Ex.2 – Due end of 6.6

<p style="text-align:center">Subject: Autoencoders (AE)</p>

1. Download data from –
   https://lhncbc.nlm.nih.gov/LHCpublications/pubs/MalariaDatasets.html
   (last paragraph – start with "The datasets are available at.."
   It contains images of parasitized and uninfected cells.
2. Train an Autoencoder on the full data (both parasitized and uninfected). Your latent layer should be of lower dimension than the input. Try different number of layers and layers sizes, until you find an optimal one.
3. After you trained the AE, use only the encoder part (including the bottleneck layer) of it for the next task.
4. Use the encoder-network on 200 (100 of each class) samples and collect the embedded vectors (whose dimension is the size of the bottleneck layer, e.g. of dimension M). Next, use t-SNE algorithm to project these M-dimensional vectors to a 2dim space and observe how well the two groups are separated by coloring each group members in a different color.
5. Next, our goal is to build a classifier that is capable of discriminating between the two types of cells.
6. By using the encoder (trained) part of the net, add a fully connected layers (try different numbers) and find optimal number of layers and number of neurons, and classification layers. Train your network, and present its test scores by using an appropriate loss function.

   Good luck!