

1. Explique, usando um fluxograma, o funcionamento de um LLM (Large Language Model) (1 página)

2. (Russell & Norvig, 3rd ed, 2021) (cap. 18)

18.22 Suppose you had a neural network with linear activation functions. That is, for each unit the output is some constant c times the weighted sum of the inputs.

- a. Assume that the network has one hidden layer. For a given assignment to the weights \mathbf{w} , write down equations for the value of the units in the output layer as a function of \mathbf{w} and the input layer \mathbf{x} , without any explicit mention of the output of the hidden layer. Show that there is a network with no hidden units that computes the same function.
- b. Repeat the calculation in part (a), but this time do it for a network with any number of hidden layers.
- c. Suppose a network with one hidden layer and linear activation functions has n input and output nodes and h hidden nodes. What effect does the transformation in part (a) to a network with no hidden layers have on the total number of weights? Discuss in particular the case $h \ll n$.

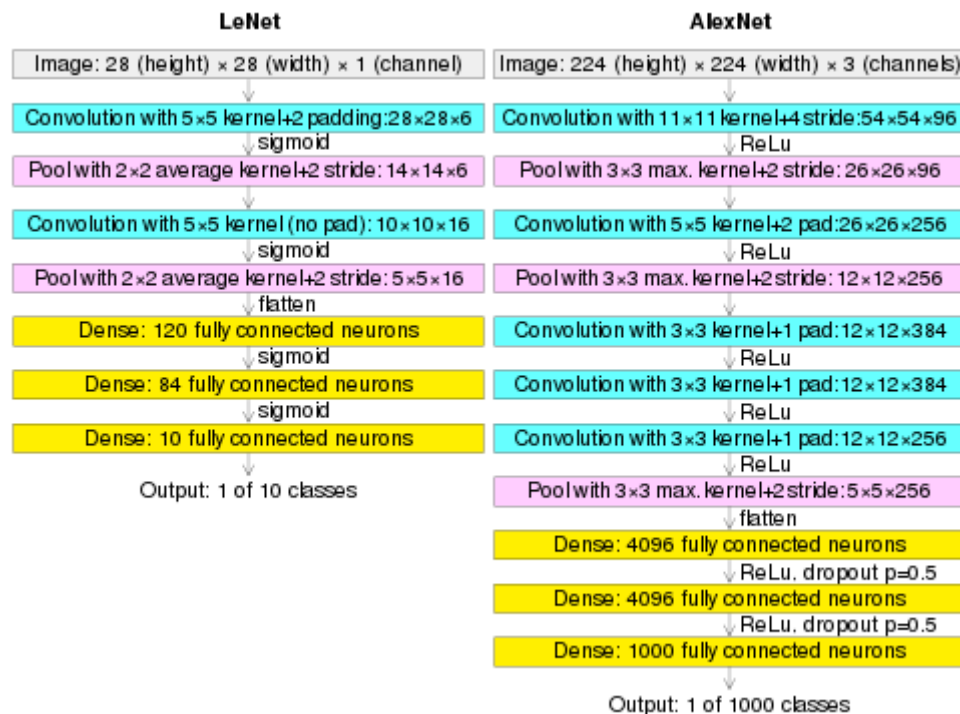
18.23 Suppose that a training set contains only a single example, repeated 100 times. In 80 of the 100 cases, the single output value is 1; in the other 20, it is 0. What will a back-propagation network predict for this example, assuming that it has been trained and reaches a global optimum? (*Hint*: to find the global optimum, differentiate the error function and set it to zero.)

18.24 The neural network whose learning performance is measured in Figure 18.25 has four hidden nodes. This number was chosen somewhat arbitrarily. Use a cross-validation method to find the best number of hidden nodes.

3. Explique modularmente as principais diferenças entre uma MLP (Multi-Layer Perceptron) e uma CNN (Convolutional Neural Network). (1 página)

4. Explique em detalhes o processo de transferência de aprendizagem em uma CNN. (1 página)

5. Descreva em python (com bibliotecas) a construção das 2 CNNs a seguir



6. Explique, e dê exemplos, o que são hiperparâmetros no contexto de deep learning. (½ página)

7. Descreva, e dê exemplos, das tarefas de Visão: (1 página para cada)

- classificação de objetos
- segmentação semântica
- detecção de objetos
- segmentação de instâncias

8. Explique como uma arquitetura Mask R-CNN realiza a segmentação de instâncias em imagens. (1 página)

9. Explique como uma arquitetura YOLO (ps. qualquer versão) realiza detecção de objetos em imagens. (1 página)

10. A figura a seguir mostra a estrutura de um “*transformer*”. Explique seu funcionamento, passo a passo, considerando entradas e saídas textuais (1 página)

