Homework 4

2-AIN-150, Winter 2018

Deadline: 19.11.2018, 23:59, moodle.uniba.sk/fmph

Before you start solving the homework, please read the general instruction at the end of the document. Submitted solutions should be your own. Do not copy and do not tried to find solution in literature or over the internet.

Optimalization of hyperparameters

Read carefully what you should submit.

Consider slightly modified neural network from third tutorial running over Fashion MNIST.

Our network will have several hidden layers and dropout layer between each hidden layer. (Dropout layer just randomly discard some neurons during training, which reduces the overfitting. Amount of discarded neurons is set by hyperparameter.)

But we need to set network hyperparameters first. We will set: Number and size of hidden layers (each one will be with same size), dropout rate and number of training epochs. We want to choose hyperparameters, which give lowest validation error.

Your task is to implement and compare two approaches for setting the hyperparameters.

First one is random search. It is just randomly picks values of hyperparameters, runs several times and chooses the best setting.

Second one is grid search. It selects several values for each parameter (sometimes we use linear scale, sometimes logaritmic) and then it tries all possible combinations. In some cases it even zooms in (it does another iteration around best solution), but you do not need to implemente this.

Implement both approaches and compare them. Which one gives to good enough result faster?

You should submit a notebook from tutorials, where you just add grid search and random search to the end. Also you should run them and add short commentary (and maybe some charts) about which method seems to be better. Note: You are forbided to call libraries, which do grid search and random search for you.

You are also given a file `architecture.py` which just describe architecture of the network and also gives you range of parameters you should consider.