Hands-on session 2

Feed forward neural network

Instructor: Dr. Souvik Chakraborty

APL 745: Deep Learning for Mechanics

January 24, 2024

Submission due on February 07, 2024

Instructions:

(i) You are allowed to discuss in a group; however, you must submit your own handwritten homework copy (no computer-typed submission will be accepted). Further, copying homework from your friends is forbidden. If found copied, the homework submission will be considered invalid for all the students involved and will be graded zero.

(ii) Write all the steps, including your reasoning and the formulae you referred to, if any. If sufficient reasoning is not provided and steps are unclear, step marks will not be given.

(iii) For practical submissions, the codes are accepted in *.ipynb* or *.py* format.

(iv) Unless mentioned otherwise, only *numpy*, *pytorch* and *matplotlib* libraries may be used. Direct commands for algorithms to be implemented are not allowed.

(v) The *.rar* file containing all submission related files shall be named in the format, **Name_Entrynumber.rar**

Question 1.

Prerequisites:

1) Download MNIST dataset using *torchvision*. The documentation for the same can be found in this link.
2) Transform the labels provided in the dataset using one hot encoding.
3) Create three datasets with following specifications,
   - Training dataset: 5000 samples
   - Validation dataset: 1000 samples
   - Testing dataset: 2000 samples

Networks:

1) Network **A**: Train a network with 2 hidden dense layers. Each hidden layer must have 50 nodes and ReLU activation. The output layer should have softmax activation.
2) Network **B**: Train a network with 2 hidden dense layers. Each hidden layer must have 150 nodes and ReLU activation. The output layer should have softmax activation.

Tasks:

1) Train network **A** for 500 epochs. Use ADAM optimizer with learning rate of 0.001 and the datasets are NOT to be normalized.
2) Train network **A** for 500 epochs. Use ADAM optimizer with learning rate of 0.001 and normalize the datasets. Compare results with Task 1 ans report your findings.
3) Add a batch normalization layer after each hidden layer of network **A** and keep other specifications same as previous task. Compare the results with previous task and report your findings.
4) Add a dropout layer between the two layers of network **A** and keep other specifications same as the 2^{nd} task. Compare the results with those obtained in 2^{nd} task and report your findings.
5) Train network **B** for 500 epochs using ADAM optimizer and a learning rate of 0.01.
6) Train network **B** for 500 epochs using ADAM optimizer and a learning rate of 0.001. Compare the results with previous task and report your findings.

Additional notes:

- Pytorch may be used to perform the above tasks.
- Accuracy obtained for all the datasets must be reported. For training and validation datasets, accuracy obtained at only the final epoch may be reported.
- Cross entropy loss is to be used while training and a plot showing the epoch vs training loss must be plotted.
- The number of nodes in the input and output layers must be selected according to the dataset under consideration.