# Artificial Neural Networks

2nd Assignment - Shahid Beheshti University - Master's Program

March 20, 2023

**Due date: April 7**

## Exercise 1

What is the difference between L1 and L2 regularization? Write their formulas and compare them with each other.

## Exercise 2

What is exploding and vanishing gradients in neural networks? How can Adam optimizer help us with them? Is LeakyReLU better or ReLU in avoiding vanishing gradients?

## Exercise 3

What is BCE loss? When should we use BCE and when should we use MSE?

## Exercise 4

How can we avoid overfitting? Name 3 methods and explain them in detail.

## Exercise 5

Does dropout slow down training? Does it slow down inference (i.e., making predictions on new instances)? What about MC dropout?

## Exercise 6

What is the problem that Glorot initialization and He initialization aim to fix?

## Exercise 7

Using FashionMNIST data, a sample dataset is created(**Link**) with all of the pixels in the center column of the photos set to zero. Also, their real values are extracted and saved in a CSV file for each image. It is expected that you:

a. Split the data into train and test sets.
b. Implement an MLP in order to predict the missing values in the images.
c. Report the accuracy of the model on the test set and visualize the final images predicted by the model.
d. Utilizing various enhancing techniques try to boost the performance of the model including:
    i. Batch Normalization layers
    ii. Dropout layers
    iii. Different activation functions and comparing their performance
    iv. Learning rate scheduling
    v. L1, L2 Regularization
    vi. Different Weight initialization (*EXTRA PONIT*)
    vii. Early stopping (*EXTRA PONIT*)
    viii. Utilizing different optimizers and comparing their performance (*EXTRA PONIT*)