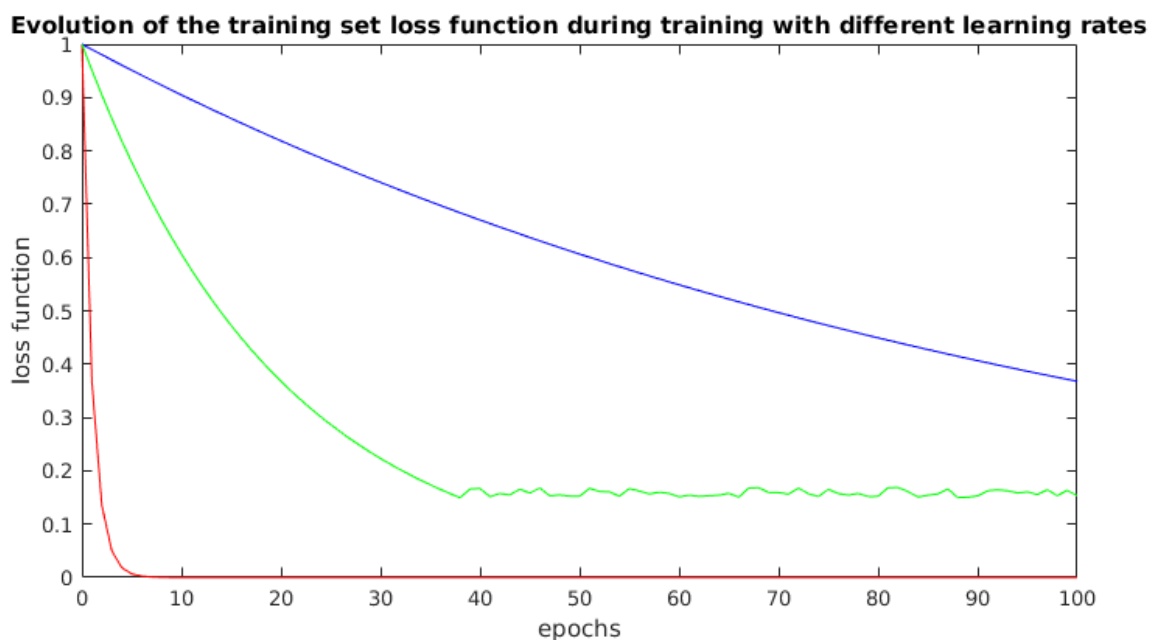
The following graph represent the evolution of the training set loss function while training the neural network introduced in question 2.



2.1) What is the difference between stochastic gradient descent and gradient descent? Which of these two algorithms do you think was used to produce the graph?

2.2) How do you interpret this curve? Do you think the neural network has learned well?

2.3) Explain what is the Expected Prediction Error. Can you guess (and draw on the graph) what is the form of the EPE during training?



2.4) The last graph presents the training set loss function evolution during training, using different learning rates. Which of the curves do you think has the highest learning rate? The lowest? Justify your answer