Assignment 1 - Multi Layer Perceptron + Backpropagation

Advanced Topics in Neural Networks

Due date: 14-10-2024

Homework - 20 points

Implement a Multi Layer Perceptron (MLP) and the Backpropagation algorithm using raw PyTorch Tensor operations. Avoid using Autograd and APIs such as **torch.nn** or **torch.functional**. You are also not allowed to use neither **torch.data** nor **torchvision.transforms**. You can use torchvision.datasets.MNIST to load the MNIST data.

The MLP architecture should consist of 784 input neurons, 100 hidden neurons, and 10 output neurons. Use the MNIST dataset to evaluate your implementation. Measure the loss and the accuracy for both training and validation. For loss calculation, you are allowed to use torch.nn.functional.cross_entropy(). The tasks for this assignment and their respective points are as follows:

1. Implement forward propagation. **(2 points)**
2. Implement the backpropagation algorithm using the chain rule. Refer to the provided resources. **(5 points implementation + 5 points explanation)**
3. Utilize batched operations for enhanced efficiency. **(4 points)**
4. You must achieve at least 95% accuracy on the validation split of the MNIST dataset. Ensure reproducibility on Google Colab, in under 5 minutes of runtime. **(4 points)**
   - Things you may try to improve the accuracy:
     (a) Weight initialization
     (b) Replacing Sigmoid with other activation functions
     (c) Changing the learning rate
     (d) Augmenting the training dataset
   - More advanced things you can try: Dropout, Batch Normalization, Learning Rate Decay, Regularization, Weight Decay, Gradient Clipping.

- **Attention**, you are allowed to use only raw tensor operations, without using Data Augmentation libraries or high level APIs like torch.nn and torch.functional.

Try to run your model both on CPU and GPU and check the runtime. If you do not have a GPU available, use Google Colab and copy your script in a notebook cell (your script should be able to run on both CPU and GPU depending on a device parameter passed from the outside).

Bonus: Strive for higher accuracy (not graded, but always try to be better).