National University of Sciences and Technology

School of Electrical Engineering and Computer Science

Department of Computing

CS 405 Deep Learning

Fall 2023

# 

# Lab 13

**Auto-encoders in PyTorch**

**Name: Alina Nasir**

**Qalam ID: 342350**

Date: 21st Dec 2023

**Dataset:**

The dataset used in this lab is CIFAR-10 dataset that contains the images of the real world objects whereas the MNIST dataset contained images of the numbers. Moreover, the MNIST dataset contains the images of grayscale whereas CIFAR-10 dataset contains the images based on RGB channel. Based on these differences the network architecture and the viewer functions are modified.

**Network Architecture:**

The network input was changed from 1 to 3 to include the three channels of RGB in the images. The output of the decoder was then modified accordingly. Moreover, in the decode function the **output\_size** parameter is not explicitly specified. The spatial size of the output tensor is determined based on the default behavior of the transposed convolution layer. The default behavior calculates the output size based on the input size, stride, and padding.

The output size formula for a transposed convolution is given by:

output size= (input size−1) × (stride−2) × padding + kernel size + output padding

Here, **input size** is the size of the input tensor, **stride** is the stride of the transposed convolution, **padding** is the padding added to the input, **kernel size** is the size of the convolutional kernel, and **output padding** is the additional padding added to the output.

The spatial size calculated using this formula is determined to cover the entire input spatial extent.

The output function is also changed to tanh inorder to normalize the input between -1 and 1.

**Viewer Functions:**

The viewer function is changed to use np.transpose instead of np.squeeze for making the image data compatible with plt.imshow. Transpose can be used to permute the dimensions of an array, effectively changing the layout of the data. This helped to rearrange the dimensions for compatibility.