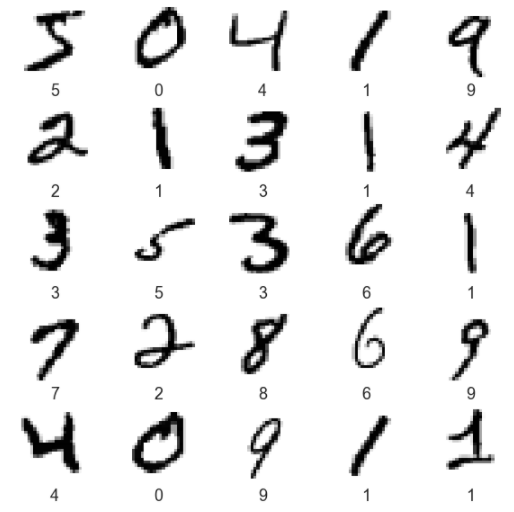
**DAT405: Introduction to data science and AI**

Module 7

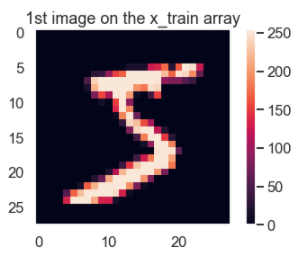


BROU BONI Joël

SANOGO Ibrahim Bechir

# Preprocessing

In this project, we are working with the mnist dataset from the keras package. Mnist is a database of handwritten figures. It is a dataset widely used in machine learning. These are standardized black and white samples (60000), centered with 28 pixels on each side. Therefore, our x\_train has a shape (60000,28,28). In the dataset we can also find and array full of labels that indicate what is the number represented on each picture.



Since data scalling is and important step when working with deep learning learning neural networks it was necessary to do it during the pre-processing. These MNIST samples are represented as an array of numbers whose values range from [0, 255] of type uint8. So in order to scale them between 0 and 1, we divided x\_train by 255 and transformed it into a float32 array.

# Part 3: