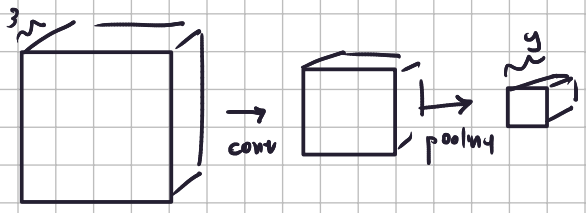
EXERCISE 1

Consider the problem of estimating the age of people from high-resolution color images of faces and assume you have a suitable dataset.

1. Provide the full design of a Convolutional Neural Network (CNN): **completely** describe all the layers (including input and output ones), for each layer, describe its type (convolutional, pooling, fully-connected, etc.) and all the relevant parameters for such a layer (including the activation functions).
2. Explain the output activation function and the loss function needed for training.

We model this as a regression problem (even if the age is in a finite set), the first layer takes as input $128 \times 128 \times 3$ tensor (RGB images). The first convolutional layer have 9 tensors of $65 \times 65 \times 3$, and produces in output a depth map of $64 \times 64 \times 9$, by applying a sigmoid activation and an average pooling of stride 2 we yet a $32 \times 32 \times 9$ tensor.



We apply then a new conv. layer with 20 $32 \times 32 \times 9$ tensor to yet a $1 \times 1 \times 20$ 1D tensor, we then apply a sigmoid act. These 20 units are then used in a fully connected output layer with linear activation: $W^T x + b \Rightarrow$ one real value.

$W \in \mathbb{R}^{20}$, b = bias, tensor $1 \times 1 \times 20 \in \mathbb{R}^{20}$

The loss will be the MSE: $\frac{1}{2} \sum_{i=1}^{N} (t_i - s(x_i))^2$

EXERCISE 2

The manager of a forum about cooking recipes asks you to devise a system to check whether a post in the forum is appropriate You have a set of posts as plain text, each labelled as *appropriate* or *inappropriate*.

POST	APPROPRIATE?
How do you cook spaghetti?	Y
I sell my old car	N
I sell a pan to fry potatoes and vegetables	Y
Visit my new blog about space travels	N
...	...

1. Identify the problem you need to solve, describe a suitable approach, and explain why you preferred it over other alternatives.
2. Describe the application of the chosen approach to the specific problem, showing in particular the model, any kind of data processing or management needed to apply the chosen approach, and the training procedure. Describe how to classify a new post after training.

it is a binary classification problem, we use an approach similar to the spam recognition, where the post are expressed in a vector form, and we use the naive bayes classificator.

each post is a vector of d component, where d is the number of total possible words (vocabulary). If $x \in D$, then $x_i = 1$ if the word i appears in the post. The model classify as follows:

$$c^* = \underset{c \in \{Y, N\}}{\operatorname{argmax}} \; \mathbb{P}(c \mid D) \prod_{i=1}^{d} \mathbb{P}(x_i = 1 \mid c, D)$$

EXERCISE 3

Consider a data set with binary (black & white) images of resolution 32×32, each one obtained by applying one out of 5 predefined linear transformations of a single base image.

1. Explain what is the dimensionality of the data space (i.e., the size of all the possible input images) and what is the intrinsic dimensionality of the given data according to the way in which they are generated.
2. Suppose you apply PCA on a set of such images $D = \{\mathbf{x}_1, \ldots, \mathbf{x}_N\}$ and find that the data can be fully described using M principal components, namely $\mathbf{u}_1, \ldots, \mathbf{u}_M$. Describe **formally** how the original data can be written in the space defined by these M principal components.
3. In this example, is M going to be equal to the number of intrinsic dimensions? Explain why.

the dataspace is in dimension $\{0,1\}^{32\times32}$. the intrinsic dim is 1, since we apply one transformation on a base image.

if $w_1 \cdots w_M$ are the principal component, let $x \in \{0,1\}^{32\times32} \subset \mathbb{R}^{1024}$ to be the vector, we consider the matrix $U = \begin{bmatrix} w_1^T \\ \vdots \\ w_M^T \end{bmatrix} \in Mat(M \times 1024)$. the point on the new space is $Ux \in \mathbb{R}^M$.

M is decided by the designer.