

Marked exercises, lecture 4

This is the third set of marked exercises, which are a part of Submission 1. In this assignment, you will get familiar with initialisation, knowledge distillation and quantization of the model. Detailed instructions and explanations are provided in the accompanying Jupyter notebook. In total, you can get 20 points out of 60 points for Submission 1 for completing all exercises related to lecture 4.

1.4.1 [5 points] Implement a small student network that is later going to be trained for a 10-class classification. Its weights should weight $\geq 2\text{MB}$ and $\leq 4\text{MB}$

1.4.2 [6 points] Initialisation of the model. Think which is the best initialisation for each layer/ type of layer and implement the corresponding initialisation. You need to initialise every trainable parameter of the network with your code.

1.4.3 [7 points] Implement training with knowledge distillation.

1.4.4 [2 points] Apply the quantization to the student model.