Practical work 03 – 07/03/2022
Shallow Networks

Objectives

The main objectives of this Practical Work for Week 3 are the following :

a) Implement MBGD and Softmax and learn what it means to choose hyper-parameters such as learning rate, batch size or number of epochs.

b) Further deepen your skills in python and numpy.

Submission

— **Deadline** : Tuesday 14 March, 3pm

— **Format** :

  — Exercise 1 (MBGD and Softmax)

    — Jupyter notebook `MNIST_softmax_classifier_stud.ipynb` completed with your solution and comments.

  Submission of all files in a single zip-file using the naming convention (for team of two students #1, #2) :

  `family name_given name #1- family name_given name #2.zip`

Exercise 1 MBGD and Softmax

Implement Mini-Batch Gradient Descent for Softmax. Do this on the basis of the Jupyter notebook `MNIST_softmax_classifier_stud.ipynb`. As in PW 02, do this by only using numpy functionality (scikit learn used only for loading the data and splitting it into train and test sets). Look out for the suitably marked sections that you need to implement.

Note that you will train softmax for the original MNIST. Since this will take more CPU and RAM, you need to be more careful in efficiently implementing the code. Make sure to properly use numpy array arithmetics! All the training runs should complete in at most a couple of minutes.
For debugging purposes you also have the posibility to use MNIST light. Set `mnist = 0` in Cell number 2.

Proceed as follows :

(a) Implement the update rules for a softmax layer when using MBGD. As in PW 02, keep an eye on the shapes of the numpy arrays defined in the input and to be provided as output.
You can use the unit tests at the end of the notebook to test your implementation of the individual methods.

(b) Run the training by using MBGD and cross-entropy cost and plot the learning curves : Cost, Error Rate - both for the training dataset and the test dataset. Play with the hyper-parameters

- learning rate
- number of epochs
- batch size

Specify your choice of these parameters and justify why your choice is best suited.