EEL6935 Spring 2017 – Project 2

April 6, 2017

Due: April 26, 2017, 11:59 PM

This project will be graded with a letter grade with respect to presentation (25%), methods (35%) and results (40%).

In this project, you should train a Multilayer Network (MLP) and a Convolutional Neural Network (CNN) on the MNIST dataset. The MNIST dataset (28×28) contains 60 k images for training and 10 k images for testing. The last 10 k images in the training dataset are used as validation dataset for cross-validation, early stopping and hyper parameter choices.

1. Download the MNIST dataset and code from the website. (courtesy of ufdl.stanford.edu)
2. Load images and labels using *readMNIST* script provide on the website. To read the training images and labels you can use

   [trainImg,trainLbl]=readMNIST('train-images.idx3-ubyte','train-labels.idx1-ubyte',60000,0).

   To read the test images and labels you can use

   [testImg,testLbl]=readMNIST('t10k-images.idx3-ubyte','t10k-labels.idx1-ubyte',10000,0).
3. Train MLPs and CNNs for classification. Use the Softmax (Categorical Cross-entropy) cost function discussed in class. For the MLP, downsample the image to decrease the number of inputs or use PCA to project to a subspace.
4. Use samples 50000 to 60000 from the training dataset as the validation dataset for early stopping.
5. Report results on the test set.
6. Show learning curves comparing training and validation cost function. Do the same for the accuracy.
7. Discuss differences in results between MLP and CNN.
8. Compare performance with SGD and SGD-momentum and RMS Prop.

Hints

- You should be getting better results with CNN.
- Use initialization techniques explained in class for better results.

- Use ReLU activation function for better results and faster convergence.
- Use dropout regularization with probability 0.5 for better results on the test set.

The project requires a report explaining the experimental procedures you followed and you must include detection data, figures, and tables to support your conclusions. Please use the format of an IEEE Transactions paper (limited to 7 double column pages). This means you have to write a brief intro to the theory, explain well the methods and present carefully the results (see below) and conclusions. Remember that any scientific paper should, by definition, contain sufficient information such others can replicate your results. A scientific paper must also contain ORIGINAL material only. If you happen to use text or equations from other source you have to reference what you cut and paste (this is not allowed in a normal publication, but here it is OK provide you reference). Of course, we expect the results to be done by the student alone.